CONSERVATION STRATEGIES FOR PORTUGUESE CROP LANDRACE DIVERSITY

By

MARIA JOÃO SANTOS DE ALMEIDA

A thesis submitted to the University of Birmingham for the degree of
DOCTOR OF PHILOSOFY

School of Biosciences
College of Life and Environmental Sciences
The University of Birmingham

May 2018

ABSTRACT

Crop landraces (LR) are threatened generally due to socio-economic and climate changing factors. However, LR often present traits able to mitigate the aforementioned threats.

Several countries ratified treaties that promote conservation and sustainable use of their LR and documents, such as National Strategic Action Plans (NSAP), are outlined to coordinate such conservation actions. Nevertheless, in most countries this is rather incipient.

With this work a Portuguese LR inventory was created with 14,813 LR entries corresponding to 7,492 different LR within 36 families and 130 taxa of 123 crops. In addition, a LR threat assessment methodology was developed including a *LR Threat Risk Calculator* which enables automatic calculation of LR threat risk levels. Finally, genetic work using 11 microsatellites was carried out for *Phaseolus vulgaris* L. LR in order to assess its genetic diversity and providing heterozygosity information to be used in threat assessment.

Nei's genetic distance calculation identified, out of three populations, population two and three LR as more closely related to each other. Population two with predominantly mainland and Madeira LR, and population three with the fewer number of LR, mainly from Azores.

ACKNOWLEDGMENTS

I would like to thank to Dr. Nigel Maxted and Dr. Joana Magos Brehm for the provided help and all the positive comments and talks throughout the work, and to Eng. Ana Maria Barata also for the support and particularly for allowing me to use the Banco Português de Germoplasma Vegetal (BPGV) facilities to continue part of this work. To Dr. Lyndsey Leach, thank you for the talks and for the supportive feedback regarding chapter IV. To Dr. Iain Johnston, for all the discussions and for the feedback concerning chapter III. Thank you.

My thank and deep appreciation also goes to all Agricultural departments of Azores archipelago, to all head of departments and highly qualified and motivated personnel that accompanied me during glorious field work:

Fernando Sousa (Regional Director of Agriculture)

Carlos Eduardo Costa Santos (Director of the Direction of Agricultural Services – S. Miguel)

Ana Catarina Faria Silva Durão (Direction of Agricultural Services – S. Miguel)

Hélder Rui Machado Amaral (Direction of Agricultural Services – S. Miguel)

José Adriano Mota (Direction of Agricultural Services – S. Miguel)

António Quintanilha (Agricultural Development Service of S. Miguel)

Fábio Carvalho (Agricultural Development Service of S. Miguel)

Janyne Sousa (Director of the Agricultural Development Service of Sta. Maria)

Cláudia Monteiro (Agricultural Development Service of Sta. Maria)

Jorge Martins (Agricultural Development Service of Terceira – Chief of Division)

Izabel Armas (Agricultural Development Service of Terceira)

Hélio Avla (Agricultural Development Service of Terceira)

Ana Catarina Coimbra (Director of the Agricultural Development Service of Graciosa)

José Orlando Santos (Agricultural Development Service of Graciosa)

Vagner Paulos (Director of the Agricultural Development Service of Pico)

Raúl Jorge (Agricultural Development Service of Pico)

João Santos (Director of the Agricultural Development Service of S. Jorge)

Jorge Azevedo (Agricultural Development Service of S. Jorge)

Tieres Vieira (Director of the Agricultural Development Service of Faial)

Ângelo Duarte (Agricultural Development Service of Faial)

Rigoberto Gomes (Diretor of the Agricultural Development Service of Flores)

Ivan Castro (Agricultural Development Service of Flores)

José Dias (Agricultural Development Service of Flores)

Valdemar Freitas (Agricultural Development Service of Flores)

Ashley Domingos (Agricultural Development Service of Corvo)

José Jorge Vieira (Agricultural Development Service of Corvo)

To ISOPlexis researchers that made me feel very welcome and guide me in ISOPlexis facilities as also helped in the collecting mission:

Dr. Miguel Carvalho (ISOPlexis coordinator)

Gregório Freitas (researcher)

Humberto Nóbrega (researcher)

Sara Castro (researcher)

To the hardworking and knowledgeable farmers that kindly offered their time and diligently keep landraces preserving our heritage alive.

To my friends: thank you for your kind words of encouragement. To all my office partners, particularly to Aremi Contreras, thank you for making this journey lighter and happier with your positivity and friendship.

Finally, to my family, my dearest Mum, Dad, Miguel, Tim and Stu… Your love and unlimited support (and gigantic patience) gave me the strength needed.

Thank you.

TABLE OF CONTENTS

Chapter I. GENERAL INTRODUCTION 1

1.1 Biodiversity and Agrobiodiversity 2

1.2 The definition of Landrace 3

1.3 Nomenclature and other proxy definition 4

1.4 The value of Landraces 5

1.5 Threats to Landraces 6

1.6 National strategic action plan for Landrace conservation 7

1.7 Landrace Conservation approaches 10

1.7.1 *In situ* Conservation 10

1.7.2 *Ex situ* Conservation 11

1.7.3 A Complementary Conservation 12

1.8 Diversity conservation and use of Landraces in Portugal 12

1.9 Aim and Objectives 17

1.10 Work programme and outcomes 18

Chapter II. PORTUGUESE INVENTORY OF FOOD AND OTHER AGRICULTURAL CROP LANDRACES 20

2.1 Abstract 21

2.2 Introduction 22

2.3 Materials and Methods 29

2.3.1 LR working definition 29

2.3.2 LR survey and data collection ... 29

2.3.3 Selection of sites ... 30

2.3.4 Field work general results ... 33

2.4 Results ... 33

2.5 Discussion ... 43

2.6 Conclusion ... 44

Chapter III. A THREAT ASSESSMENT METHODOLOGY FOR CROP LANDRACES ... 46

3.1 Abstract ... 47

3.2 Introduction ... 48

3.3 Materials and Methods ... 51

3.3.1 Description of the general methodology – LR assessment process ... 51

3.3.2 Target questions and questionnaire ... 55

3.3.3 Criteria, sub-criteria and indicators of threat - description ... 59

3.3.4 Criteria, sub-criteria and indicators of threat - organization ... 63

3.3.5 Description of the methodology basis ... 65

3.3.6 Application to a case study to test the method – Common bean LR ... 71

3.3.7 Comparing methodologies ... 74

3.4 Results ... 75

3.5 Discussion ... 77

3.6 Conclusion ... 80

Chapter IV. ASSESSING THE GENETIC DIVERSITY OF A GROUP OF PORTUGUESE COMMON BEAN (*P. VULGARIS* L.) LANDRACES USING MICROSATELLITES 81
4.1 Abstract 82
4.2 Introduction 83
4.3 Material and Methods 85
4.3.1 Plant material 85
4.3.2 DNA extraction 88
4.3.3 PCR conditions 89
4.3.4 Primers selection and testing 89
4.3.5 Fragment analysis 91
4.3.6 Data analysis 91
4.3.7 Description of the script 92
4.4 Results 94
4.5 Discussion 106
4.6 Conclusion 107
Chapter V. GENERAL DISCUSSION 109
5.1 Summary 110
5.2 Inventory 110
5.3 Threat Assessment 111
5.4 Molecular work 112
5.5 Recommendations 113

5.6 Future work proposals 114

5.7 Conclusion 115

References 118

Supplementary Material 133

LIST OF FIGURES

Figure 1.1.: Model for a National Strategic Action Plan for LR Conservation (from Maxted *et al.*, 2013)……………………………………………………………………………………9

Figure 1.2.: Traditional banana varieties – Madeira island…………………………………...13

Figure 1.3.: Various traditional fruits – Madeira island……………………………………….13

Figure 1.4.: Traditional tomatoes and potatoes – Madeira island……………………………..13

Figure 1.5.: Various traditional fruits – Madeira island……………………………………….14

Figure 1.6.: Traditional common bean varieties - Madeira island…………………………….14

Figure 1.7.: Typical small size fields in northern continental Portugal………………………..14

Figure 1.8.: Farmer in backyard garden in northern continental Portugal…………………….15

Figure 1.9.: Farm in southern continental Portugal…………………………………………...15

Figure 1.10.: Backyard garden with common bean – S. Miguel, Azores archipelago………...15

Figure 2.1.: Map of Portugal (from Encyclopaedia Britannica online, published by Encyclopaedia Britannica, https://www.britannica.com/place/Portugal/images-videos/media/471439/61694, accessed 20/09/2018)……………………………………….23

Figure 2.2.: Portugal NUTS II (adapted from image created by Rei-artur with the permission under the terms of the GNU Free Documentation License, https://commons.wikimedia.org/wiki/File:Portugal_NUTS_II.svg, accessed 20/09/2008)..…24

Figure 2.3.: ELC map with collecting points………………………………………………….32

Figure 2.4.: Distribution of landraces in the Portuguese territory according with the collated data………………………………………………………………………………………….40

Figure 2.5.: Distribution of the common top crop in Portuguese territory…………………….41

Figure 2.6.: Distribution of landraces per crop group (crop group according to FAO, 2005)…41

Figure 2.7.: Distribution of landraces per crop group in the Portuguese territory……………..42

Figure 3.1.: LR threat assessment methodology summary……………………………………52

Figure 3.2.: Process of assessing a LR threat risk level……………………………………….53

Figure 3.3.: Input answers process into *LR Threat Risk Calculator*. Direct input when there is only one LR (one population – LR A) and input of the calculated averages when there are several populations of the same LR (LR B1 and LR B2)……………………………………....53

Figure 3.4.: Questionnaire applied to farmers during field work……………………………...56

Figure 3.5.: Example of synopsis obtained with the *LR Threat Risk Calculator*……………...58

Figure 3.6.: Histogram of threat risk values from the random sampling process described in the text…………………………………………………………………………………………67

Figure 4.1.: Germination trays…………………………………………………………………..86

Figure 4.2.: Detail of labelled germination trays under shade screen…………………………86

Figure 4.3.: Collected leaves drying in silica gel……………………………………………...87

Figure 4.4.: BPGV laboratory where leaves were weighed…………………………………...88

Figure 4.5.: Sampled material in collection microtubes………………………………………88

Figure 4.6.: Allele sharing per loci per population graph……………………………………..95

Figure 4.7.: Allele patterns across the studied populations…………………………………...98

Figure 4.8.: Principal coordinate analysis (PCoA) of the 3 populations, developed using Nei's genetic distance values matrix…………………………………………………………….101

Figure 4.9a.: UPGMA dendrogram based on Nei's genetic distance for the 274 studied samples (see CD at the end of this thesis). Different populations highlighted with different colours……………………………………………………………………………..………...103

Figure 4.9b.: UPGMA dendrogram based on Nei's genetic distance for the 274 studied samples (see CD at the end of this thesis). Different regions' samples highlighted with different colours…………………………………………………………………………...………..104

LIST OF TABLES

Table 2.1.: Main crops of the Portuguese territory (adapted from INE, 2017)…………………...27

Table 2.2.: Number of landraces per family……………………………………………………..35

Table 2.3.: Number of landraces per genus……………………………………………………...36

Table 2.4.: Number of landraces per taxa (species and/or infra-specific taxa level)…………..37

Table 2.5.: Number of landraces that have only been classified at genus level………………..39

Table 2.9.: Number of landraces per crop group…………………………………………….....42

Table 3.1.: Calculations necessary to be performed, for the use of *LR Threat Risk Calculator*.53

Table 3.2.: Target Questions – specific questions to be used in questionnaires and in every collecting mission. ……………………………………………………………………………57

Table 3.3.: List of criteria, sub-criteria and indicators of threat………………………………64

Table 3.4.: Criteria, sub-criteria, indicators of threat and calculations example for *Amarelo* bean LR (ID 20, 21) and corresponding result on *LR Threat Risk Calculator*………………………73

Table 3.5.: Indicators, risk level and associated score adopted by Porfiri *et al*. (2009), to assess LR genetic erosion from Lazio region……………………………………………………...74

Table 3.6.: Assessment of threat risk level of the selected LR using the *LR Threat Risk Calculator*……………………………………………………………………………………76

Table 3.7.: Assessment of threat risk level of the selected LR using Porfiri *et al*. (2009) method…………………………………………………………………………………...77

Table 4.1.: Label, multiplex group, linkage group (LG), forward and reverse sequence and product size of the selected markers……………………………………………………...90

Table 4.2.: Populations' size……………………………………………………………………94

Table 4.6.: Polymorphic information content (PIC) per locus per population (1, 2 and 3) and F-statistics and total number (n.) of alleles per locus…………………………………………96

Table 4.7.: Mean allelic patterns and standard error (SE) across studied populations. Total number of alleles and percentage of polymorphic loci per population………………………97

Table 4.9.: Pairwise population F_{ST} values…………………………………………………..100

Table 4.10.: Pairwise population matrix of Nei Genetic Distance…………………………...100

SUPPLEMENTARY MATERIAL

Supplementary Figure 2.1.: Questionnaire applied to farmers during field work…………..133

Supplementary Figure 4.10a.: UPGMA dendrogram based on Nei's genetic distance for the 274 studied samples (272 *P. vulgaris* L. and 2 *P. coccineus* L.) with external LR in red and commercial varieties in blue (see CD at the end of this thesis)……………………………….134

Supplementary Figure 4.10b.: UPGMA dendrogram based on Nei's genetic distance for the 274 studied samples (272 *P. vulgaris* L. and 2 *P. coccineus* L.) with duplicates samples in red (see CD at the end of this thesis)…………………………………………………………..135

Supplementary Table 2.6.: Number of landraces per taxa per district and per island in the Azores and Madeira archipelagos……………………………………………………………136

Supplementary Table 2.7.: Number of landraces per taxa per undefined districts………….147

Supplementary Table 2.8.: Number of landraces per crop group per district……………….147

Supplementary Table 4.3.: Allele frequencies, sample size (N) and total number of alleles per locus (TL) and per populations (TP)…………………………………………………………151

Supplementary Table 4.4.: Number of samples per allele per population and number of alleles per locus per population………………………………………………………………...154

Supplementary Table 4.5.: Summary of private alleles per population…………………….157

Supplementary Table 4.8.: Sample Size (N), No. Alleles (Na), No. Effective Alleles (Ne), Information Index (I), Observed Heterozygosity (Ho), Expected and Unbiased Expected Heterozygosity (He, uHe), and Fixation Index (F)…………………………………………...158

LIST OF ABBREVIATIONS AND ACRONYMS

BCBA	Banco do Centro de Biotecnologia dos Açores
BPGV	Banco Português de Germoplasma Vegetal
CIAT	Centro Internacional de Agricultura Tropical
CWR	Crop Wild Relatives
DRAPAlg	Direção Regional de Agricultura e Pescas do Algarve
DRAPC	Direção Regional de Agricultura e Pescas do Centro
DRAPN	Direção Regional de Agricultura e Pescas do Norte
ELC	Ecogeographic Land Characterization
FAO	Food and Agriculture Organization of the United Nations
GIS	Geographic Information System
INE	Instituto Nacional de Estatística
INIAV	Instituto Nacional de Investigação Agrária e Veterinária
ISOPlexis	Banco de Germoplasma da Madeira
ITPGRFA	International Treaty on Plant Genetic Resources for Food and Agriculture
IUCN	International Union for the Conservation of Nature
LR	Landraces
NGOs	Non-governmental organizations
NSAP	National Strategic Action Plan
NUTSII	Nomenclature of Territorial Units for Statistics in a regional level
PCoA	Principal Coordinate Analysis
PDO	Protected Designation of Origin
PGI	Protected Geographical Indication
PGR	Plant Genetic Resources

PGRFA	Plant Genetic Resources for Food and Agriculture
PIC	Polymorphic Information Content
PPB	Participatory Plant Breeding
RF	Random Forest Method
SE	Standard Error
UAA	Utilized Agricultural Area
UPGMA	Unweighted Pair Group Method with Arithmetic Mean
USDA, ARS, GRIN	U.S. National Plant Germplasm System platform

CHAPTER I. GENERAL INTRODUCTION

1.1 BIODIVERSITY AND AGROBIODIVERSITY

Biodiversity can be defined as the diversity within and between animals, plants, fungi, prokaryotes and protists, and their environments (CBD, 1992) while the subcomponent used in agriculture and as food is referred as agrobiodiversity (FAO, 2005a). Agrobiodiversity is formed by species and varieties used for human and animal consumption, as power source, fibres and in pharmaceutical industry. It also encompasses species that assists food production (e.g. pollinators) and species that underpins production systems as all agricultural ecosystems (FAO, 2005a). Within agrobiodiversity, plant genetic resources for food and agriculture (PGRFA) are *"any genetic material of plant origin of actual or potential value for food and agriculture"* (FAO, 2009). It includes modern and outdated cultivars, breeding lines, crop wild relatives (CWR) and landraces (LR) (Maxted *et al.*, 2011). LR have value to farmers that use them, benefiting of their organoleptic properties and possible income from cultivation (e.g. Brush, 1992; Negri, 2003; Torricelli & Negri, 2015; Veloso *et al.*, 2015a, b) and value as potential source of genes for crop improvement, providing genes resilient to biotic and abiotic pressures and conferring a higher nutritional value (e.g. Newton *et al.*, 2010; Mendes-Moreira & Vaz Patto, 2012; Leitão *et al.*, 2013). Therefore, LR have great importance in human welfare and food security. In the uncertainty brought by the expected human population increase (United Nations, 2015) and by the potential climate change effects (Jarvis *et al.*, 2010), LR need to be studied and conserved so humankind may cope with these challenges and future food security will not be hampered.

1.2 THE DEFINITION OF LANDRACE

A century ago, European agriculture was grounded on traditional farming systems and farmers kept an amount of the harvested seeds to cultivate in the following year. This recurrence, in different regions and carried out by different people with different necessities, originated LR (Negri *et al.*, 2009).

Several definitions of LR were attempted over time, trying to encompass all the inherent complexity (see Zeven, 1998). Apparently the first attempt was from von Rümker in 1908 that specified LR as a variety with a name given after a certain location where it is adapted, and where it grows with no human interference within a large timeframe (Zeven, 1998). More recently Camacho-Villa *et al.*, (2005), after researching landrace bibliography and conducting key informant interviews defined distinct traits connected to LR: *historical origin*, *recognizable identity*, *lack of formal genetic improvement*, *high genetic diversity*, *local genetic adaptation* and *association with traditional farming systems*.

Concerning origin, LR usually have a long period of existence associated, usually more than one generation. They can be autochthonous or allochthonous according to being cultivated in one region for a long period of time or introduced in another location and gradually adapted over time, respectively. *Creole* LR are originated from bred varieties selected and sowed repeatedly by farmers over time (Camacho-Villa *et al.*, 2005).

LR usually have morphologically recognizable identity and are also recognised for their common name, which is problematic due the number of homonyms and synonyms connected - often the same name is used to identify different LR (homonyms) or different names are used to identify the same LR (synonyms) (Camacho-Villa *et al.*, 2005).

In addition, LR are not usually developed in the formal plant breeding system and are normally highly genetically diverse, adapted to the environment where they were selected. They may also have a connection to the traditional farming systems within where they were established (Camacho-Villa *et al.*, 2005).

However, numerous and historically referenced crop LR (e.g. apple, pear, grape LR) would not be identified as such if all these characteristics ought to be used. For instance, vegetative propagated crops, such as fruit trees, originate more genetically uniform organisms and LR used in large scale organic farming systems are not associated with traditional systems (Camacho-Villa *et al.*, 2005). Hence, the authors proposed a broader definition, concluding that every trait has the same relevance and different combinations of traits should be used for the definition of different LR:

"A landrace is a dynamic population(s) of a cultivated plant that has historical origin, distinct identity and lacks formal crop improvement, as well as often being genetically diverse, locally adapted and associated with traditional farming systems." (Camacho-Villa *et al.*, 2005)

1.3 NOMENCLATURE AND OTHER PROXY DEFINITION

The homonyms and synonyms often generate confusion when working with LR. Homonyms (i.e. different LR with equal names) and synonyms (i.e. the same LR has more than one different name) are problematic if we do not have access to a sample of the crop in farmer's field, a photograph or any type of register to clarify the situation. The Portuguese cowpea LR (*Vigna unguiculata* (L.) Walp.) and the chickling pea LR (*Lathyrus sativus* L.) both called *Chícharo*, and the common bean LR (*Phaseolus vulgaris* L.) called *3 Luas* and *Catarino*, are examples of homonyns and synomyms, respectively. Molecular work could be helpful to differentiate LR, but this is not always possible due to time and financial constraints as well as a lack of skilled researchers.

Usually researchers use nomenclatural or/and experts' proxies in their studies (Maxted *et al.*, 2009; Negri *et al.*, 2009). Some studies are in agreement with this nomenclatural proxy where it is assumed the connection between nomenclature and LR genetic diversity (e.g. Harlan, 1992), which means LR with same name are the same LR, and others do not completely agree (e.g. de Haan *et al.*, 2007). Likewise, expert´s opinion (e.g. farmers expertise) is used to clarify and distinguish LR. The two proxies should be used together when genetic information from molecular work is not available.

1.4 THE VALUE OF LANDRACES

The undeniable value of LR encompasses benefits to farmers, breeders and all agrobiodiversity (as they are part of systems that should be balanced). Farmers are not the only guardians of LR. Gardeners and general enthusiasts can, likewise, keep LR. For that reason, it would be preferable to use *LR maintainers* as it is a broader definition. Nevertheless, *farmer* is the term extensively used in bibliography (Maxted *et al.*, 2013), so it will be used throughout my thesis.

Farmers have been relying directly on LR, profiting from their cultivation, for a long time. LR are cultivated for the organoleptic characteristics and for possible extra income by selling them mostly in local markets (e.g. Brush, 1992; Negri, 2003; Torricelli & Negri, 2015; Veloso *et al.*, 2015a, b). They are usually genetically diverse organisms, adapted to the location where they developed and particularly resilient to marginal environments, thus less susceptible to diseases or pest attacks in the entire yield (Harlan, 1975; Frankel *et al.*, 1995). They may also be appealing for its singularity to niche markets (Brush, 1992) and with possible direct use in sustainable, low-input subsistence and commercial agricultural systems (organic agriculture). Furthermore, breeders may have a probable source of new important traits (Veteläinen *et al.*, 2009; Maxted *et al.*, 2013) that in the advent of an increase of the world population and the potential threats of climate change (Jarvis *et al.*, 2010; United Nations, 2015) may be relevant to sustain food security.

1.5 THREATS TO LANDRACES

Despite their value, several factors threaten LR:

- the replacement by new, genetically uniform cultivars (Frankel & Hawkes, 1975; Harlan, 1975);
- the use of intensive, high-input agricultural systems (Negri, 2005);
- the possible effects of climate change (Jarvis *et al.*, 2010);
- ageing of LR maintainers and non-passage of traditional knowledge (Negri, 2003);
- the desertion of the land caused by migration from rural areas to cities (Negri, 2005);
- obstructive existing legislation (Lorenzetti & Negri, 2009);
- the globalization of food systems and pressure of changing markets (Joshi *et al*., 2004; Maxted *et al.*, 2013).

With the *Green Revolution* from the early to mid-twentieth century, new technologies as new irrigation systems; mechanization; fertilizers, pesticides and herbicides development, enabled great changes in agriculture with consequent LR abandonment (Ceccarelli, 2012). With the advances in genetics and plant breeding, new cultivars were created with higher yields and with considerable range of geographic adaptation, nonetheless genetically homogenous. The new cultivars favoured the development and dispersal of high-input agricultural systems where these varieties could thrive. Despite the high yields may have fed millions, the poorer farmers that could not afford this development were side-stepped, landraces diversity started to diminish due to being substituted by new cultivars, and the environment impacted negatively by the abusive use of chemicals like fertilizers and pesticides (Ceccarelli, 2012). The social and economic changes in society similarly allowed the preference of new cultivars, due the globalization of markets and pressure for standardized products. Legislation that constraints LR commercialization, due to lack of uniformity required for being able for marketing, creates a situation where LR are set aside for registered varieties that obeys the standard parameters of size and homogeneous appearance (e.g. Negri *et al.*, 2009; Maxted *et al.*, 2013). The socio-economic changes in society also led to a general rural migration to the cities, in search for an improvement in living conditions, leaving the main LR maintainers nowadays older people (Negri, 2003, 2005), which makes it difficult to ensure LR conservation in long term.

1.6 NATIONAL STRATEGIC ACTION PLAN FOR LANDRACE CONSERVATION

The goal of a National Strategic Action Plan (NSAP) for plant genetic resources for food and agriculture (PGRFA) (Figure 1.1.) is the establishment of a system to sustainably conserve and use the plant genetic resources of a country, as LR and crop wild relatives (CWR). It encompasses several steps to develop a document where (particularly concerning LR) a list of LR with their location, farming and use traditions and LR maintainer information should be collected. Similarly, a threat assessment and data concerning LR conservation should also be gathered in this document as LR conservation actions (Maxted *et al.*, 2013). The steps are i) the development of a LR checklist: a list of the extant LR in a country or a list of nominated crops, ii) the development of a LR inventory which can be prepared simultaneously with point i), and should include geographic, ecological and ethnographic information about LR and the region where it developed, iii) threats identification and assessment of LR, iv) LR prioritization (if the resources are not enough to cover the conservation of all national LR), v) genetic work of LR selected in the previous point, vi) gap analysis, to assess gaps in *ex situ* and on-farm LR conservation and vii) establishment of the plan, with *ex situ* and on-farm conservation actions (Maxted *et al*., 2013).

As signatory of international treaties as the Convention on Biological Diversity (CBD) and the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA), the Portuguese Ministry of Agriculture, Forestry and Rural Development (*Ministério da Agricultura, Florestas e Desenvolvimento Rural*) and other responsible authorities are aware of the importance of LR conservation and there is a National strategic action plan for plant genetic resources conservation (*Plano Nacional para os Recursos Genéticos Vegetais*) outlined (INIAV *et al.*, 2015).

The Portuguese NSAP encompasses an overview of the importance of PGRFA and to what extent this document can contribute to their sustainable conservation and utilization. It presents the international and national legal framework for these resources, as well as their current state of conservation and use in the country. Finally, it presents the aim, and the strategies/activities expected to promote the conservation and use of the Portuguese PGRFA (e.g. development of an inventory; promotion and improvement programmes with the attribution of quality seals, as PDO and PGI, to LR products; training and workshops for researchers/technicians working with PGRFA; update of the Portuguese PGRFA legislation). Nonetheless, and still concerning LR, some of the stages as the LR inventory development, the threat assessment of Portuguese LR or the genetic analysis of the extant LR in the country or a complete genetic study of a crop variety, are still not developed and the final aim of the strategic action plan for LR conservation not yet achieved.

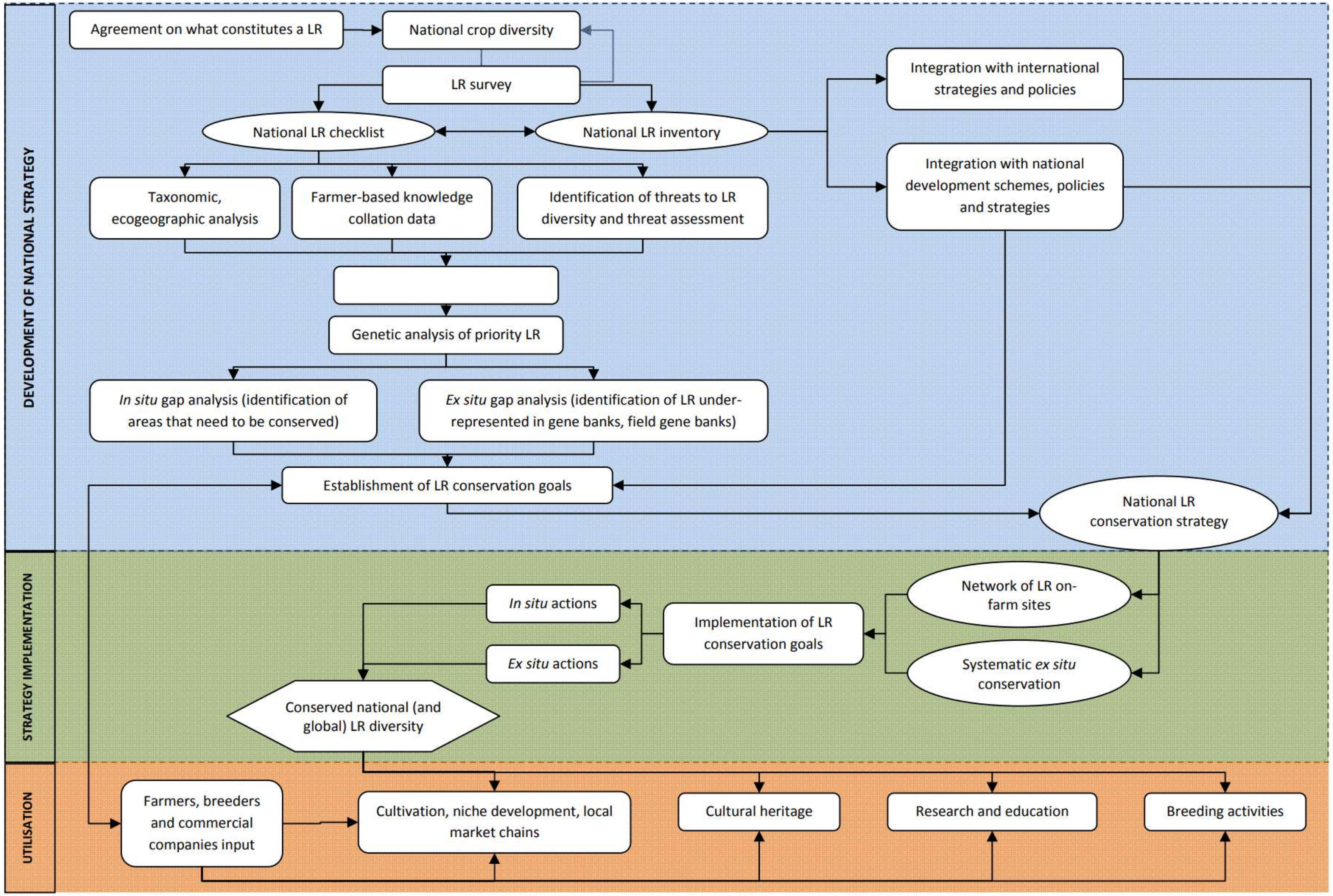


Figure 1.1.: Model for a National Strategic Action Plan for LR conservation (from Maxted *et al.*, 2013).

1.7 LANDRACE CONSERVATION APPROACHES

There are ethical and functional motives when addressing biological conservation. Humankind has an obligation to conserve species and ecosystems regarding the well-being of future generations. The functional motives are related to all aspects of humankind subsistence as it encompasses the conservation of base material for medicine, as food and for healthy ecosystems functioning (Hawkes *et al.*, 2000a). We have been using LR for centuries but regardless of their value (see section 1.4 of this chapter), their extinction and globalized threats have been growing especially in the last century (Negri *et al.*, 2009). Thus, the conservation of these PGR is essential for future food security.

LR can be conserved *in situ* and *ex situ*. Each of the strategies has an array of techniques, with benefits and drawbacks that should be weighted when choosing the conservation strategy. Nevertheless, both *in situ* and *ex situ* conservation strategies should be undertaken and complementary. (Maxted *et al.*, 1997).

1.7.1 *In situ* Conservation

According to the Article 2 of the CBD *"In situ conservation means the conservation of ecosystems and natural habitats and the maintenance and recovery of viable populations of species in their natural surroundings and, in the case of domesticates or cultivated species, in the surroundings where they have developed their distinctive properties."* (CBD, 1992)

This type of conservation can be *passive* or *active*. To be considered active *in situ* conservation it is required the sustainable management and monitoring of the specific taxa and environments where it is found (Maxted, 2000). *In situ* conservation of LR can be carried out *on-*

farm or in *home gardens*. On-farm conservation relies on farmers' willingness to cultivate LR. It is dynamic since it sustains the natural evolution of crops concerning biotic and abiotic changes (Maxted *et al.*, 1997). Nonetheless, it is susceptible to changes in society (e.g. migration in search for different jobs and way of living) and in agricultural practices (e.g. adoption of high-input farming systems and modern varieties) (Maxted *et al.*, 1997; Negri, 2005). Home garden conservation is a scale-down on-farm conservation in terms of size, though more species diversity may be found. The advantages and disadvantages of this technique might be similar to on-farm (Maxted *et al.*, 1997).

1.7.2 *Ex situ* Conservation

Article 2 of the CBD also defines:

"Ex situ conservation means the conservation of components of biological diversity outside their natural habitats." (CBD, 1992).

Ex situ conservation comprises collecting samples and transport of seeds, vegetal material, to an institution where adequate maintenance conditions are required. It can be applied to wild or domesticated species and may be carried out via: *seed storage*; *in vitro storage*; *DNA storage*; *pollen storage*; *field gene banks* and *botanical gardens*. LR are mostly maintained *ex situ* within seed storage and field gene bank techniques. Vegetatively propagated crops and species with recalcitrant seeds may also be conserved *ex situ* using *in vitro* conservation techniques. *Ex situ* conservation facilitates the access to the material and are effective for medium and long-term conservation. However, seed and *in vitro* storage stops the natural evolution of crops concerning biotic and abiotic changes. Field gene banks may be more vulnerable to pest and diseases outbreaks

(Maxted *et al.*, 1997). DNA and pollen storage require specialized technicians for their maintenance, as well as appropriate facilities.

1.7.3 A Complementary Conservation

In situ and *ex situ* conservation approaches should not be individual solutions to LR conservation. Article 9 of the CBD considers that they should be used together and complement each other (CBD, 1992) since both present advantages. Some points must be addressed enabling the researcher to decide the combination of techniques to use in their work. Hawkes *et al.*, (2000b) have defined them as:

- Storage features of the selected species;
- Breeding system of the selected species;
- Timeframe of conservation;
- Possible access to the material;
- Gene bank infrastructure;
- Type of backup possible/necessary;
- Type of sustainable use of the material.

1.8 DIVERSITY CONSERVATION AND USE OF LANDRACES IN PORTUGAL

Portugal geographical location in Southwestern Europe with Mediterranean and Atlantic climatic influences, along with the geology and morphology of the country allows a great variety of species (Figure 1.2. to 1.10.) (Carvalho, 2008; Veloso, 2008). Specifically, the ecogeographic and agricultural characteristics of Portugal created the possibility of the development of a large number of LR of several crops (Veloso, 2008).

Figure 1.2.: Traditional banana varieties – Madeira island.

Figure 1.3.: Various traditional fruits – Madeira island.

Figure 1.4.: Traditional tomatoes and potatoes – Madeira island.

Figure 1.5.: Various traditional fruits – Madeira island.

Figure 1.6.: Traditional common bean varieties – Madeira island.

Figure 1.7.: Typical small size fields in northern continental Portugal.

Figure 1.8.: Farmer in backyard garden in northern continental Portugal.

Figure 1.9.: Farm in southern continental Portugal.

Figure 1.10.: Backyard garden with common bean – S. Miguel, Azores archipelago.

Nevertheless, most LR conservation nowadays is *ex situ,* being around 35% of the gene banks collections (Barata *et al.*, 2008). Concerning *in situ* conservation, there might be only one project to be considered active on-farm conservation in Portugal, the VASO (Sousa Valley) Project (Mendes Moreira *et al.*, 2009; Barata, personal communication 2016).

The *Banco Português de Germoplasma Vegetal* (BPGV) in Braga, *Banco de Germoplasma ISOPlexis* (ISOPlexis) in Madeira island and *Banco do Centro de Biotecnologia dos Açores* (BCBA) in Azores archipelago are the main institutions in charge of the maintenance of *ex situ* seed storage collections (Barata *et al.*, 2008). BPGV also maintain *in vitro* storage collections.

Concerning the principal *ex situ* field gene bank collections, there is a national vineyard collection (*Coleção Ampelográfica Nacional*) and a reference *Olea* collection (*Coleção Portuguesa de Referência de Cultivares de Oliveira*) within INIAV (*Instituto Nacional de Investigação Agrária e Veterinária*) jurisdiction. The *Direção Regional de Agricultura e Pescas do Norte* (DRAPN) maintain a national apricot, almond and *Olea* collections. The main *Malus* and *Pyrus* collections are kept by *Direção Regional de Agricultura e Pescas do Centro* (DRAPC) and *Citrus*, almond, carob, figs, loquat and pomegranate collections are sustained by *Direção Regional de Agricultura e Pescas do Algarve* (DRAPAlg) (INIAV *et al.*, 2015). There are field collections of vineyards and a *Citrus* field collection in the Azores Autonomous Region, (Santos, Sousa and Paulos, personal communication 2017 and Santos, personal communication 2017). In the Madeira archipelago the *Centro Experimental do Farrobo* presents a field collection of traditional varieties of cereals (Freitas, personal communication 2017). Nevertheless, the collections are being assessed by BPGV and INIAV experts to a more accurate and updated perspective of their status.

Based on the CBD (1992), Maxted *et al.* (1997) and Maxted (2000) on-farm conservation definition, there is no active on-farm conservation of LR in Portugal, except for VASO (Sousa

Valley) Project (see Mendes Moreira *et al.*, 2009). It is a Participatory Plant Breeding (PPB) programme where farmers, breeders, cooperatives are intertwined in the production of improved varieties of maize based on selected LR varieties. However, LR of two major crops, vineyards and olive trees, are highly cultivated and economically significant (e.g. Cardoso & Maxted, 2009; Veloso *et al.*, 2015b). Additionally, the INIAV develops plant breeding programmes mostly with forage and cereals crops, which results on the register of several varieties in the National Varieties Catalogues (*Catálogo Nacional de Variedades – CNV*) (INIAV *et al.*, 2015). The product certification as quality products according to EU policy is also present in Portugal. Currently, there are 32 products designated as Protected Designation of Origin (PDO) and 23 products as Protected Geographical Indication (PGI), formed with crop LR varieties, within Class 1.5 and 1.6 and wine products (IVV, 2016; European Commission, 2017).

Likewise, some municipalities organize various events dedicated to specific LR of their region. For example, Alvaiázere municipality celebrates their chickling pea LR (*Lathyrus sativus* L.) with a food festival (*Alvaiázere Capital do Chícharo*, see http://www.alvaiazerecapitaldochicharo.pt/) and São Pedro do Sul municipality organizes a show cooking and a traditional farmer's market where common bean LR (*Phaseolus vulgaris* L.) are sold (Feijão.com(e) – *Festival do Feijão*, see http://www.cm-spsul.pt/conteudo.asp?idcat=293).

1.9 AIM AND OBJECTIVES

The aim of this research is to improve Portuguese LR knowledge and conservation, contributing to the promotion of the Portuguese LR NSAP application. The Portuguese NSAP, document where guidelines for conservation and promotion of sustainable use of Plant Genetic Resources are outlined, needs to achieve their objectives, to have the end point of sustaining the future of genetic resources and food security.

To achieve the aim described previously, we formulated the subordinate objectives:

a) To research bibliographic sources: scientific papers, statistical data, "grey" literature, pertinent websites, for LR data gathering for posterior inventory development;
b) Field work: landrace surveys through questionnaire answering and collection of *Phaseolus vulgaris* L. LR, for LR data gathering for posterior inventory development and molecular work;
c) Develop a comprehensive inventory of LR in Portuguese territory, mainland, Azores and Madeira archipelagos;
d) Develop a standardized threat assessment methodology;
e) Undertake molecular work on *P. vulgaris* L. LR collected throughout field work in order to assess their genetic diversity and parameters that enables the threat assessment;
f) Understand the diversity of *P. vulgaris* L. landraces populations;

1.10 WORK PROGRAMME AND OUTCOMES

Subsequently, the work programme and outcomes will be described:

a) Acquire *Colher para Semear* NGO publications;
b) Develop field work in continental Portugal, Azores and Madeira archipelagos: applying questionnaires and collecting *P. vulgaris* L. samples;
c) Create a database with a comprehensive Portuguese Landrace Inventory;
d) Develop a Landrace Threat Assessment Methodology;

e) Develop a Landrace Threat Risk Calculator;

f) Inventory of *ex situ* and *on farm* Portuguese landraces paper;

g) Landrace Threat assessment methodology paper;

h) Understanding of the diversity of *P. vulgaris* L. landraces populations;

i) Summary with future work suggestions.

CHAPTER II. PORTUGUESE INVENTORY OF FOOD AND OTHER AGRICULTURAL CROP LANDRACES

2.1 ABSTRACT

The threats to agrobiodiversity ultimately affect our future food security. Countries bonded to national and international biodiversity conservation legislation and treaties should conserve and sustainably use their agrobiodiversity. Landraces (LR) are an element of agrobiodiversity largely endangered, partially due to obstacles in LR's inventory development due lack of a standard LR definition, conservation methodologies and nomenclature synonyms/homonyms. However, knowing the extant LR in a country would help effective conservation strategies and could be achieved with a comprehensive inventory. The inventory can be developed by collating data from various sources: literature and media search and farmers, through questionnaire answering and local media. The sites to visit and apply questionnaires to farmers can be selected by using the knowledge of experts and/or using an ecogeographic diversity approach and Geographic Information System (GIS) software. Our work aimed at developing a comprehensive inventory of food and other agricultural LR crops for Portugal. We compiled a total of 14,813 LR entries with 7,492 different LR. *Vitis*, *Phaseolus*, *Zea* and *Brassica* genera as *Vitis vinifera* L., *Phaseolus vulgaris* L., and *Zea mays* L. subsp. *mays* taxa had the highest number of LR with the gathered data. Bragança and Faro were the districts with more LR entries. Our work is a first approach to an inventory that should be updated, reviewed and improved whenever new information is available.

Keywords: Agrobiodiversity, landraces (LR), conservation, management, ELC maps.

2.2 INTRODUCTION

Despite the value of landraces (LR) to farmers and plant breeding, they are disappearing owed to numerous social and economic factors (e.g. Negri, 2005; Negri *et al.*, 2009). For the countries which have signed treaties such as the Convention on Biological Diversity (CBD) and the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA), it is necessary to develop inventories of the plant genetic resources (PGR) they hold (e.g. LR) as the first stage towards achieving the goals of promoting PGR conservation and sustainable use (Negri *et al.*, 2009; INIAV *et al.*, 2015). A LR inventory provides the knowledge of what LR exist, their location within a country and farmer and farming details, which can then be used as a basis for developing efficient conservation actions (Maxted *et al.,* 2013). Nonetheless, a LR inventory can comprise a wider or a narrow array of information, i.e., it can include all LR of a country or LR of a specific area or LR of a crop group. However, an inventory of all LR existing in a country is advisable whenever possible, though remains rare. The gathered data should contain ecogeographic information; farm and farmer facts; nomenclature, cultivation and habits as other possible LRs' specificities (Maxted *et al.*, 2009). This information can be surveyed from several sources such as: experts' meetings; companies and non-governmental organizations (NGOs) connected to agriculture and development; literature (e.g. scientific papers, historical documents and "grey" literature); relevant websites (e.g. national and international gene banks) and farmers (e.g. through questionnaire answering) (Maxted *et al.*, 2009).

For the selection of sites to apply questionnaires to farmers and collect data and samples, we can use the researched literature, experts' knowledge, or an ecogeographic approach.

The latter requires development of an Ecogeographic Land Characterization (ELC) map to identify diverse environmental areas, with different abiotic features (Parra-Quijano *et al.*, 2012). Thus, each area can be visited and sampled systematically with less effort, saving time and resources.

Portugal, located Southwest Europe in the Iberian Peninsula, includes the territories of continental Portugal, the Azores and Madeira Autonomous Regions, and it has a total area of 92,212 km^2. The Azores and Madeira archipelagos are located in the Atlantic Ocean, with the Azores archipelago comprising nine islands, São Miguel, Sta. Maria, Graciosa, Terceira, São Jorge, Faial, Pico, Flores and Corvo, and Madeira archipelago formed by Madeira, Porto Santo, Desertas and the sub-archipelago of Selvagens islands, with the Desertas and Selvagens islands being uninhabited (Figure 2.1.).

Figure 2.1.: Map of Portugal (from Encyclopaedia Britannica online, published by Encyclopaedia Britannica, https://www.britannica.com/place/Portugal/images-videos/media/471439/61694, accessed 20/09/2018).

The Nomenclature of Territorial Units for Statistics in a regional level (NUTSII) divides Portugal in 7 regions: North, Centre, Lisbon Metropolitan area, Alentejo, Algarve and the Autonomous Regions of Azores and Madeira) (Figure 2.2.).

Figure 2.2.: Portugal NUTS II (adapted from image created by Rei-artur with the permission under the terms of the GNU Free Documentation License, https://commons.wikimedia.org/wiki/File:Portugal_NUTS_II.svg, accessed 20/09/2008).

The 2013 Farm Structure Survey (INE, 2014) recorded 240,527 farms in mainland, mainly in the North and Centre regions (98,824 and 86,291 farms, respectively), 11,825 in the Azores and 12,068 in the Madeira Autonomous Regions. A total of 3,517,740 ha of utilized agricultural area (UAA) was recorded for the mainland, with Alentejo presenting the largest value (58,9%), followed by the North (17,8%) and Centre (15,3%) regions. Alentejo's farm average size is 56,9 ha of UAA per farm, while North, Centre and Madeira Autonomous Region register the lowest average size with 6,5 ha of UAA per farm in North and Centre mainland and 0,4 ha of UAA per farm in Madeira (INE, 2014).

As reported by *Instituto Nacional de Estatística* (INE) the main annual crops cultivated in continental Portugal are divided in five categories, i.e. grain cereals; grain legumes; oleaginous and horticultural crops and potatoes; while the main perennial crops are grouped in seven categories: fresh fruits; berries; sub-tropical fruits; nuts/dried fruits; citrus fruits; olive groves and vineyards (Table 2.1.) (INE, 2017).

In general, mixed cropping farms are present throughout the mainland and generally the aforementioned main crops can be found and produced in the different regions of the continental territory (Carvalho, 2008). Nonetheless, there are predominant crops concerning different regions.

Following, we will present the crops that in one or more regions fulfil ≥50% mainland total production. In the North region: fodder maize; rye; beans; potatoes; apple; cherry; kiwi; walnut; almond; chestnut; table olives; grapes (wine production), with rye, fodder maize, kiwi, chestnuts and almond being more cultivated here than in other regions. Centre region has predominantly maize (grain); beans; potatoes; apple; peach; plum; pear; cherry and grapes (wine production). Peach and pear crops have more relevance in this region than in others. Lisbon Metropolitan area is characterized by having tomato for industry as the main productive crop. In Alentejo region, maize (grain); rice; oat; fodder oat; wheat; barley; chickpeas; tomatoes (industry); olives (table olive and oil production); sunflower and table grapes have more significance than in other regions. Likewise, grapes for wine production, walnut and plum are cultivated in the region. The citrus fruits as oranges and tangerines are the relevant crops in Algarve region (INE, 2017). The Autonomous regions of Azores and Madeira have their major crops referred in Table 2.1.

Portugal's ecogeographic settings with varied landscape and Atlantic and Mediterranean climatic influences along with traditional agricultural characteristics allowed the development of a significant range of LR of the numerous cultivated crops (Carvalho, 2008; Veloso, 2008; Veloso *et al.,* 2008). Nevertheless, Portugal follows the global trend of LR disappearing (Veloso, 2008).

There are three main institutions in Portugal maintaining *ex situ* PGR collections: *Banco Português de Germoplasma Vegetal* (BPGV) in Braga, *Banco de Germoplasma ISOPlexis* (ISOPlexis) in Madeira and *Banco do Centro de Biotecnologia dos Açores* (BCBA) in Azores. The BPGV has the most comprehensive collection, comprising more than 70% of national collections. In the national total, LR material accounts for around 35% of the gene bank collections, with cereals and grain legumes having the highest representation (Barata *et al.*, 2008).

Concerning *ex situ* field collections, the *Coleção Ampelográfica Nacional* (National Ampelographic Collection) has the broader vineyard collection in the country; the *Estação Agrária de Viseu* and *Centro de Formação Profissional de Vidago* have the main *Malus* and *Pyrus* collections; the only national collection of apricots is placed in *Direção Regional de Agricultura e Pescas do Norte* (DRAPN), whereas the *Direção de Agricultura e Pescas do Algarve* (DRAPAlg) presents the main *Citrus* collection with other fruit trees such as carob and fig trees, loquat and pomegranates. DRAPN and DRAPAlg also have almond field collections (INIAV *et al.*, 2015). There are field collections of vineyards in the Azores archipelago, namely the varieties *Terrantês*, *Verdelho* and *Arinto dos Açores* in S. Miguel, Graciosa and Pico islands, respectively (Jorge, Sousa and Paulos, personal communication 2017). There are as well *Citrus* field collections in Pico and Terceira islands in the Azores (Jorge, personal communication 2017). In the Madeira archipelago, the *Centro Experimental do Farrobo* maintains a field collection of traditional varieties of cereals, namely of wheat, oat, barley and rye (Freitas, personal communication 2017).

Table 2.1.: Main crops of the Portuguese territory (adapted from INE, 2017).

Mainland		Azores Autonomous Region		Madeira Autonomous Region	
Annual Crops	Perennial Crops	Annual Crops	Perennial Crops	Annual Crops	Perennial Crops
Grain cereals	Fresh fruits	Beans	Apple	Beans	Apple
Barley	Apple	Beet	Banana	Broad bean (green)	Avocado
Common wheat	Apricot	Broad bean	Chestnuts	Broccoli	Bananas
Durum wheat	Cherry	Fodder maize	Custard apple	Cabbage (*Couve-repolho*)	Cherry
Maize	Fig	Maize (grain)	Orange	Carrot	Chestnuts
Oat	Peach	Potatoes	Passion fruit	Cauliflower	Cider apple
Rice	Pear	Sweet potatoes	Pineapple	Green beans	Custard apple
Rye	Plum	Tobacco	Tea	Lettuce	Kiwi
Triticale	Berries	Yam		Maize (for corncob)	Lemon
Grain legumes	Blackberries			Onion	Mango
Beans	Blueberries			Potatoes	Papaya
Chikpeas	Redcurrant			Pumpkin	Passion fruit
Potatoes	Raspberries			Strawberry	Pear
Oleaginous crops	Sub-tropical fruits			Sugar cane	Plum
Sunflower	Kiwi			Sweet potatoes	Tangerine
Horticultural crops	Citrus fruits			Tomato	Vineyards
Broad bean	Lemon			Turnip	
Broccoli	Orange			Yam	
Cabbage (*Couve-repolho*)	Tangerine				
Cabbage (*Couve-tronchuda*)	Nuts/dried fruits				
Carrot	Almond				
Cauliflower	Chestnut				
Courgette	Hazelnut				
Garlic	Walnut				
Green bean	Olive groves				
Leek	Olives (olive oil)				
Lettuce	Table olives				
Melon	Vineyards				
Onion	Table grapes				
Pea	Wine grapes				
Pepper					
Pumpkin					
Savoy cabbage					
Spinach					
Strawberry					
Tomato (fresh)					
Tomato (industry)					
Turnip greens (turnip+greens)					
Watermelon					

Concerning on-farm conservation, there is no active on-farm conservation of LR in Portugal, except for VASO (Sousa Valley) Project (Mendes Moreira *et al.*, 2009), if we consider Maxted (1997a) definition: *"the sustainable management of genetic diversity of locally developed* *traditional crop varieties, with associated wild and weedy species or forms, by farmers within traditional agricultural, horticultural or agri-silvicultural cultivation systems"* (Maxted *et al.*, 1997a) and being performed actively or passively, depending on whether management and monitoring of the local populations is a part of the process or not. Critically, the farmer having the main role in the maintenance and conservation of LR (Maxted *et al.*, 1997a). However, there is a wide array of LR of vineyards and olive trees, that are highly cultivated and economically relevant (e.g. Cardoso & Maxted, 2009; Veloso *et al.*, 2008).

The EU policy on product quality certification is also applied in Portugal. Protected Designation of Origin (PDO) and Protected Geographical Indication (PGI) have been given to those products produced with LR varieties, stated as Class 1.5. and 1.6. (European Commission, 2017; and see *Portuguese inventory of food and other agricultural crop landraces*, in the CD at the end of this thesis) and wine products (IVV, 2016). There are 55 crop certified products out of which 32 and 23 are PDO and PGI certifications (IVV, 2016; European Commission, 2017). The certification entities were contacted to ascertain the commercialization of products under the quality seal. From 12 obtained answers, 10 products remained certified and sold and 2 are no longer commercialized with the quality label. The certified wine products continued to be commercialized as PDO and PGI (IVV, 2016). The necessary conditions to have quality certification may change and a product might be granted or withdraw from it.

This chapter presents the development of a Portuguese on-farm and *ex situ* LR inventory as the basis to develop a national strategy for the conservation of Portuguese LR.

2.3 MATERIALS AND METHODS

2.3.1 LR working definition

A definition of LR was applied as proposed by Camacho-Villa *et al.*, (2005) and LR distinction based on nomenclature and in agreement with BPGV experts' opinion: LR with the same name and collecting site are the same LR, whereas LR with different names or same name but different collecting sites are different LR.

2.3.2 LR survey and data collection

Published literature, "grey" literature, relevant websites, and field work were the sources of information about Portuguese LR. The websites included *Instituto da Vinha e do Vinho* (IVV, 2016) website and the USDA, ARS, GRIN (2017) platform (USDA, ARS, GRIN, 2017). The USDA, ARS, GRIN (2017) platform was used to standardize the taxonomy of the inventory as well as to obtain the accessions from the Portuguese gene banks (BPGV and ISOPlexis) (up to March 16th, 2017). The National Catalogues, *Catálogo Nacional de Variedades* (CNV, 2016) and *Catálogo Nacional de Variedades - Fruteiras* (CNVF, 2016) were also considered and the varieties registered as *conservation* or *traditional varieties* were collated. Nevertheless, these varieties were not accounted for the results as they do not have collection site or date of collection.

The data recorded were LR names, crop scientific names, collection site and year of collection, mode of consumption and other available remarks as for example adaptation to altitude and drought.

Additionally, field work was carried out in April and between July and September 2015 and a total of 165 farms were surveyed (Supplementary Figure 2.1.). Simultaneously with interviewing the farmers, samples of *Phaseolus vulgaris* L. (common bean) LR were requested for posterior molecular work. A database based on the descriptors of FAO/Bioversity Multi-Crop Passport Descriptors V.2.1 (Alercia *et al.*, 2012) was created to compile the collected data in a standard manner, therefore facilitating its understanding and interchangeability.

2.3.3 Selection of sites

To make effective use of the available time and funding, the sites to survey were selected according to two different approaches: (i) in Madeira and Azores Autonomous regions, experts from Madeira University (ISOPlexis Genebank) and from the Azores Agricultural Departments of each island selected the field sites to be visited therefore, using local experts' experience and knowledge; (ii) in mainland Portugal, with a broader area to survey, an ecogeographic approach was carried out using the CAPFITOGEN tools version 2.0 (Parra-Quijano *et al.*, 2016) to identify diverse ecogeographic areas. Each identified ecogeographic area was visited systematically covering the mainland. An ELC (ecogeographic land characterization) map was developed for common bean LR in Portugal based on ecogeographic variables selected using a two-step process: experts' consultation (from BPGV) and the use of SelecVar tool of CAPFITOGEN tools with the Random Forest Method (RF) (Parra-Quijano *et al.*, 2016). A set of ecogeographic variables (Parra-Quijano *et al.*, 2016) were chosen by BPGV experts', in accordance with the characteristics of the species to be collected, and then tested using RF method:

Bioclimatic variables:

- Average temperature from March to August (usual cropping period of common bean in Portugal)
- Minimum temperature of the coldest month of the cropping period (March)

- Maximum temperature of the warmest month of the cropping period (August)
- Average rainfall of the cropping period months (March to August)

Edaphic variables:

- Gravel, sand, clay and silt contents in surface soil
- Apparent bulk density reference in surface soil
- Surface soil pH
- Organic carbon content in surface soil
- Cation exchange capacity in surface soil
- Salinity of surface soil

Geophysical variables:

- Altitude
- Longitude
- Latitude

With the RF method, all previous variables were analysed using a 0.8 threshold (standard value applied by the programme), and the variables not correlated accessed from the correlation table developed by the programme, and selected for the ELC maps development:

Bioclimatic variables:

- Average rainfall for August
- Maximum temperature for August
- Average temperature for March

Edaphic variables:

- Cation exchange capacity in surface soil

- Silt and clay content in surface soil
- Organic carbon content in surface soil

Geophysical variables:

- Altitude
- Longitude
- Latitude

The final map presented 22 different ecogeographic areas and a degree of resolution was 1x1 km cell size (30 arc-sec) (Figure 2.3.).

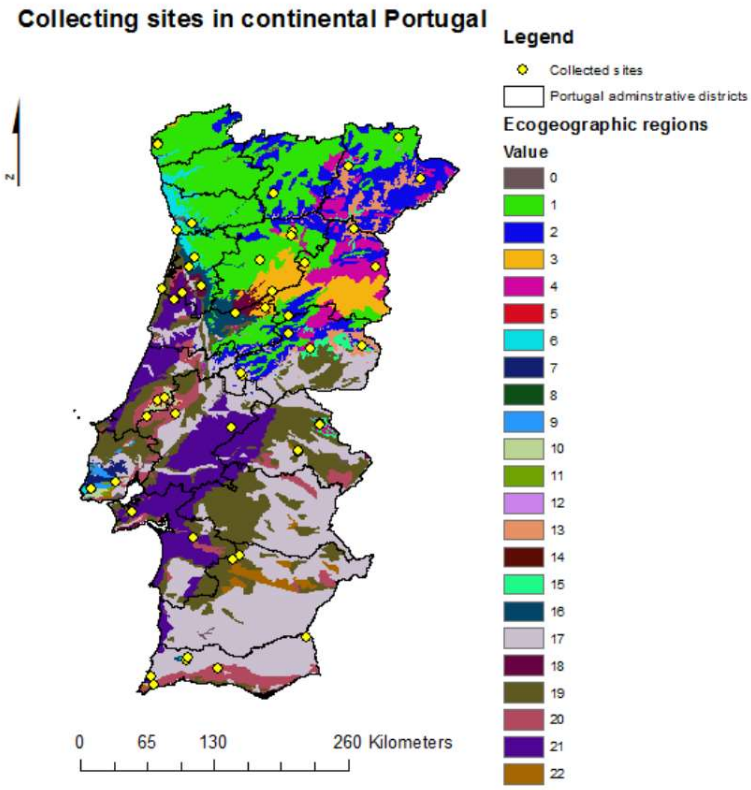


Figure 2.3.: ELC map with collecting points.

2.3.4 Field work general results

The results of the field work survey are comparable with other works (e.g. Negri 2003, 2005), regarding farmers' age, LR's consumption mode, reasons for LR maintenance and decline.

Throughout the Portuguese territory, the majority of LR maintainers are over 65 years old. Farmers are keen on maintaining their LR varieties mostly due to the associated traditions (e.g. gastronomical traditions) and the superior flavour presented when compared to commercial varieties. LR are cultivated predominantly for self-consumption, in backyards or small gardens and not so frequently in small farms (up to 0,5 ha).

However, one can observe a general decline on LR cultivation. Farmers are older and unable to maintain the formerly LR varieties. At the same time, their families have other occupations, generally lacking knowledge and time to maintain the family LR varieties. Nonetheless, there is still a good range of LR diversity requiring strong conservation measures.

2.4 RESULTS

The *Portuguese inventory of food and other agricultural crop landraces* (in the CD at the end of this thesis) contains 14,813 LR entries of LR, identified as being gathered from field work, from gene bank entries or from bibliography. However, after applying the LR working definition, (i.e. same name and collecting site are same LR, different names or same name but different collecting sites are distinct LR), resulted in a total of 7,492 different LR. These LR were grouped within 36 families, 88 genera, 130 taxa and 123 crops (assessed at species level, nonetheless some crops may include different sub-taxa). The four families with highest number of LR were *Fabaceae*, *Vitaceae*, *Poaceae* and *Rosaceae* (Table 2.2.). *Vitis*, *Phaseolus*, *Zea* and *Brassica* are the top genera with the highest number of LR (Table 2.3.). The 130 taxa include both species and infra-specific taxa, and 7,200 LR are classified to the species or infra-specific taxa level, whereas 292 are classified only to genus level, due to the information being collated from USDA, ARS, GRIN (2017) platform (Table 2.4. and Table 2.5.). The taxa with the largest number of LR are grapevine (*Vitis vinifera* L.), beans (*Phaseolus vulgaris* L.) and maize (*Zea mays* L. subsp. *mays*) (Table 2.4.).

Table 2.2.: Number of landraces per family.

Family	Number of landraces
Fabaceae	1,916
Vitaceae	1,470
Poaceae	934
Rosaceae	674
Cucurbitaceae	534
Brassicaceae	523
Solanaceae	435
Amaryllidaceae	251
Asteraceae	148
Moraceae	126
Oleaceae (95), Rutaceae (94), Apiaceae (49), Convolvulaceae (49), Linaceae (32), Araceae (30), Chenopodiaceae (24), Betulaceae (17), Lythraceae (14), Fagaceae (12), Lamiaceae (12), Musaceae (12), Passifloraceae (8), Juglandaceae (7), Ebenaceae (5), Myrtaceae (5), Portulacaceae (3), Annonaceae (2), Caricaceae (2), Ericaceae (2), Lauraceae (2), Asparagaceae (1), Bromeliaceae (1), Plantaginaceae (1), Rubiaceae (1), Theaceae (1)	481
Total	**7,492**

Table 2.3.: Number of landraces per genus.

Genus	Number of Landraces
Vitis	1,470
Phaseolus	1,406
Zea	585
Brassica	478
Cucurbita	309
Solanum	265
Malus	264
Allium	251
Pyrus	227
Prunus	168
Capsicum	166
Triticum	157
Cucumis	147
Lactuca	142
Pisum	125
Ficus	124
Vigna	116
Vicia	115
Olea (95), *Citrus* (94), *Avena* (60), *Citrullus* (56), *Lupinus* (51), *Ipomoea* (49), *Secale* (49), *Cicer* (46), *Lathyrus* (37), *Lolium* (36), *Linum* (32), *Colocasia* (30), *Raphanus* (25), *Hordeum* (23), *Coriandrum* (22), *Lagenaria* (21), *Corylus* (17), *Petroselinum* (17), *Punica* (14), *Beta* (13), *Castanea* (12), *Musa* (12), *Spinacia* (11), *Lepidium* (10), *Nasturtium* (9), *Sorghum* (9), *Cydonia* (8), *Daucus* (8), *Passiflora* (8), *Eriobotrya* (7), *Juglans* (7), *Ornithopus* (7), *Setaria* (6), *Diospyrus* (5), *Ocimum* (5), *Psidium* (5), *Arachis* (4), *Helianthus* (4), *Holcus* (4), *Lens* (3), *Physalis* (3), *Portulaca* (3), *Saccharum* (3), *Satureja* (3), *Annona* (2), *Arbutus* (2), *Carica* (2), *Ceratonia* (2), *Cichorium* (2), *Medicago* (2), *Mentha* (2), *Morus* (2), *Ananas* (1), *Apium* (1), *Asparagus* (1), *Cajanus* (1), *Camellia* (1), *Coffea* (1), *Foeniculum* (1), *Glycine* (1), *Laurus* (1), *Luffa* (1), *Melissa* (1), *Nicotiana* (1), *Origanum* (1), *Oryza* (1), *Panicum* (1), *Persea* (1), *Plantago* (1), *Sinapis* (1)	977
Total	**7,492**

Table 2.4.: Number of landraces per taxa (species and/or infra-specific taxa level).

Taxa	Number of Landraces	Taxa	Number of Landraces	Taxa	Number of Landraces	Taxa	Number of Landraces
*Vitis vinifera**	1,470	*Prunus dulcis*	38	*Eriobotrya japonica*	7	*Cucumis melo var. reticulatus*	2
*Phaseolus vulgaris**	1,358	*Cucurbita moschata*	37	*Juglans regia*	7	*Lathyrus tingitanus*	2
Zea mays subsp. mays	585	*Linum usitatissimum*	32	*Avena strigosa*	6	*Morus nigra*	2
Malus domestica	264	*Colocasia esculenta*	30	*Brassica oleracea var. capitata*	6	*Ocimum basilicum*	2
Solanum lycopersicum var. lycopersicum	246	*Lathyrus sativus*	29	*Capsicum chinense*	6	*Ornithopus compressus*	2
*Pyrus communis**	227	*Prunus persica**	29	*Setaria italica**	6	*Physalis peruviana*	2
*Lactuca sativa**	142	*Lolium multiflorum*	27	*Triticum aestivum subsp. compactum*	6	*Prunus cerasus*	2
*Brassica rapa**	137	*Capsicum frutescens*	26	*Brassica oleracea**	5	*Solanum betaceum*	2
*Pisum sativum**	125	*Prunus avium*	24	*Capsicum baccatum**	5	*Vicia ervilia*	2
*Allium cepa**	124	*Raphanus sativus**	24	*Diospyrus kaki*	5	*Ananas comosu var. cayene*	1
Ficus carica	124	*Hordeum vulgare subsp. vulgare*	23	*Psidium cattleyanum**	5	*Apium graveolens**	1
*Capsicum annuum**	121	*Coriandrum sativum*	22	*Arachis hypogaea*	4	*Asparagus officinalis*	1
*Allium sativum**	114	*Cucurbita ficifolia*	22	*Brassica oleracea var. sabauda*	4	*Cajanus cajan*	1
*Cucurbita pepo**	106	*Lagenaria siceraria*	21	*Helianthus annuus*	4	*Camellia sinensis*	1
*Vicia faba**	103	*Brassica napus**	18	*Holcus lanatus*	4	*Cichorium endivia*	1
*Vigna unguiculata**	100	*Corylus avellana*	17	*Ornithopus sativus*	4	*Cichorium intybus*	1
*Olea europeae**	95	*Petroselinum crispum**	17	*Passiflora edulis**	4	*Foeniculum vulgare**	1

Brassica rapa subsp. Rapa	91	*Vigna unguiculata subsp. sesquipedalis*	16	*Solanum melongena*	4	*Glycine max*	1
*Cucumis melo**	89	*Triticum turgidum subsp. durum*	15	*Vicia sativa**	4	*Lathyrus cicera*	1
Brassica oleracea var. acephala	86	*Citrus limon*	14	*Brassica oleracea var. italica*	3	*Laurus nobilis*	1
Triticum aestivum subsp. Aestivum	73	*Punica granatum*	14	*Lens culinaris**	3	*Lolium perenne*	1
*Prunus domestica**	72	*Beta vulgaris subsp. vulgaris*	13	*Portulaca oleracea*	3	*Luffa aegyptiaca*	1
Citrus sinensis	68	*Solanum tuberosum**	13	*Prunus armeniaca*	3	*Melissa officinalis*	1
Brassica oleracea var. costata	59	*Castanea sativa*	12	*Saccharum officinarum*	3	*Nicotiana tabacum*	1
Citrullus lanatus	56	*Citrus reticulata*	12	*Satureja hortensis*	3	*Origanum vulgare subsp. virens*	1
*Cucumis sativus**	56	*Spinacia oleracea*	11	*Allium ascalonicum*	2	*Oryza sativa*	1
*Ipomoea batatas**	49	*Lepidium sativum*	10	*Allium schoenoprasum*	2	*Panicum miliaceum**	1
*Secale cereale**	49	*Lupinus luteus*	10	*Annona cherimola*	2	*Passiflora ligularis*	1
*Cucurbita maxima**	48	*Nasturtium officinale*	9	*Arbutus unedo*	2	*Persea americana**	1
Avena sativa	47	*Sorghum bicolor**	9	*Carica papaya*	2	*Plantago lanceolata*	1
*Phaseolus coccineus**	47	*Cydonia oblonga*	8	*Ceratonia siliqua*	2	*Sinapis alba**	1
Cicer arietinum	46	*Daucus carota**	8			*Triticum turgidum subsp. turgidum*	1
Lupinus albus	38	*Allium ampeloprasum**	7			*Vicia articulata*	1
						Total	**7,200**

* Taxa not specified

Table 2.5.: Number of landraces that have only been classified at genus level.

Taxa (Genus level)	Number of Landraces
Cucurbita sp.	96
Brassica sp.	69
Triticum sp.	62
Musa sp.	12
Capsicum sp.	8
Lolium sp.	8
Avena sp.	7
Lathyrus sp.	5
Vicia sp.	5
Lupinus sp.	3
Ocimum sp.	3
Passiflora sp.	3
Allium sp.	2
Medicago sp.	2
Mentha sp.	2
Coffea sp.	1
Ornithopus sp.	1
Pasheolus sp.	1
Physalis sp.	1
Raphanus sp.	1
Total	**292**

In Portugal mainland, the districts with the highest number of genera are Aveiro and Évora (43), Bragança and Faro (42), Santarém (39) and Setúbal (38). The highest number of LR are recorded in Bragança (855), Faro (555), Aveiro (537) and Viseu (467) districts (Figure 2.4. and Supplementary Table 2.6.). Some accessions were recorded in the inventory but not included in our regional analysis, because they were from undefined districts and only general localities were known (Supplementary Table 2.7. and 2.8.).

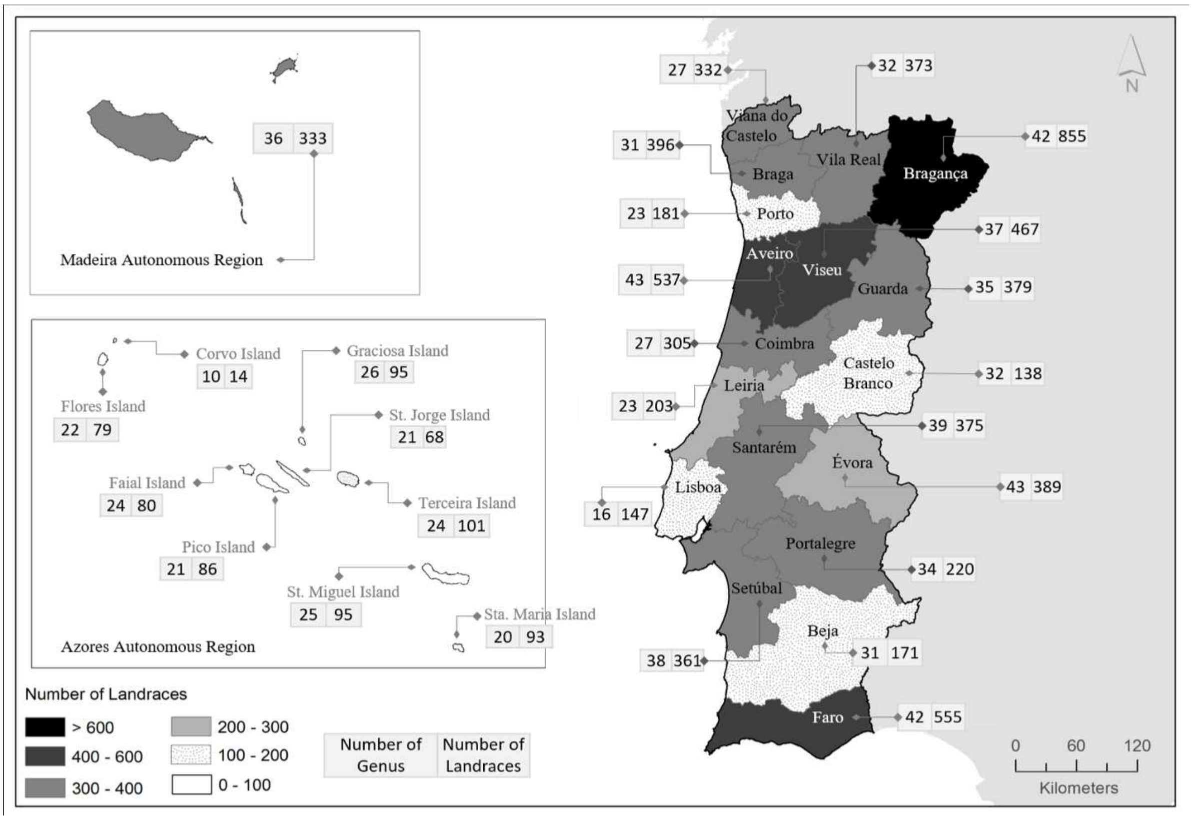


Figure 2.4.: Distribution of landraces in the Portuguese territory according with the collated data.

Beans and grapevine are the most common crops in continental Portugal, while in the Azores and Madeira archipelagos, beans, sweet orange and sweet potato are the most common crops (Figure 2.5.).

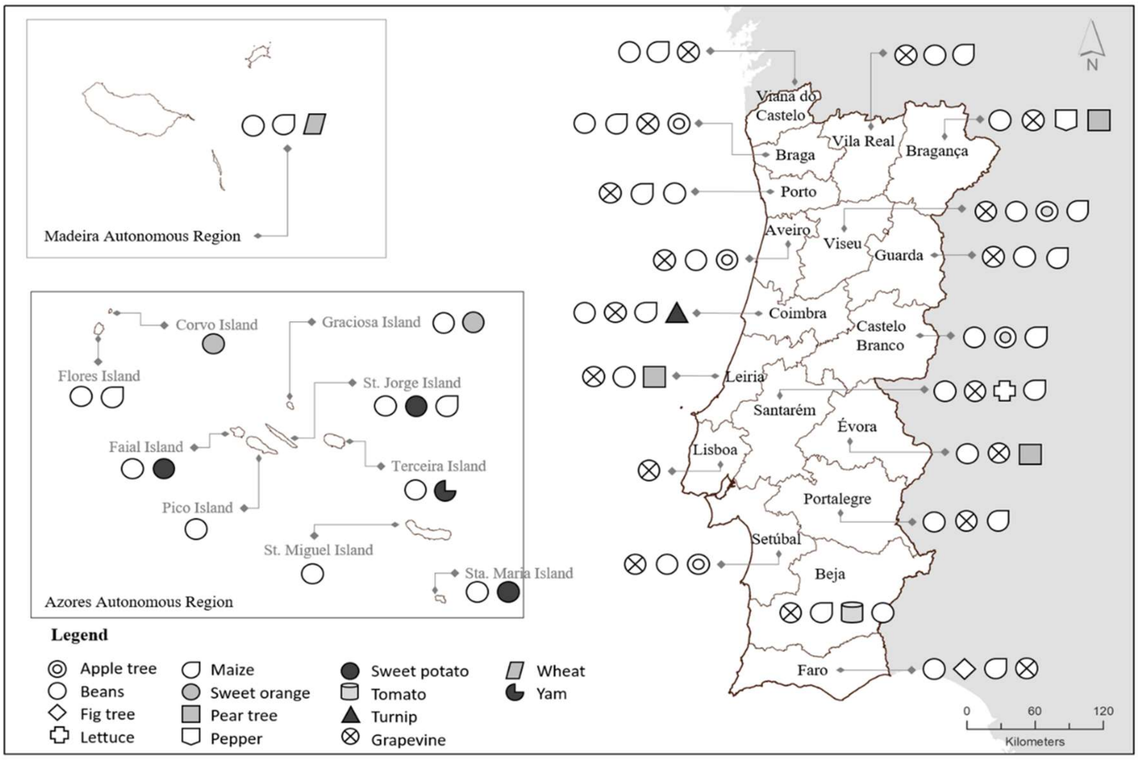


Figure 2.5.: Distribution of the common top crops in Portuguese territory.

The most represented crop groups using FAO Indicative Crop Classification version 1.0 (FAO, 2005) are the leguminous crops and vegetables and melons group (25% and 24% respectively) (Figure 2.6. and Table 2.9.).

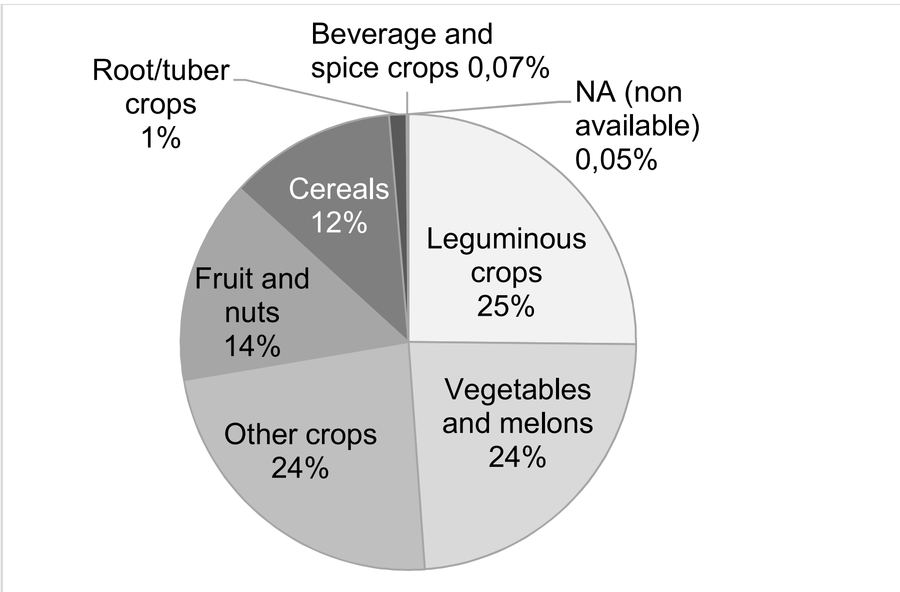


Figure 2.6.: Distribution of landraces per crop group (crop groups according to FAO, 2005).

Table 2.9.: Number of landraces per crop group.

Group	Number of landraces
Leguminous crops	1,885
Vegetables and melons	1,774
Other crops	1,759
Fruit and nuts	1,082
Cereals	891
Root/tuber crops	92
Beverage and spice crops	5
Other	4
Total	**7,492**

The districts with the highest number of LR, were also those where generally leguminous crops were one of the most common (Figure 2.7. and Supplementary Table 2.8.).

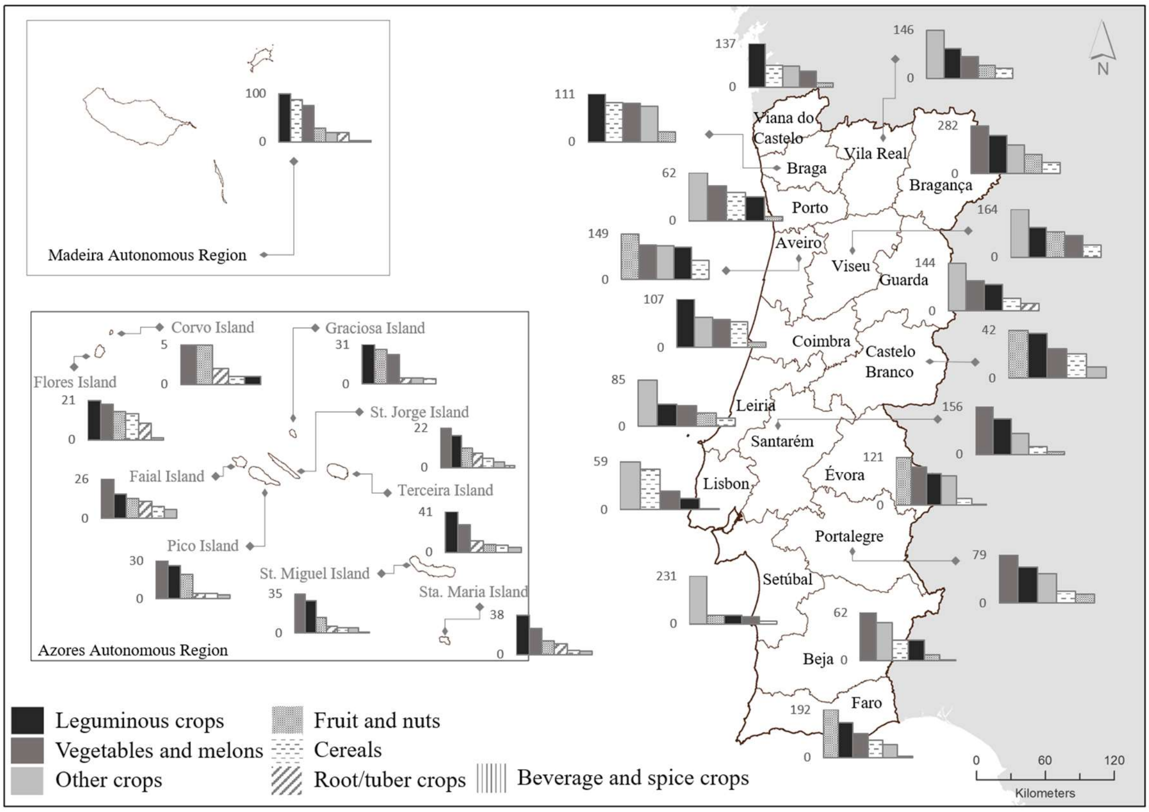


Figure 2.7.: Distribution of landraces per crop group in the Portuguese territory.

2.5 DISCUSSION

Information concerning existing LR is diffuse and its collation is time consuming, but the existence of at least a first iteration of a national LR inventory with 14,813 entries for a country known to have signification agrobiodiversity wealth is a noteworthy advance in LR conservation that will aid LR diversity utilization. The methodology used to collate the inventory data, combining literature and media surveys, ELC maps and field work proved successful and such an approach can be applied in other countries and regions. It was noticeable that all farmers visited, once the study aim was explained to them, immediately grasped the value of the study in helping them preserve their and their neighbours' agricultural heritage. The farmers were then generous with their time and always tried to provide samples whenever possible.

Agrobiodiversity inventories are a "photography", a "snapshot" according to Maxted *et al.*, (2009) of diversity defined by the techniques applied and resources available, but the critical point is that they should be periodic and integrative, repeated regularly to provide time series and thus allow monitoring. Subsequent surveys can then extend the range of LR recorded. Thus, each survey and revised inventory builds toward a more complete inventory over time (Maxted *et al.*, 2009). Therefore, the inventory (a) supplies a useful tool for conservation planning (Veteläinen *et al.*, 2009; Maxted *et al.*, 2009) as it helps future collecting missions to be more systematic, addressing possible gaps and plan on-farm conservation projects, but also (b) provides a means of monitoring LR diversity with possible identification of threats to LR and threatened LR, and (c) promotes utilization of LR diversity by farmer and breeders (Maxted *et al.*, 2009).

However, we should acknowledge the inherent characteristics of non-conservationist based conservation: LR farmers are not conservationists and may have other priorities, habits and traditions for the LR they grow, than ensuring unique alleles are maintained. All which make enforcing standards in LR definition, threat assessment and conservation methodologies and nomenclature synonyms problematic. Further, are we certain all entries included in the Portuguese LR inventory are LR and are they actually distinct LR? Are they all still extant in the country? To ensure precision these questions require extensive genomic analysis and additional field work. But even though the objectivity of some records in the inventory could be queried, what cannot be queried is the value of the inventory itself in providing the conservation tool outlined above. The fact is that LR diversity in Portugal is threatened (Veloso, 2008), and if we wait until sufficient resources are available for the necessary genomic analysis, a significant Portuguese LR diversity will be surely lost. A loss that has implications for future food security in times of growing human population and environmental instability; a loss certainly increased by recent fires in Portugal that have impacted farms and farmers and undoubtedly led to further LR extinction. Through the bibliographic survey and collation of farmer's knowledge, a significant step has been taken to more systematically plan LR conservation and so avert LR loss.

Finally, we want to stress what has been referred previously: that an inventory is a work in progress, with the requirement to repeat and expand periodically, so a web enabled inventory can be updated regularly and have maximum use potential. Therefore, being available at national/international relevant agricultural websites (e.g. agricultural department websites).

2.6 CONCLUSION

A comprehensive *Portuguese inventory of food and other agricultural crop landraces* has been produced that will form a platform for future agrobiodiversity conservation and use. Nonetheless, it is always a work in progress, and research, molecular work, field work and sampling should be performed in a regular basis, to strengthen and continue improving the inventory thus, sustaining future conservation actions. However, the 14,813 LR entries belonging to 7,492 LR included in the inventory demonstrate both significant regional and crop-based patterns of diversity that could help sustain Portuguese national food security and well-being.

CHAPTER III. A THREAT ASSESSMENT METHODOLOGY FOR CROP LANDRACES

3.1 ABSTRACT

The risk of landrace (LR) erosion compel us to act and delineate effective measures to assess the risks they may endure in order to develop successful conservation actions. The IUCN Red List Categories and Criteria is used worldwide for threat assessment at species level, but a standardized methodology to assess the threat status at infra-specific level, i.e. for LR, remains unavailable. In this chapter a standardized and objective methodology for LR threat assessment was developed. A literature survey was complemented with field work to gather LR-related data and screen threats to LR. An online *LR Threat Risk Calculator* was created based on the data gathered to calculate automatically LR threat risk level resulting in a synopsis where all data assembled for each LR is presented along with the calculated risk levels. The synopses of target LR can then be used to help in the development of a threat assessment report, and the information collected used in the production of LR Red Lists. The proposed methodology can be applied at national, regional or global level and to any crop group.

Keywords: Landraces; threat assessment; methodology; IUCN; Threat Risk Calculator.

3.2 INTRODUCTION

One in nine people were predicted to experience chronic hunger worldwide in 2014-16 (FAO, IFAD and WFP, 2015), with the number of people facing severe food constraints increasing 35% in 2015-16 (FSIN, 2017). As the human population is estimated to rise to 9.7 billion by 2050 (United Nations, 2015), it is predicted that global food production will need to grow by around 60% compared to the 2005/2007 period to meet this growing demand (Alexandratos & Bruinsma, 2012). At the same time, crop production may decrease over 25% if crop varieties are not adapted to the changing environment (IPCC, 2014). To meet this challenge, plant breeders will require increased breadth of genetic diversity to increase production (Litrico & Violle, 2015). This diversity is often found in the traditionally grown, genetically diverse, crop landraces (LR), which have not been bred for trait uniformity like modern cultivars.

Camacho-Villa *et al.*, (2005) defined a landrace as *"a dynamic population(s) of a cultivated plant that has historical origin, distinct identity and lacks formal crop improvement, as well as often being genetically diverse, locally adapted and associated with traditional farming systems"*. The importance of the utilization of LR is well recognized, as they often contain unique trait diversity due to their adaption to the location where they developed. Adaptive trait diversity that can sustain yield in marginal environments and mitigate diseases or pest attacks (Harlan, 1975; Frankel *et al.*, 1995), presenting probable traits of significant interest to plant breeding (Harlan, 1975; Frankel *et al.*, 1995; Veteläinen *et al.*, 2009) as for example drought, frost and salinity tolerance (e.g. Newton *et al.*, 2010).

Nevertheless, LR diversity is threatened by various factors: (i) replacement by novel, genetically uniform cultivars (Frankel & Hawkes, 1975; Harlan, 1975; Negri, 2005); (ii) the possible general effects of climate change (Jarvis *et al.*, 2010); (iii) ageing of farmers and ineffective transmission of knowledge related to LR (Negri, 2003); (iv) the desertion of the land caused by migration from rural areas to cities (Negri, 2005); (v) restrictive existing legislation where seeds need to be registered, at a cost often expensive to farmers, to be enabled for sale (Maxted *et al.*, 2013); and (vi) the internationalization of food systems and pressure of changing markets with restrictive food standards (Negri, 2003; Joshi *et al.*, 2004; Maxted et al., 2013). Thus, LR diversity is at risk of both genetic (Hammer *et al.*, 1996; Negri, 2005) and cultural erosion and/or extinction, impacting food security and cultural growth (Negri, 2005).

Conservation planning normally involves some form of prioritization because the conservation target is too extensive to avoid focusing conservation action. One commonly applied means of prioritisation is relative threat assessment, assessing the risk of extinction (Maxted *et al.,* 2013). At the species level, the IUCN Categories and Criteria are universally recognised and used for threat assessment (IUCN, 2001). However, adapting the IUCN Categories and Criteria to LR is problematic, as they do not take into consideration the need to retain genetic as well as taxonomically recognized diversity. Thus, no standardized LR threat assessment methodology is currently widely accepted and easily applied, though proposals have been made for their development as well of agricultural crops (e.g. Joshi *et al.,* 2004; Porfiri *et al.,* 2009; Padulosi & Dulloo, 2012).

Joshi *et al.,* (2004) proposed categorizing LR based on: *population*, *ecological* and *social* (taken from Brush, 2000), and *use* and *modernization* criteria. The authors' LR categorization has parallels to the IUCN categories. Hammer & Khoshbakht, (2005) developed a list of threatened crop species by correlating mainly the list in the 3rd edition of Mansfeld's Encyclopaedia of Agricultural and Horticultural Crops (Hanelt & IPK, 2001) with the IUCN Red List of Threatened Plants (Walter & Gillett, 1998), organising the threatened crop species within the categories: (i) Extinct, (ii) Extinct/Endangered, (iii) Endangered, (iv) Vulnerable, (v) Rare and (vi) Indeterminate. Alternatively, in order to rationally apply regional funds for sustaining landrace cultivation, Porfiri *et al.,* (2009) assessed LR threat level using five criteria: (i) *Presence of the product on the market*, (ii) *Presence in the catalogues of seed companies/nurseries*, (iii) *Number of cultivating farmers*, (iv) *Areas under cultivation* (*as a percentage of the total regional area for the species*), (v) *New dedicated area trend* (presence of new areas reserved to LR cultivation). Antofie *et al.,* (2010) extended the work of Hammer (1991) and Hammer & Khoshbakht (2005) and tried to adapt the Red Listing approach for LR. The authors produced a data sheet for each LR including crop and LR vernacular and scientific names; seed origin; cultivation and location details; conservation status; photographs; authors and references. The data sheet presented information that would help identify LR Red Lists. Further, Voegel (2012) advocates using diverse crop information (e.g. historical material; statistical registers; lists/inventories of cultivars; scientific literature) to formulate a Red List system, based on the continuity of cultivation and use of a crop and cultivars over time in a certain location. A five-step system was suggested by Padulosi & Dulloo (2012) for monitoring agrobiodiversity in order to develop a Red List of cultivated plant species/varieties that involved: i) *Step 1 – General Assessment and Inventory*; ii) *Step 2 – Red List and Vulnerable Variety List Establishment* (with 4 development stages); iii) *Step 3 - First Validation of Red Lists*; iv) *Step 4 - Second Validation of Red Lists*, and v) *Step 5 – Documentation and Monitoring*. Despite the merits of these approaches, most of the proposals do not fully address the requirement to assess LR infra-specific level of threat nor have been widely applied. Also, the lack of information about LR (e.g. LR checklists; LR statistical registers) in most countries, does not enable the use of some of the approaches.

The present work proposes a standardized and quantitative methodology that can be applied objectively to assess LR threat risk nationally. The methodology is illustrated using Portuguese LR. A *LR Threat Risk Calculator* was created to automatically calculate LR threat risk level and was tested using 26 *Phaseolus vulgaris* L., common bean, LR at a national level. In the end, a synopsis for each LR in the study area is obtained. The synopsis presents the information assembled for each LR, as well as the risk level calculated. Therefore, the synopses can be used in the threat assessment of LR and production of a LR threat assessment report. Finally, a LR Red List can be produced with the information gathered previously.

3.3 MATERIALS AND METHODS

3.3.1 Description of the general methodology – LR assessment process

The proposed standardized and quantitative methodology to assess the degree of LR threat composes the following process, also summarized in Figure 3.1. and 3.2.:

1. Pre-threat assessment:

   a. *LR definition*: Agree what constitutes a LR in the assessment and a working definition that can be applied to identify which entities are to be assessed;

   b. *Crop scope*: Agree what crop categories will be assessed, all crop categories or one or more assigned crop categories (FAO, 2005);

   c. *Geographic scope*: Agree the area to be covered, e.g. sub-national (district/region/municipality/town/village) or national level;

2. Threat assessment:

   a. *LR survey*: collate data from farmers using a questionnaire, bibliography or other relevant sources (e.g. official websites) for each recognized LR. Each LR should have a questionnaire associated with specific questions. The assessment criteria include (Figure 3.2.):

      i. Range (A)

ii. Trend (B)

iii. Market and farmer (C)

iv. Conservation status (D)

b. *LR Threat Risk Calculation*: import assessment data into the *LR Threat Risk Calculator* to analyse the level of threat and categorize LR (Figure 3.3. and http://landracethreatriskcalculator.com/).

c. *Compile LR assessment report*: the *LR Threat Risk Calculator* will generate a synopsis for each LR in study. The synopses gathered in the assessment will help to produce a threat assessment report.

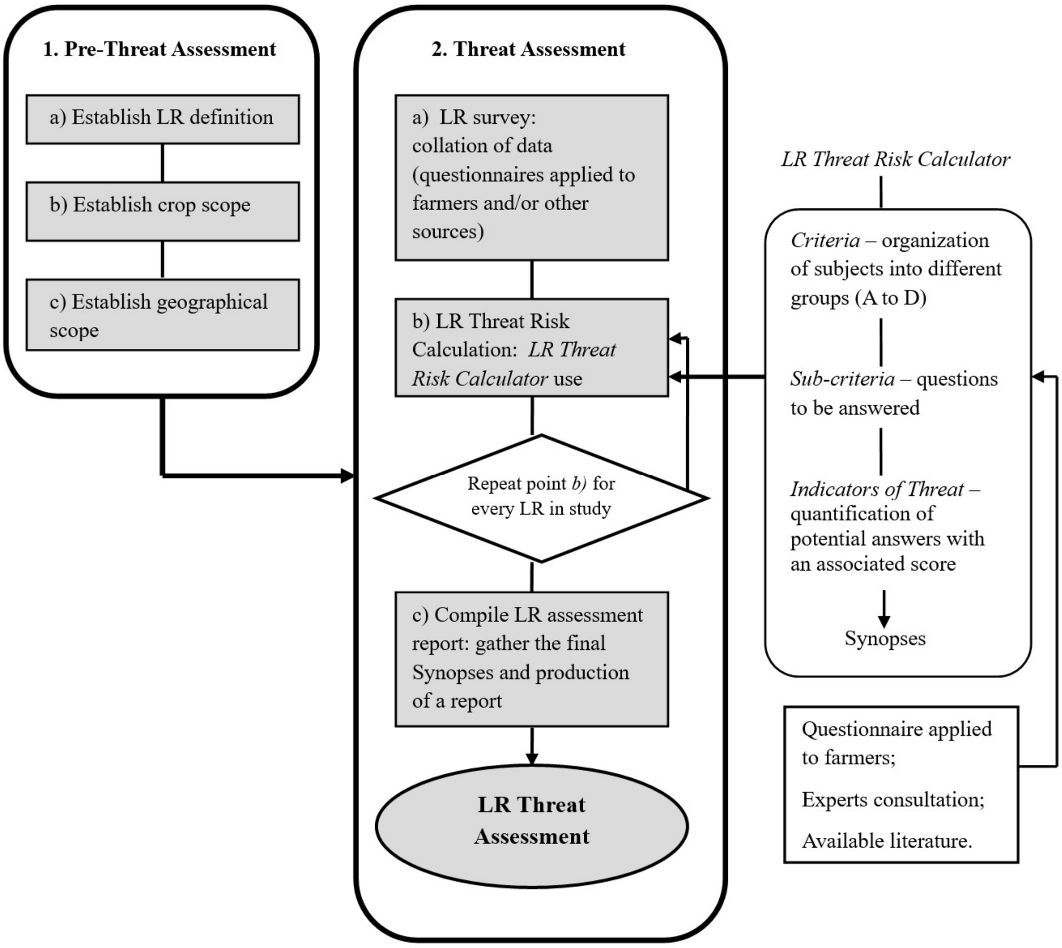


Figure 3.1.: LR threat assessment methodology summary.

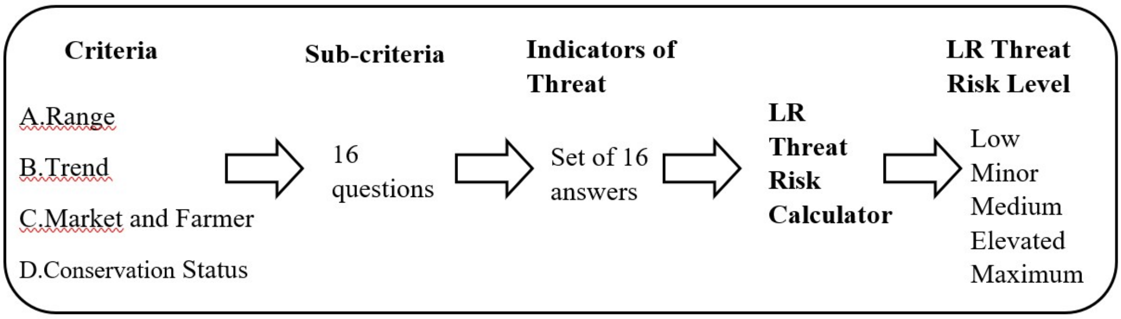


Figure 3.2.: Process of assessing a LR threat risk level.

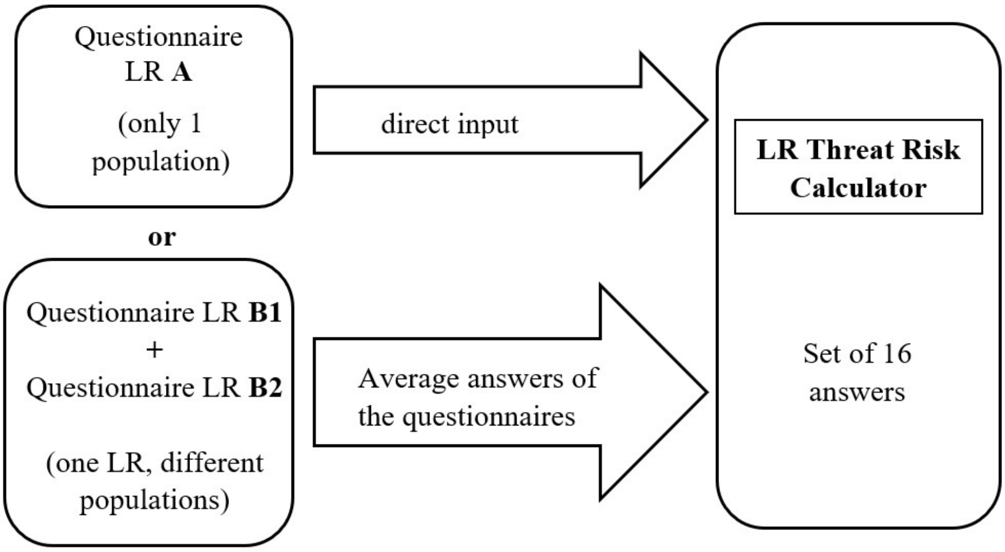


Figure 3.3.: Input answers process into *LR Threat Risk Calculator*. Direct input when there is only one LR (one population – LR A), and input of the calculated averages when there are several populations of the same LR (LR B1 and LR B2).

If there are several populations of the same LR (Figure 3.3.), some calculations should be carried out and them results inserted into the *LR Threat Risk Calculator* (Table 3.1.).

Table 3.1.: Calculations necessary to be performed, for the use of *LR Threat Risk Calculator*.

Calculations (for the necessary indicators of threat)
Percentage of farmers that exchange/save LR seed/material *No. of farmers that exchange, save LR seed or material ÷ No. of sites with the LR × 100*

Percentage of farmers that answered that the area occupied by LR decreased over the past 5 years *No. of farmers answering area occupied by LR decreased ÷ No. of sites with the LR × 100*
Percentage of farmers that answered that LR maintainers are decreasing over the past 5 years *No. of farmers answering LR maintainers decreased ÷ No. of sites with the LR × 100*
The average of the heterozygosity values for a crop LR in the visited sites *∑ heterozygosity values ÷ number of sites with the LR*
Percentage of farmers that answered that LR seed/material are abundant/easily propagated *No. of farmers answering LR seed/material are abundant/easily propagated ÷ No. of sites with the LR × 100*
Percentage of farmers that answered that their families have interest in maintaining LR *No. of farmers' families interested in maintaining LR ÷ number of sites with the LR × 100*
The average of the number of varieties of the crop left and not cultivated nowadays for the visited sites *∑ No. of varieties of the crop left and not cultivated nowadays ÷ No. of sites with the LR*
Percentage of farmers who use modern cultivars of the LR crop *No. of farmers using modern cultivars ÷ No. of sites with the LR × 100*
Percentage of farmers that sell their LR/LR products in the regional or global markets *No. of farmers selling LR/LR products in regional/extra-regional markets ÷ No. of sites with the LR × 100*
Average of farmers' age (When a group of questionnaires is for a LR) *∑ farmers' age ÷ No. of sites with the LR*

Percentage of farmers with sustainable commercial farming (sustainable low input farming system) instead small LR plots *No. of farmers with sustainable commercial farming systems ÷ No. of sites with the LR × 100*
Percentage of farmers using chemical herbicides and fertilizers *No. of farmers using chemical herbicides or fertilizers ÷ No. of sites with the LR × 100*
Percentage of visited sites located in mountain peaks or coastal areas; flood prone areas or with occurrence of droughts known to have happened in 2 or more consecutive years *Nº. of visited sites subject to regular adverse stochastic conditions ÷ No. of sites with the LR × 100*

3.3.2 Target questions and questionnaire

A field survey was conducted in mainland Portugal as well as in the archipelago of Madeira and Azores, during April 2015 and July to September 2015 in order to survey for common bean LR and collect data regarding threats affecting them and LR in general. The information gathered was then used to assess common bean LR threat risk level and test the LR threat assessment methodology. A total of 165 farm sites were visited: 54 in the mainland, 22 in Madeira archipelago and 89 across the Azores archipelago.

A questionnaire was applied to the farmers in each site. It was produced based on those delineated by Kell *et al.,* (2009), Fonseca (2004) and of the *Banco Português de Germoplasma Vegetal* (BPGV) as well as considering the descriptors for *Phaseolus vulgaris* L. (IBPGR Executive Secretariat, 1982). Data on farmers (e.g. age, gender); socio-economic conditions; cultivated crops; cultural practices; qualities of LR; local physiography and topology; soil texture and seed characteristics were collected using the questionnaire (Figure 3.4.).

Questionnaire to farmers

1. Farmers' information

Date: ____________ Altitude: ________________

Latitude: __________ Longitude: ________________

Solar exposition: N / S / E / W / NE / SE / NW / SW

ID: _______ Age: ______ Contact later: y/ n

Name: ______________________________________

____

Address: ____________________________________

District: ____________ Municipality: ____________

Parish: ____________ Locality: ________________

Teleph.: ____________ Mobile: ________________

Fax: _________ Email: ______________________

Member of association? ______________________

Are you interested in maintaining LR long term/ years that been maintaining LR? ______________________

2. Social-economics information

Do yo work in other activity besides agriculture?

____

Farms´area: <0,5ha 0,5–2ha 2,5–5ha >5ha

Nr. people working on the farm: _____ Relation: ________

Are family members willing to maintain LR?

____

3. Crop data

Genus	Species	Local name

Collecting point: field / warehouse / other

____

Purpose: self-consumption / locally sold / sold outside community / partially sold/ other ________________

Used part of plant: stem / leaves / seed / fruit

Plant uses: human consumption/ animal feeding/ other

____

No. of people in the locality that still maintain LR? ______________________________________

Picture(s): y/ n ______________________________

Value of LR in the farm: main crop / harvest before main crop/ harvested after main crop / backyard small productions/ other __________________________

Sowing period	Harvest period

4. Cultural practices

Irrigation / rotation / organic fertilizers / inorganic fertilizers / animal traction / mechanization

5. Qualities of LR

market request / tradition / flavor/ diseases and plagues resistance / high yield / good storage / other

____

6. Local physiography

Plain / valley / mountain / hill / other ________________

7. Local topography

Plan / slop / mountainous / other

____

8. Soil texture: sandy/ clay/ balanced

9. Seed descriptors

Seed coat patterns:

1 / 2 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10 / 99 (see descript. for *P. vulgaris* 4.3.1)

Seed coat darker colour:

1 / 2 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10 / 11 / 12 / 13 / 14 / 99 (see descript. for *P. vulgaris* 4.3.2)

Seed coat lighter colour:

1 / 2 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10 / 11 / 12 / 13 / 14 / 99 (see descript. for *P. vulgaris* 4.3.3)

Brilliance of seed: Matt/ medium/ shiny

Seed shape:1 / 2 / 3 / 4 / 5 (see descript. for *P. vulgaris* 4.3.5)

Observation: ________________________________

____

____

____

____

____

____

____

____

Figure 3.4.: Questionnaire applied to farmers during field work.

The field survey and information gathered with the questionnaire allowed the establishment of Target Questions (Table 3.2.), that should be used in the questionnaires or online questionnaires and answers fed into the *LR Threat Risk Calculator* (Figure 3.3.; http://landracethreatriskcalculator.com/). Thus, enabling the assessment of the LR risk level and categorization, with the proposed methodology.

Table 3.2.: Target Questions – specific questions to be used in questionnaires and in every collecting mission.

1	If the farmer exchanges/saves LR seed/material;
2	Maintainer perception of LR area cultivated decreasing;
3	Maintainer perception of LR maintainers population decreasing;
4	Calculation of the average of heterozygosity values for LR (molecular estimate);
5	If the farmer has the perception that LR seed/material are abundant/easily propagated;
6	If maintainers' families have interest in maintaining LR;
7	The number of varieties of the crop left and not cultivated nowadays for the visited sites;
8	If farmers' use modern cultivars of the LR crop in study;
9	If LR is a certified product;
10	If there are a local, regional, extra-regional market for LR seed/ LR derived product;
11	Maintainers' age;
12	If the LR is maintained in active *in situ* conservation program;
13	If the LR is maintained in active *ex situ* conservation program;
14	If the cultivation systems are sustainable commercial farming (sustainable low input farming system) instead of small LR plots;
15	If the LR's maintainer use chemical herbicides and fertilizers;
16	Is the LR maintained in an area subject to regular adverse stochastic conditions.

In this methodology the *LR Threat Risk Calculator* creates a synopsis for each LR, that are a summary of the data inserted in it. It presents researcher's details, LR scientific and common names, location of study, photographs, remarks, LR threat risk level as scores for each given answer in the calculator and for each LR (e.g. Figure 3.5.: example synopsis of *Amarelo* LR; after field work it was possible to answer the questions on the *LR Threat Risk Calculator* that automatically calculated LR threat risk level).

Landrace Threat Risk Calculator

Synopsis

Your information

Name: Maria Almeida

Institution: University of Birmingham

Location of study: Portugal; Azores Autonomous region - Terceira island; Angra do Heroísmo; S. Sebastião

LR name: Amarelo (ID 20 and 21)

Scientific name: Phaseolus vulgaris L.

Photographs: N/A

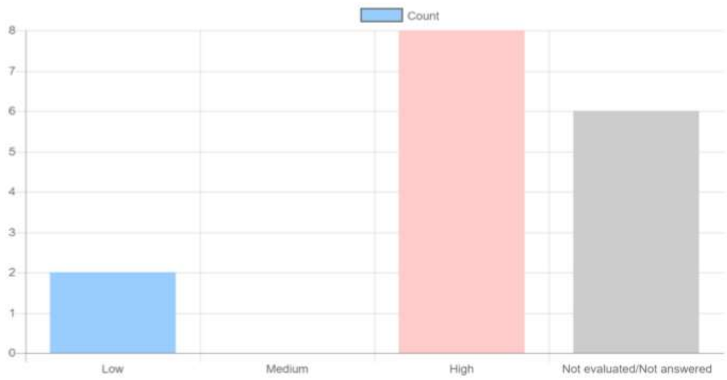
Remarks:

Landrace Threat Risk Level Calculated

Data deficient (2.6 out of 3)

Number of indicators of threat per risk level

Level	Count	Percentage
Low	2	12.50%
Medium	0	0.00%
High	8	50.00%
Not evaluated/Not answered	6	37.50%



Number of indicators of threat per risk level per criteria

Criteria	Low	Medium	High	Not evaluated/Not answered
Criteria: A. Range	1	0	2	2
Criteria B: Trend	0	0	1	2
Criteria: C. Market and Farmer	0	0	3	0
Criteria: D. Conservation Status	1	0	2	2

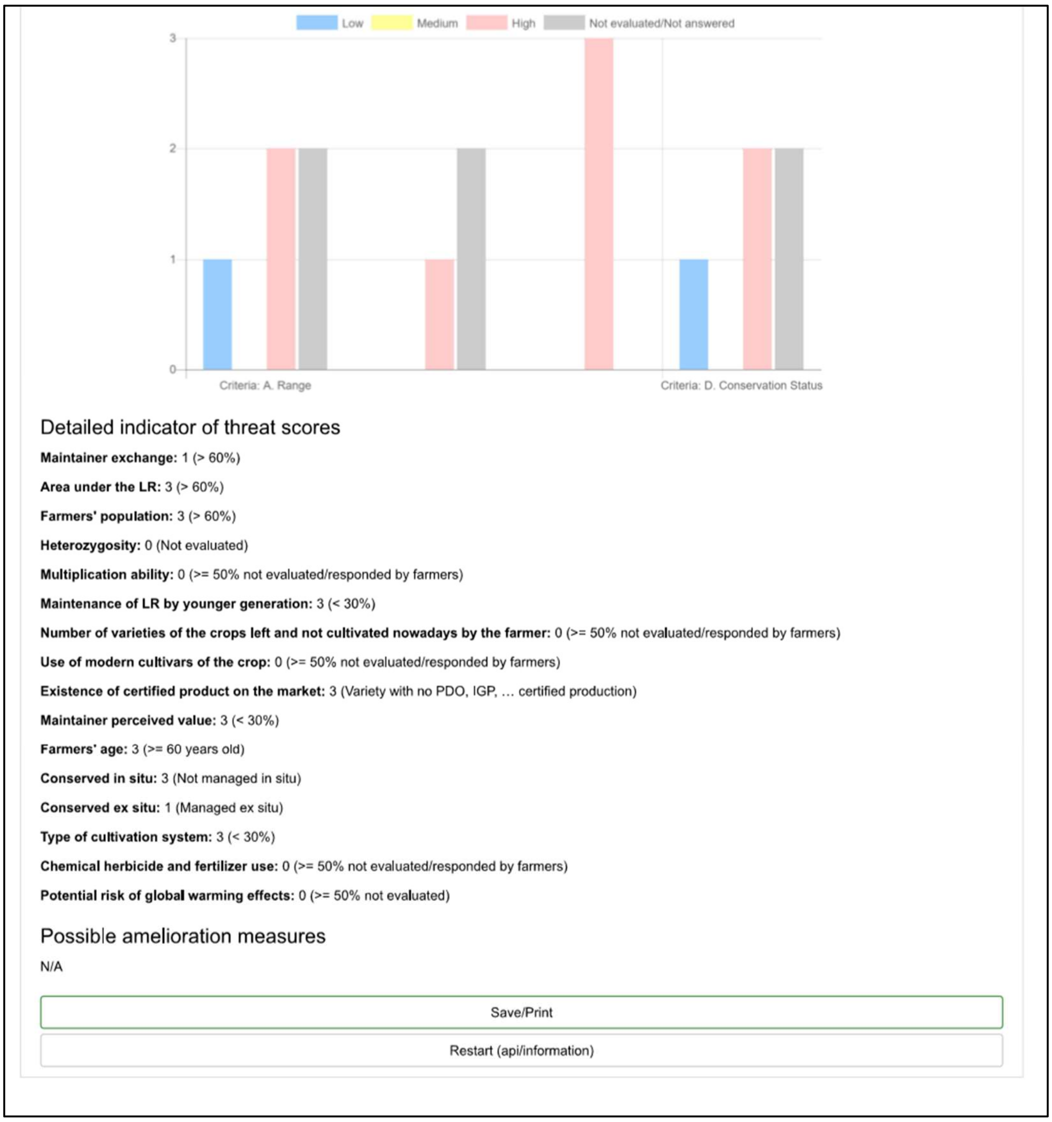


Figure 3.5.: Example of synopsis obtained with the *LR Threat Risk Calculator*.

3.3.3 Criteria, sub-criteria and indicators of threat - description

The field work and questionnaire (Figure 3.4.) enabled to record the adverse conditions and factors that may pose a threat to LR diversity and a set of criteria, sub-criteria and indicators of threat were generalised from those developed for application in Portugal, also considering experts' knowledge and available literature (e.g. Guarino, 1995; Porfiri *et al.,* 2009).

3.3.3.1 Range - Isolation

3.3.3.1.1 Maintainer exchange

The genetic diversity in plants is influenced by its synergy with humans (Maxted and Guarino, 2006). The probable outcome of limited or non-existent LR sharing, if needed, and/or lack of farmers keeping LR in a community lead to loss of LR diversity and possibly their extinction (Maxted *et al.,* 2013). Thus, this sub-criterion would be the number of farmers still maintaining LR seeds/planting material and/or who continue exchanging them.

3.3.3.1.2 Range of LR cultivation

The range of LR cultivation is an estimation of LR cultivated in a determinate area (adapted from Guarino, 1995). A decrease in cultivated area means a higher risk of extinction. Therefore, this sub-criterion records the number of farmers that registered a decrease in the area occupied by LR in the previous five years.

3.3.3.1.3 Number of farmers maintaining LR

As the previous sub-criterion, the number of farmers maintaining LR (adapted from Guarino, 1995) generally provides an estimate of the range of LR cultivated in a certain location. A decrease in LR maintainers means a higher risk of extinction. Therefore, this sub-criterion is the number of farmers answering LR maintainers decreased in the previous five years.

3.3.3.2 Range – Heterozygosity

3.3.3.2.1 Heterozygosity

Genetically similar populations have higher risk of extinction due their limited genepool and their inherent lack of resilience and resistance. Therefore, assess heterozygosity value: 0-1 value, with higher the heterozygosity higher genetic variability. The same molecular marker should be used when assessing heterozygosity of different populations of the same LR, in order to have comparable estimates.

3.3.3.3 Range - Population multiplication

3.3.3.3.1 Multiplication potential

A significant amount of seeds or easily propagated vegetative material is more likely to avoid erosion or extinction (Maxted *et al.,* 2013). Hence, the assessment of the LR seed/material abundance/propagation in this sub-criterion.

3.3.3.4 Trend - Production sustainability

The sub-criterion production sustainability refers to the trend for continued LR production and considers: the interest of farmers' families in maintaining LR over multiple generations (Negri, 2003); the number of varieties known to be lost from cultivation; and the relative accessibility and use of modern cultivars (Negri, 2005; adapted from Guarino, 2005). If the LR maintainers have no long-term, multi-generational interest in maintaining LR, take-up of novel varieties of the crops and stop cultivating LR varieties, will each result in higher extinction risk.

3.3.3.5 Market and Farmer - Market prospects

This sub-criterion assesses if the LR has quality certification and if it is sold in regional and extra-regional markets. If a variety is certified or has been used in the production of certified products, e.g. Protected Designation of Origin (PDO) and Protected Geographical Indication (PGI) (https://ec.europa.eu/agriculture/quality_pt), has better economic perspectives and so, conservation wise, more probability of being maintained and cultivated. Also, if it is sold in regional or extra-regional markets, it has better prospects of long-term cultivation by farmers (e.g. Hannukkala, 2009; Porfiri *et al.,* 2009).

3.3.3.6 Market and Farmer – Farmer generation

This sub-criterion would be the age of the LR maintainers. Farmer ageing threatens LR owed to the prospects of stopping LR maintenance (Negri, 2003). With the field work it was clear that the majority of farmers over 60 years old are more prone to stop cultivating some LR.

3.3.3.7 Conservation Status - Existing conservation action

A consistent and effective plant genetic resources conservation strategy relies on complementary *in situ* and *ex situ* conservation strategies (Maxted *et al.,* 1997). If a LR is a target of these conservation strategies, it has better perspectives of not being extinct. Hence, this sub-criterion would be if the LR are being kept *in situ* and/or *ex situ*.

3.3.3.8 Conservation Status - Cultivation system

LR are adapted to low input farming (not dependent on chemical herbicides and fertilizers) and usually found in subsistence farming systems (gardens, backyards, small plots or allotments). Nonetheless, a subsistence farming system could threaten LR maintenance due to reduced output and lower possibility of widespread, commercial sales. A sustainable, commercial farming system on the other hand, would present lower threat risk and increase farmers' interest in LR maintenance (Negri, 2003).

3.3.3.9 Conservation Status - Climate change, stochastic events and human activity

Climate change is impacting agriculture at a global scale and undoubtedly LR diversity is going to be affected (Jarvis *et al.,* 2010). Certain areas, such as mountain and coastal areas are predicted to be more affected by changes in climate. Drought, for instance, is a consequence of climate change and constitutes a threat to LR (adapted from Guarino, 1995). Therefore, this sub-criterion would be if the area where the LR was collected is subject to any regular adverse stochastic conditions.

3.3.4 Criteria, sub-criteria and indicators of threat - organization

In this proposed methodology, the specific values presented for some sub-criteria and indicators of threat are based on experience gained from field work and adapted from a survey of the literature. It is appreciated that the application of the methods presented in this work, in different geographical contexts, may require some adjustment of these values to capture specific local features; the result presented here is in order to establish the structure of this globally-applicable methodology.

The criteria organize the subjects to be studied into different groups, from A to D (A – Range; B – Trend; C – Market and Farmer; D – Conservation Status), and have 16 equally important sub-criteria associated, identified with numbers, representing the questions made to the farmers. They present the data that should be recorded in order to evaluate the threats to LR at a time and location. The indicators of threat are a group of potential answers, that should be assessed to estimate the relative threat to LR. Each indicator of threat has a linked score of 1, 2 or 3 which grades the threat associated to LR as low, medium or high level, respectively (Table 3.3.). We also employ a score of 0 labelling any category for which no data is available. Not each category has all scores available.

LR diversity can decrease due to several factors and the criteria, sub-criteria and respective indicators of threat reflect and quantify them. If the threats to a LR in a determinate location decrease or cease, that LR can be moved from a high threat to a lower risk level. Similarly, if LR's threats in a certain area expand the threat risk level will increase.

Table 3.3.: List of criteria, sub-criteria and indicators of threat.

CRITERIA	SUB-CRITERIA	INDICATORS OF THREAT (Scores)
A. RANGE	1. ISOLATION 1.1) Maintainer exchange Percentage of farmers that exchange/save LR seed/planting material. 1.2) Range of LR cultivation Percentage of farmers that answered that the area occupied by LR decreased over the past 5 y. 1.3) Number of farmers maintaining LR Percentage of farmers that answered that LR maintainers are decreasing over the past 5 years.	<33% (**1** for 1.2) and 1.3); **3** for 1.1)) 33 to 66% **(2)** >66% (**1** for 1.1); **3** for 1.2) and 1.3)) >=50% not evaluated/responded by farmers **(0)**
	2. HETEROZYGOSITY 2.1) Heterozygosity The average of heterozygosity values for the visited sites.	0 to 0.25 **(3)** >0.25 and <0.75 **(2)** >0.75 to 1 **(1)** Not evaluated **(0)**
	3. POPULATION MULTIPLICATION 3.1) Multiplication potential Percentage of farmers that answered that LR seed/material are abundant/easily propagated.	<33% **(3)** 33 to 66% **(2)** >66% **(1)** >=50% not evaluated/responded by farmers **(0)**
B. TREND	1. PRODUCTION SUSTAINABILITY 1.1) Maintenance of LR by younger generations Percentage of farmers that answered that their families have interest in maintaining LR.	<33% **(3)** 33 to 66% **(2)** >66% **(1)** >=50% not evaluated/responded by farmers **(0)**
	1.2) Number of varieties of the crops left and not cultivated nowadays by the farmer The average number of varieties of the crop left and not cultivated nowadays for the visited sites.	0 to 1 variety **(1)** 2 to 3 varieties **(2)** > 3 varieties **(3)** >=50% not evaluated/responded by farmers **(0)**
	1.3) Use of modern cultivars of the crop Percentage of farmers using novel cultivars.	<33% **(1)** 33 to 66% **(2)** >66% **(3)** >=50% not evaluated/responded by farmers **(0)**
C. MARKET AND FARMER	1. MARKET PROSPECTS 1.1) Active certified product on the market.	Variety with PDO*, PGI**,… certification **(1)** Variety with no PDO, PGI,… certification **(3)** Not evaluated **(0)**
	1.2) Maintainer perceived value Percentage of farmers that sell their LR/LR products in the regional or extra-regional markets.	<33% **(3)** 33 to 66% **(2)** >66% **(1)** >=50% not evaluated/responded by farmers **(0)**
	2. FARMER GENERATION 2.1) Farmers' age Farmers' age or average of farmer's age (if there is more than 1 population of the same LR, use average of farmers' age).	>=60 years old **(3)** 50 to 59 years old **(2)** <50 years old **(1)**
D. CONSERVATION STATUS	1. EXISTING CONSERVATION ACTIONS 1.1) Conserved *in situ*	Managed *in situ* **(1)** Not managed *in situ* **(3)** Not evaluated **(0)**
	1.2) Conserved *ex situ*	Managed *ex situ* **(1)** Not managed *ex situ* **(3)** Not evaluated **(0)**
	2. CULTIVATION SYSTEM 2.1) Type of cultivation system Percentage of farmers with sustainable commercial farming instead of small LR plots. 2.2) Chemical herbicide and fertilizer use Percentage of farmers using chemical herbicides and fertilizers continuously. 3. CLIMATE CHANGE, STOCHASTIC EVENTS and HUMAN ACTIVITY 3.1) Global warming effects potential risks and stochastic events Percentage of visited sites located in mountain/coastal areas; flood prone areas/with occurrence of droughts, human made or wild fires, volcanic activity, known to have happened in 2 or more continuous years within a period of 5 years.	<33% (**1** for 2.2) and 3.1); **3** for 2.1)) 33 to 66% **(2)** >66% (**1** for 2.1); **3** for 2.2) and 3.1)) >=50% not evaluated/responded by farmers **(0)**

*PDO – Protected Designation of Origin; ** PGI – Protected Geographical Indication
Scores: 0 (no data available); 1 (low level threat); 2 (medium level threat); 3 (high level threat)

3.3.5 Description of the methodology basis

3.3.5.1 LR Threat Risk Calculator - threat risk level calculation process

The *LR Threat Risk Calculator* was developed to estimate the threat risk level of LR (see http://landracethreatriskcalculator.com/).

The LR threat risk calculation is based on quantified information about LR's potential threats – indicators of threat. It includes 16 questions, the sub-criteria, linked to four criteria. Each of the 16 questions have a set of answers which are, the indicators of threat. Only one answer (indicator of threat) can be chosen for each question (sub-criteria) (Table 3.3.). Indicators of threat have an associated score of 1, 2 or 3, which categorize the relative threat as low, medium or high-risk level (not all are available for each sub-criterion; see Table 3.3.). Score 0 implies that the indicator of threat was not evaluated or had no answer by the farmer consequently, is categorized as data deficient. The *LR Threat Risk Calculator* combines the scores of each scored indicator of threat (zeroes are ignored in the statistical analysis) to calculate the LR threat risk level score for each LR in study. To this end, we take the mean of all non-zero scores across sub-criteria. If more than 5 indicators of threat were not answered by farmers or the information not accessed by the researcher, the LR threat risk level is classified as *Data Deficient*.

The LR threat risk level score provides a LR threat risk level category. It was decided to categorize the LR threat risk level within *Low*, *Minor*, *Medium*, *Elevated* or *Maximum* risk. Since every answer (indicator of threat) for each sub-criterion falls in the interval [1, 3], the LR threat risk level – the mean of these scores – would also have an associated score within the interval [1, 3]. Hence, the need to determine where is the interval for *Low*, *Minor*, *Medium*, *Elevated* or *Maximum* risk. The score is within [1, 3] interval therefore we could divide the interval in five equal parts; however, the division in equal parts would mean the calculation results of LR threat risk level would fall mainly in central intervals. To adequately assess and weight the diversity in possible responses, we need to control for the distribution of possible scores. Thus, we needed to know how many times each LR threat risk level would happen, given every possible set of answers to all 16 indicators of threat groups of answers (*sets*), which implies many possibilities. To perform the calculations, it was used the R script (R Core Team, 2014) presented in section 3.3.5.2. (a).

Due the large number of possible LR threat risk levels (1 358 954 496), 1 000 000 randomly sampled combinations of *sets* (16 indicators of threat groups) were generated and scored. Every possible response to each sub-criterion, including data deficiency, was allowed equal weighting in this sampling process. Only those sampled sets with <= 5 data-deficient fields were scored (following the protocol), with the score being the mean of the non-zero fields[1]. The distribution of the possible scores is given as a histogram (Figure 3.6.). The histogram allows to see that the threat values would fall mostly in central values, with a substantial mode around a threat level of 2.

---

1 Mathematically, this process is similar to finding the distribution of the mean of *N* discrete uniform random variates on [1,3], where *N* is a random variable drawn from binomial distribution with n=16 and p=0.75, corresponding to the number of non-zero fields in the sampled set. If every field had the same set of possible response (0, 1, 2, 3) this comparison would be exact.

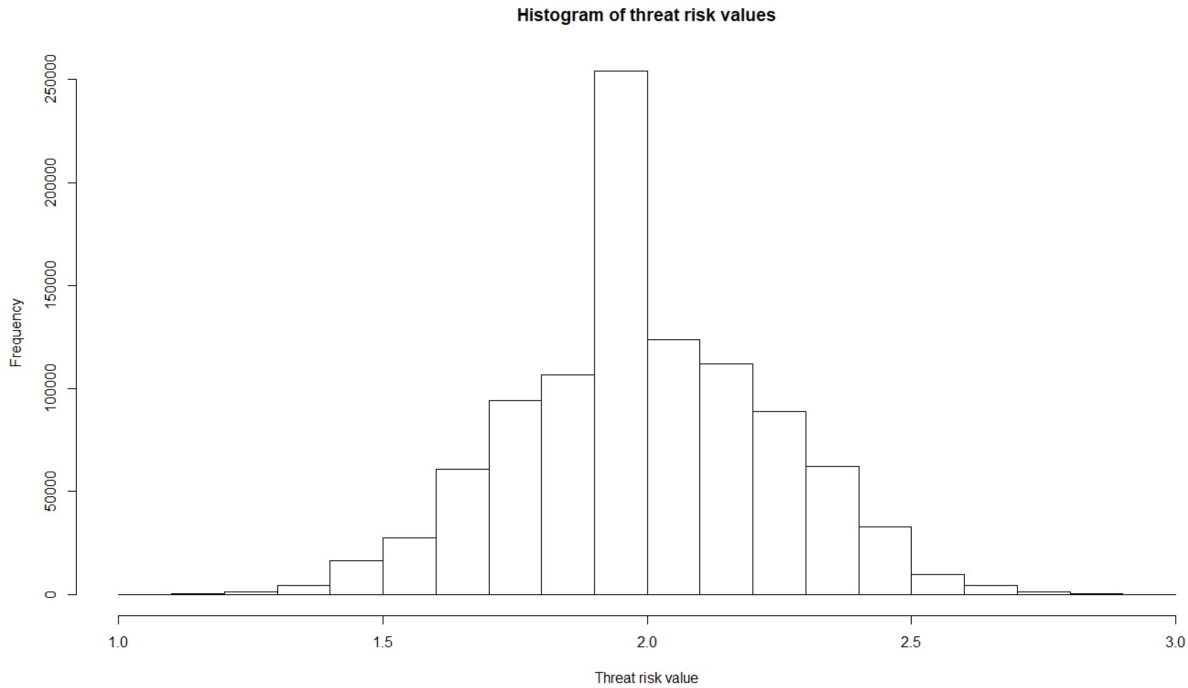


Figure 3.6.: Histogram of threat risk values from the random sampling process described in the text.

To account for the structure of this distribution of possible scores, we divided the distribution into five parts of equal probability mass and assigned each of these regions to one of our categories. In other words, we identified the threat score intervals corresponding to the five 20% quantiles of the distribution and assign these intervals to our ordered categories. An R script was used for this purpose (R Core Team, 2014) [(see section 3.3.5.2., (b)], yielding intervals reflecting quantiles within our expected "null" distribution:

Low [1, 1.80[

Minor [1.80, 1.92[

Medium [1.92, 2.08]

Elevated]2.08, 2.20]

Maximum]2.20, 3]

3.3.5.2 Description of the script

The subsequent R script was developed for the calculation of the total number of the LR threat risk level. It calculates all possible combinations of threat risk levels. The hashtags (#) have the explanation of the following code:

(a)

```
# define a list of all possible indicator values (each sub-criterion has a group of possible answers)
   indicator.value.list = list(
   c(0,1,2,3),
   c(0,1,2,3),
   c(0,1,2,3),
   c(0,1,2,3),
   c(0,1,2,3),
   c(0,1,2,3),
   c(0,1,2,3),
   c(0,1,2,3),
   c(0,1,3),
   c(0,1,2,3),
   c(1,2,3),
   c(0,1,3),
   c(0,1,3),
   c(0,1,2,3),
   c(0,1,2,3),
   c(0,1,2,3)
)
# for each value list, calculate the length (to assess the number of answers for each indicator of threat)
   length.list = lapply(indicator.value.list, length)

# convert length list into length array to apply prod function next
# (function unlist is used as a requirement of R to perform the next step, with the prod function)
   length.array = unlist(length.list)

# multiply all values (to assess all possible combinations)
   total.lr.threat.risk.levels = prod(length.array)
# total number of LR threat risk levels possible: 1358954496
```

The R script presented next was developed to calculate the scores within the defined intervals for threat risk level. Similarly, the hashtags (#) have the explanation of the following code:

(b)

```
# Defining some variables and functions:
# define a list of all possible indicator values (each sub-criterion has a group of possible answers)
   indicator.value.list = list(
   c(0,1,2,3),
   c(0,1,2,3),
   c(0,1,2,3),
   c(0,1,2,3),
   c(0,1,2,3),
   c(0,1,2,3),
   c(0,1,2,3),
   c(0,1,2,3),
   c(0,1,3),
   c(0,1,2,3),
   c(1,2,3),
   c(0,1,3),
   c(0,1,3),
   c(0,1,2,3),
   c(0,1,2,3),
   c(0,1,2,3)
)
# define how to calculate data deficiency
   is.data.deficient = function (answer) {
# get all zeros from the answer and count how many they are (count the zeros in a set of indicators of threat answers)
   number.of.zeros = length(answer[answer == 0])
# return if number.of.zeros is larger than 5 or not
   return ((number.of.zeros > 5))
}
# define how to create a random answer
   create.random.answer = function (indicator.value.list) {

# define how to get a random sample from a list of answer options
# (chooses randomly one answer within the possible group of indicators of threat)
   extract.one.random.element = function (answer.option.list) {
      return (sample(answer.option.list, 1))
   };
   repeat {
# on each possible indicator value extract a random sample
```

```
# (chooses randomly one answer within the possible group of indicators of threat, for each one of the 16 sub-criteria)
    list.of.random.elements = lapply(indicator.value.list, extract.one.random.element)

# convert the list.of.random.elements into a vector (for efficiency)
    random.answer.with.zeros = unlist(list.of.random.elements)

# if the random answer is NOT data deficient, leave this cicle
    if (!(is.data.deficient(random.answer.with.zeros))) {
            break
        }
    }
# remove zeros from the random answer and return that
# we remove zeros because they are not part of an answer
# zeros are the absense of answer to a criterion, therefore
# can't be used to calculate risk
    random.answer = random.answer.with.zeros[random.answer.with.zeros != 0]
    return (random.answer)
}
# Functions are created. Start calculating random answers:
# create a vector
    result = numeric()

# create all random answers and store in the vector (creates randomly 1000000 sets of indicators of threat answers)
    answer.pool.size = 1000000
    print(sprintf("creating %d random answers...", answer.pool.size))

for (i in 1:answer.pool.size) {
# create a random set of indicators of threat
# (randomly chooses 1 answer in each group of indicators of threat, connected to a sub-criteria)
    random.answer = create.random.answer(indicator.value.list)
# calculate the overall risk associated to this answer
    risk = mean(random.answer)
# put the calculated risk in the end of the result vector
    result = c(result, risk)
}
# Random answers generated and risk calculated. Those are all in the result vector.
# Now calculate statistics.

# if we start from 1 and count all risk values until 3 we know we have 100% probability of finding any risk value.
# To get the risk value where the probability matches the expected value per risk level:
# low:   0%  -> 20%
# minor:    20% -> 40%
# medium:   40% -> 60%
```

```
# elevated: 60% -> 80%
# maximum:  80% -> 100%
minor_low = quantile(result, 0.2)
medium_low = quantile(result, 0.4)
elevated_low = quantile(result, 0.6)
maximum_low = quantile(result, 0.8)

print(sprintf("lowest:   [1, %.2f[", minor_low))
print(sprintf("minor:    [%.2f, %.2f[", minor_low, medium_low))
print(sprintf("medium:   [%.2f, %.2f]", medium_low, elevated_low))
print(sprintf("elevated: ]%.2f, %.2f]", elevated_low, maximum_low))
print(sprintf("maximum:  ]%.2f, 3]", maximum_low))
#print function shows on screen the values.
```

3.3.6 Application to a case study to test the method – Common bean LR

As a first step, the concept of LR was defined:

- LR of a crop with different names are different LR;
- LR of a crop with the same name but gathered in different areas (district, city, village) are different LR;
- LR of a crop with the same name located in the same area are the same LR.

Common bean (*P. vulgaris* L.) remains a staple food in Portugal with cultural and gastronomical importance, and farmers still maintain LR of this crop sometimes helping in the household income as they are often sold in local markets (Martins *et al.,* 2006). Given its importance at national level, it was used as a case-study to test the threat assessment methodology proposed in this work.

Secondly, twenty-six samples of *Amarelo*, *Canário*, *Cor de carne*, *Da Ribeira Quente*, *Mocho* and *Touquinho* were selected based on their distribution in all Portuguese territory (mainland and islands) and maintainers needed to have seed saved for at least one human generation.

Thirdly, the resulting data collated for the 26 samples were then imported into the *LR Threat Risk Calculator* to obtain the LR threat risk level. *Amarelo* LR, ID 20 and 21 data are presented as an example in Table 3.4. and having in mind Table 3.1. regarding the calculations necessary for some indicators of threat. These two samples were considered the same LR according to the established working LR definition. They had the same name and were collected from the same location, from two farmers. Thus, for the input of data into the *LR Threat Risk Calculator*, some simple calculations were performed. For example, the first question in the *Calculator* requires the percentage of farmers that exchanges/saves LR seed/planting material therefore:

No. of farmers that exchange, save LR seed or material ÷ No. of sites with the LR × 100

So, and since both farmers saved LR:

(1+1)/2 x 100 = 100%

With this calculation and knowing the range of answers of the *Calculator* for this question (<33%; 33 to 66%; >66% and >=50% not evaluated/responded by farmers), it was possible to choose >66% as the answer (Table 3.3. and 3.4.).

Finally, for each LR a synopsis was compiled as a final product of the *Calculator* (e.g. Figure 3.5.).

Table 3.4.: Criteria, sub-criteria, indicators of threat and calculations example for *Amarelo* bean LR (ID 20, 21) and corresponding result on *LR Threat Risk Calculator*.

CRITERIA and SUB-CRITERIA	INDICATORS OF THREAT	CALCULATIONS (Q1 = questionnaire 1; Q2 = questionnaire 2)	Answer in LR Threat Risk Calculator
A. 1. 1) A. 1. 2) A. 1. 3)	<33% 33 to 66% >66% >=50% not evaluated/responded by farmers	Q1=1; Q2=1 (1+1)/2 x 100 =100% Q1=1; Q2=1 (1+1)/2 x 100 =100% Q1=1; Q2=1 (1+1)/2 x 100 =100%	>66% >66% >66%
A. 2. 1)	0 to 0.25;>0.25 and <0.75; >0.75 to 1;Not evaluated		Not evaluated
A. 3. 1)	<33% 33 to 66% >66% >=50% not evaluated/responded by farmers		>=50% not evaluated/responded by farmers
B. 1. 1)	<30% 33 to 66% >66% >=50% not evaluated/responded by farmers	Q1=0; Q2=0 (0+0)/2 x 100 = 0%	<33%
B. 1. 2)	0 to 1 variety 2 to 3 varieties > 3 varieties >=50% not evaluated/responded by farmers		>=50% not evaluated/responded by farmers
B. 1. 3)	<33% 33 to 66% >66% >=50% not evaluated/responded by farmers		>=50% not evaluated/responded by farmers
C. 1. 1)	Variety with PDO, IGP,… certification Variety with no PDO, IGP,… certification Not evaluated		Variety with no PDO, IGP,… certification
C. 1. 2)	<33% 33 to 66% >66% >=50% not evaluated/responded by farmers	Q1=0; Q2=0 (0+0)/2 x 100 = 0%	<33%
C. 2. 1)	>60 years old 50 to 59 years old <50 years old	Q1=66; Q2=72 (66+72)/2 = 69	>60 years old
D. 1. 1)	Managed in situ/Not managed in situ/Not evaluated		Not managed in situ
D. 1. 2)	Managed ex situ/Not managed ex situ/Not evaluated		Managed ex situ
D. 2. 1) D. 2. 2)	<33% 33 to 66% >66% >=50% not evaluated/responded by farmers	Q1=0; Q2=0 (0+0)/2 x 100 = 0%	<33% >=50% not evaluated/responded by farmers
D. 3. 1)			>=50% not evaluated/responded by farmers

3.3.7 Comparing methodologies

Porfiri *et al.* (2009) presented a model and applied it to Lazio region, Italy. The model encompasses questions within five group of criteria: A – presence of the product on the market; B – presence in the catalogues of the seed companies/ nurseries; C – number of cultivating farmers; D – areas under cultivation (as percentage of the total regional area for the species); E – new dedicated area trend. Each question has a connected score of 1, 2 or 3, which categorize the LR threat as low, medium or high, respectively (Table 3.5.).

Table 3.5.: Indicators, risk level and associated score adopted by Porfiri *et al*. (2009), to assess LR genetic erosion from Lazio region.

Indicator	Description[2]	Risk level	Score
A Presence of the product on the market	Markets and/or producer's cooperatives Sector: main variety in a certain DOC, DOP, IGP, IGT certified production	Low	1
	Niche market: locally limited cultivated areas Market section: secondary varieties in a DOC, DOP, IGP or IGT certified production	Medium	2
	Only some fruits/few seeds available for consumption or research No product on the market	High	3
B Presence in the catalogues of the seed companies/ nurseries	Fruits: presence in variety list A, B and C Vegetables and plants: listed in the national register of varieties Grapevine: listed in the regional register	Low	1
	Grapevine: under registration to regional register Propagation materials available at a few nurseries	Medium	2
	Fruits: not registered in the variety list Vegetables and plants: not registered in the national register of varieties Grapevine: not registered in the regional register No propagating material available out of the maintaining farm	High	3
C Number of cultivating farmers	> 100	Low	1
	30 to100	Medium	2
	< 30	High	3
D Areas under cultivation (as percentage of the total regional area for the species)	> 5 %	Low	1
	1 to 5 %	Medium	2
	< 1% Isolated plants or home garden cultivations	High	3
E New dedicated area trend	New areas dedicated to landrace present	Low	1
	No new areas dedicated to landrace present	High	3

The authors decided that, after the sum of all scores, LR present low risk of erosion if the total score is ≤9; if the total score lies within 10 to 13, LR present medium risk; and if the total value is ≥14, LR are classified having high risk of erosion.

To apply this method to the set of twenty-six samples of common bean, previously referred in section 3.3.6, the considered LR working definition was the same. Concerning the criteria C – number of cultivating farmers, criteria D – areas under cultivation and criteria E – new dedicated area trend, it was considered the number of farmers of the local of the surveyed farm(s), as well as the areas under cultivation or new dedicated areas for LR cultivation of the local of the surveyed farm(s).

3.4 RESULTS

Of 26 common bean samples included in the case-study, *Canário, Mocho* and *Touquinho* LR have one accession each, the *Cor de carne* and *Da Ribeira Quente* have 2 accessions each, and *Amarelo* has 19 accessions. Samples ID 20 and 21, ID 37, 38 and 43, and ID 39 and 41 were considered 3 different LR of *Amarelo*, all other samples were regarded as single, distinct LR. Consequently, 22 unique accessions were included in the assessment.

All answers (indicators of threat) for each unique accession were imported into the *LR Threat Risk Calculator* and the score and LR threat risk level was calculated. Considering Table 3.6. and all information gathered in the synopses of the selected group of 26 LR, eight samples (ID 13, 20, 21, 37, 38, 43, 67 and 87) were Data Deficient, i.e. there was more than 5 indicators of threat without score. One sample was categorized as Minor threat risk level (ID 1M). For all others a Maximum threat risk was registered. In all synopsis, the LR name, area of study and the threat risk level of each indicator of threat were recorded.

When applying the model presented by Porfiri *et al.* (2009), all the accessions were categorized as high risk level of erosion (Table 3.7.), within a score of 14 or 15.

Table 3.6.: Assessment of threat risk level of the selected LR using the *LR Threat Risk Calculator*.

LR ID	LR name	District	Municipality	Collection Site	Threat Risk Level
13	Amarelo	Azores Autonomous region - Sta. Maria island	Vila do Porto	St. Espírito	Data deficient (2.800 out of 3
20	Amarelo	Azores Autonomous region - Terceira island	Angra do Heroísmo	S. Sebastião	Data deficient (2.800 out of 3
21	Amarelo	Azores Autonomous region - Terceira island	Angra do Heroísmo	S. Sebastião	
87	Cor de carne	Azores Autonomous region - S. Miguel island	Vila da Povoação	Povoação	Data deficient (2.700 out of 3
37	Amarelo	Azores Autonomous region - S. Jorge island	Calheta de S. Jorge	Sto. Antão	Data deficient (2.600 out of 3
38	Amarelo	Azores Autonomous region - S. Jorge island	Calheta de S. Jorge	Sto. Antão	
43	Amarelo	Azores Autonomous region - S. Jorge island	Calheta de S. Jorge	Sto. Antão	
67	Amarelo	Azores Autonomous region - Flores island	Sta. Cruz	Sta. Cruz	Data deficient (2.375 out of 3
86	Amarelo	Azores Autonomous region - S. Miguel island	Vila da Povoação	Povoação	Maximum (2.727 out of 3)
89	da Ribeira Quente	Azores Autonomous region - S. Miguel island	Vila Franca do Campo	S. Miguel	Maximum (2.636 out of 3)
44	Amarelo	Azores Autonomous region - S. Jorge island	Calheta de S. Jorge	Ribeira Seca	Maximum (2.636 out of 3)
139	Canário	Aveiro	Oliveira de Azeméis	Pinheiro da Bemposta	Maximum (2.500 out of 3)
121	Mocho	Castelo Branco	Sertã	Sertã	Maximum (2.455 out of 3)
88	Cor de carne	Azores Autonomous region - S. Miguel island	Vila da Povoação	Furnas	Maximum (2.545 out of 3)
53	Amarelo	Azores Autonomous region - Pico island	São Roque	Sta. Luzia	Maximum (2.500 out of 3)
51	Amarelo	Azores Autonomous region - Pico island	Lajes do Pico	Piedade	Maximum (2.455 out of 3)
8	Amarelo	Azores Autonomous region - S. Maria island	Vila do Porto	Sta. Bárbara	Maximum (2.545 out of 3)
22	Amarelo	Azores Autonomous region - Terceira island	Angra do Heroísmo	Porto Judeu	Maximum (2.545 out of 3)
23	Amarelo	Azores Autonomous region - Terceira island	Angra do Heroísmo	Sta. Bárbara	Maximum (2.455 out of 3)
20M	Amarelo	Madeira Autonomous region	Sta. Cruz	Sta. Cruz	Maximum (2.545 out of 3)
143	Amarelo	Vila Real	Vila Real	Lamas de Olo	Maximum (2.500 out of 3)
41	Amarelo	Azores Autonomous region - S. Jorge island	Calheta de S. Jorge	Vila do Topo	Maximum (2.364 out of 3)
39	Amarelo	Azores Autonomous region - S. Jorge island	Calheta de S. Jorge	Vila do Topo	
6	da Ribeira Quente	Azores Autonomous region - S. Miguel island	Vila Franca do Campo	Ribeira Seca	Maximum (2.333 out of 3)
10	Amarelo	Azores Autonomous region - Sta. Maria island	Vila do Porto	Vila do Porto	Maximum (2.250 out of 3)
1M	Touquinho	Madeira Autonomous region	Calheta	Arco da Calheta	Minor (1.909 out of 3)

Table 3.7.: Assessment of threat risk level of the selected LR using Porfiri *et al*. (2009) method.

LR ID	LR name	District	Municipality	Collection Site	Threat Risk Level
13	Amarelo	Azores Autonomous region - Sta. Maria island	Vila do Porto	St. Espírito	14
20	Amarelo	Azores Autonomous region - Terceira island	Angra do Heroísmo	S. Sebastião	14
21	Amarelo	Azores Autonomous region - Terceira island	Angra do Heroísmo	S. Sebastião	
87	Cor de carne	Azores Autonomous region - S. Miguel island	Vila da Povoação	Povoação	15
37	Amarelo	Azores Autonomous region - S. Jorge island	Calheta de S. Jorge	Sto. Antão	14
38	Amarelo	Azores Autonomous region - S. Jorge island	Calheta de S. Jorge	Sto. Antão	
43	Amarelo	Azores Autonomous region - S. Jorge island	Calheta de S. Jorge	Sto. Antão	
67	Amarelo	Azores Autonomous region - Flores island	Sta. Cruz	Sta. Cruz	14
86	Amarelo	Azores Autonomous region - S. Miguel island	Vila da Povoação	Povoação	14
89	da Ribeira Quente	Azores Autonomous region - S. Miguel island	Vila Franca do Campo	S. Miguel	15
44	Amarelo	Azores Autonomous region - S. Jorge island	Calheta de S. Jorge	Ribeira Seca	14
139	Canário	Aveiro	Oliveira de Azeméis	Pinheiro da Bemposta	15
121	Mocho	Castelo Branco	Sertã	Sertã	15
88	Cor de carne	Azores Autonomous region - S. Miguel island	Vila da Povoação	Furnas	15
53	Amarelo	Azores Autonomous region - Pico island	São Roque	Sta. Luzia	14
51	Amarelo	Azores Autonomous region - Pico island	Lajes do Pico	Piedade	14
8	Amarelo	Azores Autonomous region - S. Maria island	Vila do Porto	Sta. Bárbara	14
22	Amarelo	Azores Autonomous region - Terceira island	Angra do Heroísmo	Porto Judeu	14
23	Amarelo	Azores Autonomous region - Terceira island	Angra do Heroísmo	Sta. Bárbara	14
20M	Amarelo	Madeira Autonomous region	Sta. Cruz	Sta. Cruz	14
143	Amarelo	Vila Real	Vila Real	Lamas de Olo	14
41	Amarelo	Azores Autonomous region - S. Jorge island	Calheta de S. Jorge	Vila do Topo	14
39	Amarelo	Azores Autonomous region - S. Jorge island	Calheta de S. Jorge	Vila do Topo	
6	da Ribeira Quente	Azores Autonomous region - S. Miguel island	Vila Franca do Campo	Ribeira Seca	15
10	Amarelo	Azores Autonomous region - Sta. Maria island	Vila do Porto	Vila do Porto	14
1M	Touquinho	Madeira Autonomous region	Calheta	Arco da Calheta	15

3.5 DISCUSSION

LR diversity is increasingly recognized as a critical resource for contemporary crop improvement and a means of objective threat assessment will significantly aid effective conservation planning and implementation (Veteläinen *et al.,* 2009). The intrinsic characteristics of LR, notably the range of diversity, non-standardized nomenclature, and lack of comprehensive national LR inventories, make them a challenging study object. To have a standardized method that assesses the threat risk of these plant genetic resources would help to organize effective conservation actions (Veteläinen *et al.,* 2009). The presented methodology has demonstrated such an approach to objective LR threat assessment that can be applied to a range of crops at national or sub-national level.

The use of the *LR Threat Risk Calculator* allows automatic calculation of LR threat risk level and the compilation of important information in the synopses, such as the risk level associated to a specific LR. The proposed methodology provides a means of LR threat assessment comparable with the application of IUCN Red List threat assessment for crop wild relatives and other wild taxa.

The threat assessment report, in addition to information given by the synopses as the LR threat risk level, willingness of the population for LR maintenance, abandonment of LR for new varieties, existence of certified product on the market (e.g. PDO, PGI), type of cultivation system in use (commercial sustainable farming system or high input farming system), should include other available data. For example, information about culture connected to certain species (Maxted *et al.,* 1997b) such as gastronomical uses of a crop LR; existence of successful rural tourism facilities and activities based on the presence of certain LR (Veloso, 2008); presence of activities promoting LR use and conservation by the nearest municipality (e.g. feijão.com(e) in São Pedro do Sul municipality); willingness of the municipalities to engage in conservation measures; among other accessible data, should be gathered by the researcher. It should also address the conservation activity and monitoring to be applied (Maxted *et al.,* 1997) and present possible measures to promote the use of LR such as seminars, workshops, "tasting fairs" where the promotion of LR knowledge would be adapted to the specific location being studied thus, enabling the decrease of LR threats (see Jarvis *et al.,* 2011 for an overview of possible measures to promote LR sustainable use). The cost of the conservation activities (Maxted *et al.,* 1997b) and their benefits should also be included. Subsequently, the researcher along with local and/or national stakeholders must evaluate all the information and verify if there is an infrastructure to sustainably conserve LR or the possibility to create one. Afterwards, within the monetary and human resources availability, the LR in study may be considered or not a priority for conservation.

It is also important to stress that the number of sites visited during field work and/or the gathered data should cover the territory in study. If covering the whole territory is impossible (e.g. due to financial constraints), then the use of ELC maps (Ecogeographic Land Characterization maps) (Parra-Quijano *et al.,* 2014) may be used to select the sites to survey. Alternatively, sub-national (village/town/municipality/district/region) studies can be carried out and subsequently all collated data integrated into a national level assessment.

Further, the *LR Threat Risk Calculator* and the target questions listed Table 3.2., should be consulted and used in the development of a questionnaire to be applied during a survey (see also Table 3.1. with calculations necessary for the use of *LR Threat Risk Calculator*). This facilitates the collection of sufficient raw data to generate the assessment. In our case-study, the questionnaire and the field survey were completed to assess potential threats to LR and determine criteria, and indicators of threats and develop a threat assessment methodology, and posterior integration of questions in the *LR Threat Risk Calculator* happened after field work.

The results obtained from the methodology developed in this work and from the method presented by Porfiri *et al.* (2009), were comparable. With both methods most of the accessions obtained the highest threat risk level (Maximum and High). Nonetheless, the first methodology had five accessions as Data Deficient – five or more questions were not answered by farmers or assessed by the researcher. Also, the accession 1M was categorized as Minor threat risk level. The differences observed between methodologies seem to be related to the level of detail given by the number of questions. The questionnaire and the questions used in the *LR Threat calculator* proposed in this work, allow the awareness of more detail concerning, for example, the LR maintainer and the conservation status of the LR. Thus, more thorough results.

Lastly, the measure(s) to be carried out in order to promote LR use and conservation will depend on the farmer and farmer' population availability to engage in the LR conservation (Jarvis *et al.,* 2011).

3.6 CONCLUSION

The proposed methodology presents a standardized format that can be used globally: in different countries, regions and with different crops. It would be of help in LR threat assessments and posterior LR red listing, which would be a support for LR conservation strategies. The conservation strategies should always encompass researchers; local, regional, national stakeholders and all farmers' community, to achieve success.

CHAPTER IV. ASSESSING THE GENETIC DIVERSITY OF A GROUP OF PORTUGUESE COMMON BEAN (*P. VULGARIS* L.) LANDRACES USING MICROSATELLITES

4.1 ABSTRACT

Phaseolus vulgaris L., common bean, is a general staple food with recorded importance. Its arrival through the Iberian Peninsula to Europe, combined with Portugal's ecogeography allowed the development of several landraces (LR) of this crop in the country. Farmers and plant breeding system can benefit from traits presented by LR as for example, resistance to pests and diseases and drought tolerance which, in the advent of climate change effects, is an added value favouring food security. Conservation strategies of LR should be thorough and can be enhanced by molecular studies. The study of population genetic diversity can support conservation activities, helping to establish where and what to conserve according to the diversity and population structure found. We used 11 microsatellites in order to assess the diversity and population structure of 274 samples: collected during field work throughout the Azores and Madeira archipelagos and mainland Portugal, and also four LR from Mexico, Colombia, Peru and Spain (named external LR), four commercial varieties with the same origin and two *Phaseolus coccineus* L. from Portugal and United Kingdom were also added to the study in order to be a control. Results showed Azores LR mainly in population one and three, the latter with fewer samples. Mainland and Madeira LR were mostly in population two. Population two and three LR were less genetic differentiated than population one and two.

Keywords: genetic diversity; landraces (LR); microsatellites; conservation strategies; Portugal.

4.2 INTRODUCTION

The nonwoody annual plant *Phaseolus vulgaris* L., common bean (2n=2x=22), is a largely self-pollinating species with documented economic, nutritional and agricultural relevance and is found broadly (Broughton *et al.*, 2003; Blair *et al.*, 2009; Câmara *et al.*, 2013). Its introduction in Europe through the Iberian Peninsula in the early 1500s, possible posterior re-introductions and circulation in European countries were complex movements (Angioi *et al.*, 2010). The genetic differences in the introduced common beans, different farmers' selection, and adaptation to diverse environments may be reasons for the range of common bean diversity (Asfaw *et al.*, 2009), and even why the Iberian Peninsula was considered the second centre of genetic diversity of common bean (Santalla *et al.*, 2002). In Portugal, common bean corresponds to around 76% of leguminous crops consumption (INE, 2017). For centuries, Portuguese farmers have been cultivating common bean especially for household consumption, originating great diversity of landraces (LR) (Leitão *et al.*, 2013). Likewise, LR diversity may be used in crop improvement programmes (e.g. search and use of traits of interest such as resistance to drought and resistance to pest and diseases) which, with the possibility of climate change effects, has significant value for food security and overall humankind welfare (Vetelainen *et al.*, 2009).

Conservation of LR can be *in situ* or/and *ex situ* (Maxted *et al.*, 1997). It ought to be, whenever possible, supported by molecular work since it provides thorough data about population diversity and structure. Thus, knowing a populations' genetic diversity may help conservation activities with for example the prioritization of LR (if time and money are a constraint) and improvement of LR threat assessments (Maxted *et al.*, 2013).

Molecular markers *"are genetic loci that can be easily tracked and quantified in a population and may be associated with a particular gene or trait of interest"* (Hayward *et al.*,

2015) and can be used in inter and intra-populations diversity studies (Gonçalves-Vidigal & Rubiano, 2011) presenting great variety [e.g. AFLP (Amplified fragment length polymorphism), RAPD (Random Amplified Polymorphic DNA), SSRs (microsatellites) and RFLP (Restriction Fragment Length Polymorphism)]. Nevertheless, SSRs which are short tandem DNA repeats, presents benefits: they are i) codominant: ii) highly polymorphic; iii) easily reproduced and demanding small quantities of DNA and iv) can be used across correlated species and are abundant in the genome (Moore *et al.*, 1991; Morgante & Olivieri, 1993; Saghai Maroof *et al.*, 1994).

There are examples of genetic analysis studies with microsatellites in wheat (e.g. Devos *et al.*, 1995), maize (e.g. Senior & Heun, 1993) and tomato (e.g. Vosman & Arens, 1997), among several others.

Studies that ascertain and describe genetic diversity of common bean have also been performed using microsatellites (e.g. Blair *et al.*, 2006; Blair *et al.*, 2009; Leitão *et al.*, 2017). Blair *et al.*, (2006) defined the genetic diversity and population structure of 43 common bean cultivars using 129 microsatellite markers. The differentiation among Mesoamerican and Andean common bean genepools were assessed as races within each genepool. Blair *et al.*, (2009) studied the population structure and genetic diversity of 604 accessions from Centro Internacional de Agricultura Tropical (CIAT) germplasm collection with 36 microsatellite markers. The results were generally comparable with other studies in terms of the race structure assessment of the germplasm collection. Leitão *et al.*, (2017) analysed 175 Portuguese common bean accessions in terms of diversity and structure with a group of 21 microsatellite markers. The studied accessions presented great genetic diversity and showed that most Portuguese accessions were more connected to Andean genepool than Mesoamerican.

This work aims to assess the genetic diversity among LR collected during field work in continental Portugal and the Azores and Madeira archipelagos using microsatellites, allowing the enhancement of future conservation strategies by providing information (as population structure and diversity parameters) that can be used in LR threat assessment (see Chapter III *A Threat Assessment Methodology for Crop Landraces*).

4.3 MATERIAL AND METHODS

4.3.1 Plant material

The material analysed in this work were collected during the field work undertaken in 2015 in continental Portugal and the Azores and Madeira archipelagos. A total of 274 records/accessions were analysed (272 of *P. vulgaris* LR and 2 of *Phaseolus coccineus* L.): 80 records from continental Portugal, 131 from Azores, 48 from Madeira archipelagos, five samples were randomly duplicated (collected second time for analysis), ten were requested from Centro Internacional de Agricultura Tropical (CIAT) [four LR and four commercial varieties from Mexico, Colombia, Peru and Spain, and two accessions of *P. coccineus*. from Portugal and United Kingdom (UK)]. The duplicates and CIAT accessions were used as a control to test the applied methodology.

By *sample(s)*/*record(s)* it is meant LR collected in each visited site/farm. Some visited sites provided one sample – one LR of *P. vulgaris* – whereas others provided more than one sample – more than one different LR of *P. vulgaris*. Regarding *accession*, it is meant to be the *P. vulgaris* LR, commercial varieties and *P. coccineus* asked to CIAT.

Ten seeds, randomly selected from each sample/accession, were germinated and the first leaves were collected for DNA extraction (Blair *et al*., 2009). The common bean seeds were germinated in germination trays in an equal mixture of peat and soil during August and

September 2016, in Banco Português de Germoplasma Vegetal (BPGV) facilities. The trays were placed under a shade screen, protected from direct sunlight and watered two times a day (Figure 4.1. and 4.2.).

Figure 4.1.: Germination trays.

Figure 4.2.: Detail of labelled germination trays under shade screen.

The first young leaves from each one of the 10 seeds cultivated per sample/accession were collected, weighted using an OHAUS Pioneer PA4102 Precision Balance and immediately placed to dry in a grip seal plastic bag with 0.5–1.0 mm granular self-indicating silica gel, orange to green (GeeJay Chemicals Ltd.). Silica gel is a low-priced product with good results in preserving leave material for molecular studies (e.g. Chase & Hills, 1991). The bags were identified with the ID number of each sample/accession (Figure 4.3.). After dried, an equivalent of 50 mg fresh weight was weighted using a KERN 770 Analytical Balance and each sample was placed in collection 1.2 ml microtubes (Figure 4.4. and 4.5.). Three plates of 96 wells each were filled with sampled material (third plate with 82 wells filled with sampled material) and sent to the Institute of Biological, Environmental and Rural Sciences (IBERS), Prifysgol Aberystwyth University, in Wales, for DNA extraction and analysis.

Figure 4.3.: Collected leaves drying in silica gel.

Figure 4.4.: BPGV laboratory where leaves were weighed.

Figure 4.5.: Sampled material in collection microtubes.

4.3.2 DNA extraction

The genomic DNA of all silica dried samples/accessions were extracted using the Qiagen DNeasy® 96 Plant kit (Cat. no. 69181) and eluted using 200 µl of AE Buffer as per protocol.

One µl of the genomic DNA extract was run on a 1% Agarose (Melford MB1200)/x1 TAE buffer (VWR K915-1.6L X50 TAE diluted 50 ml to 2.5 L to give x1 TAE), containing GelRed® nucleic acid stain 4.5 µl to 90 ml (Biotium 41003) for genomic DNA quality verification. The gel was visualized using the G:Box Syngene transilluminator and Syngene software.

4.3.3 PCR conditions

The PCR Master mix (15 µl PCR reaction) had 7.5 µl ImmoMix™ (Bioline BIO-25020); 5 µM of each forward and reverse primers (total 1.0 µl); 4.5 µl sterile distilled water and 2.0 µl of genomic DNA.

The PCR conditions were: 1 cycle of 95ºC for 10 minutes; and 35 cycles of denaturation at 95ºC for 1 minute, annealing at 69ºC or 60ºC dependent on annealing temperature of primer (Table 4.1.) for 1 minute and final extension at 72ºC for 10 minutes, final hold at 12ºC.

4.3.4 Primers selection and testing

The selection of microsatellite markers was based on a literature search. A total of 19 common bean primers were selected and tested based on their performance, namely those presenting good results: the highest Polymorphic Information Content (PIC) and referred in more than two research papers (Table 4.1.).

The amount of 10 µl of PCR product plus loading dye were run on a 1.5% agarose gel (4.5 g agarose/300ml x! TAE) containing 15 µl of GelRed® and 100bp DNA ladder (Promega G210A). The gel was visualized using the G:Box Syngene transilluminator and Syngene software. Results were assessed for the presence of DNA good amplification, in order to determine which primers to fluorescently label and PCR multiplex for fragment analysis. Sixteen primers were finally selected. The other three did not present good amplification (Table 4.1.).

Table 4.1.: Label, multiplex group, linkage group (LG), forward and reverse sequence and product size of the selected markers.

Marker	Label	Multiplex plate	LG	Anneling temp.(̊C)	Sequence forward	Sequence reverse	Product size (bp)	References
AG1	FAM	3	3	69	CATGCAGAGGAAGCAGAGTG	GAGCGTCGTCGTTTCGAT	120 to138	Blair *et al* . (2006; 2009); Diaz *et al* . (2011); CIAT (2006)
BM53	FAM	3	1	60	AACTAACCTCATACGACATGAAA	AATGCTTGCACTAGGGAGTT	-	Blair *et al* . (2006); CIAT (2006); Lopes *et al* . (2007)
BM151	NED	2	8	60	CACAACAAGAAAGACCTCCT	TTATGTATTAGACCACATTACTTCC	138 to 152	Blair *et al* . (2006); CIAT (2006)
BM152	FAM	1	2	69	AAGAGGAGGTCGAAACCTTAAATCG	CCGGGACTTGCCAGAAGAAC	-	Blair *et al* . (2006); CIAT (2006); Lopes *et al* . (2007)
BM187	PET	3	6	60	TTTCTCCAACTCACTCCTTTCC	TGTGTTTGTGTTCCGAATTATGA	-	Okii *et al* . (2014); Blair *et al* . (2006; 2009); Diaz *et al* . (2011); CIAT (2006)
BM210	VIC	2	7	60	ACCACTGCAATCCTCATCTTTG	CCCTCATCCTCCATTCTTATCG	159 to 185	Angioi *et al* . (2008); Blair *et al* . (2006); CIAT (2006); Lopes *et al* . (2007)
BMd1	PET	1	3	60	CAAATCGCAACACCTCACAA	GTCGGAGCCATCATCTGTTT	164 to 192	Blair *et al* . (2003; 2006; 2009); Diaz *et al* . (2011); Gioia *et al* . (2013); CIAT (2006); Lopes *et al* . (2007)
BMd15	NED	3	4	60	TTGCCATCGTTGCTTAATTG	TTGGAGGAAGCCATGTATGC	165 to 201	Blair *et al* . (2003; 2006; 2009); Diaz *et al* . (2011); CIAT (2006); Lopes *et al* . (2007)
BMd37	FAM	2	6	60	GGCACGAGCAACAATCCTT	CCATCATAGAGGGCAACCAC	126 to 140	Blair *et al* . (2003; 2006); Okii *et al* . (2014); CIAT (2006)
BMd41	FAM	2	11	60	CAGTAAATATTGGCGTGGATGA	TGAAAGTGCAGAGTGGTGGA	227 to 250	Gioia *et al* . (2013); Okii et al. (2014); Blair *et al* . (2003; 2006); CIAT (2006); Lopes *et al* . (2007)
BMd42	VIC	3	10	60	TCATAGAAGATTTGTGGAAGCA	TGAGACACGTACGAGGCTGTAT	149 to 162	Gioia *et al* . (2013); Okii *et al* . (2014); Blair *et al* . (2003; 2006); CIAT (2006)
GATS91	PET	2	2	69	GAGTGCGGAAGCGAGTAGAG	TCCGTGTTCCTCTGTCTGTG	218 to 263	Blair *et al* . (2006; 2009); Diaz *et al* . (2011); CIAT (2006); Lopes *et al* . (2007)
PVat001	FAM	1	4	69	GGGAGGGTAGGGAAGCAGTG	GCGAACCACGTTCATGAATGA	-	Blair *et al* . (2006; 2009); Diaz *et al* . (2011); Yu *et al* . (1999; 2000); CIAT (2006)
PVat007	NED	2	9	69	AGTTAAATTATACGAGGTTAGCCTAAATC	CATTCCCTTCACACATTCACCG	190 to 220	Blair *et al* . (2006); Yu *et al* . (1999; 2000); Okii *et al* . (2014); CIAT (2006)
PVcct001	VIC	1	2	69	CCAACCACATTCTTCCCTACGTC	CGCAGGCAGTTATCTTTAGGAGTG	-	Blair *et al* . (2006; 2009); Yu *et al* . (2000); CIAT (2006)
PVctt001	NED	1	4	69	GAGGGTGTTTCACTATTGTCACTGC	TTCATGGATGGTGGAGGAACAG	152 to 165	Blair *et al* . (2006; 2009); Yu *et al* . (1999; 2000); CIAT (2006); Lopes *et al* . (2007)

4.3.5 Fragment analysis

The samples were then grouped according to the fluorescent label and expected size of the fragment, forming 3 multiplexes (Table 4.1.). 1 µl of each PCR product dilution (1:40) was added to 10 µl of formamide containing GeneScan™ (-250) LIZ® standard and ran through the ABI 3730 DNA Analyzer (Applied Biosystems™) using a 48 capillary 50 cm array (Applied Biosystems 4331250). The ABI 3730 DNA Analyzer (Applied Biosystems™) conditions were set using the default GeneMapper 50_POP7 protocol: oven 63°C; buffer temperature 35°C; injection time 1.6 seconds; Rtn time 1600 seconds; run voltage 15; POP7 Performance Optimised Polymer (Applied Biosystems 4335615), buffer with 1X concentration of 3730 Buffer (X10) with EDTA (Applied Biosystems 4335613), DS-33 (Dye Set G5), Matrix STD (Applied Biosystems 4345833).

4.3.6 Data analysis

The GeneMarker® v.2.6.7 software was used for DNA fragment analysis. The electropherograms were analysed and an allele report table, with allele size for each marker and sample, was produced. After the fragment analysis we discard 5 markers (Table 4.1.) owed to bad amplification in the majority of samples.

The software Structure Version 2.3.4 (Pritchard *et al.*, 2000; Falush *et al.*, 2003, 2007; Hubisz *et al.*, 2009) was employed to infer about population structure of the 274 samples/ accessions used in this work. The parameters were set considering admixture and correlated frequency models.

With GenAlEx 6.5 software (Peakall and Smouse 2006, 2012) the number of alleles per locus (*Na*), number of private alleles, observed and expected heterozygosity (*Ho* and *He*), allele frequency, percentage of polymorphic loci, F-statistics calculations (Wright, 1978), Nei's genetic distance (D) and also Principal Coordinate Analysis (PCoA) were assessed.

An Unweighted Pair Group Method with Arithmetic Mean (UPGMA) was used and a script on R software (Paradis *et al.*, 2004; Lemon, 2006; Jombart, 2008, 2011; Guz, 2014; Kamvar *et al.*, 2014, 2015; Galili, 2015; Maechler *et al.*, 2016) to build a dendrogram based on Nei's genetic distance (see section 4.3.7.).

PIC was calculated according to the formula PIC= $1-\sum p^2_{ij}$ (Blair *et al.*, 2006), with p_{ij} being the frequency of allele j for marker i.

4.3.7 Description of the script

The following R script was created for the dendrogram development. The hashtags (#) have the explanation of the following code:

```
# In the R library:
# install.packages(c("ape", "cluster", "dendextend"))
# https://stats.stackexchange.com/questions/4062/how-to-plot-a-fan-polar-dendrogram-in-r
# http://www.stat.columbia.edu/~tzheng/files/Rcolor.pdf

library(ape)
library(cluster)
library(dendextend)
library(poppr)
library(adegenet)
library(plotrix)

# input
data.filename = "data_genalex_3pop_v9.csv"
markers.filename = "markers.csv"
colors.filename = "color_dup.csv"
dendrogram.filename = "dendrogram_data_genalex_3pop_v9_markers_color_dup.png"

# to generate data.csv (preparing the data files):
# * export merge table data as csv
# * remove rows above marker column labels
# * duplicate marker column labels
# * simplify and tify up other column labels
data = read.csv(data.filename, sep=";")
```

```
markers = read.csv(markers.filename)
for (i in rownames(markers)) {
   marker = as.character(markers[i,])
   marker.1 = paste(marker, ".1", sep="")
   data[,marker] = paste(data[,marker], data[,marker.1], sep=",")
   data[,marker.1] = NULL
}
ind.names = data$Sample
data$Sample = NULL
data$Pop = NULL
genind.obj = df2genind(data, sep=",", ind.names=ind.names,
loc.names=unlist(markers$Name), NA.char="0,0", type="codom")

colors = read.csv(colors.filename, header=T, sep=";")
hc = aboot(genind.obj, tree="upgma", dist=nei.dist, showtree=F)

hc.dendrogram = as.dendrogram(hc)

i=0
add.colors<<-function(n){
  if(is.leaf(n)){

# Take the current attributes
   a=attributes(n)

# Deduce the line in the original data, and so the treatment and the specie.
   ligne = pmatch(attributes(n)$label,colors[,1])
   color = as.character(colors[ligne, 2])

# Modification of leaf attribute
   attr(n,"nodePar")<-
c(a$nodePar,list(cex=1.5,lab.cex=1,pch=20,col=color,lab.col=color,lab.font=1,lab.cex=1))
  }
  return(n)
}
hc.dendrogram <- dendrapply(hc.dendrogram, add.colors)

plot.width = 2480
plot.height = 3508
png(dendrogram.filename, width=plot.width, height=plot.height)
par(mar=c(5,5,5,40) + .1)
plot(hc.dendrogram, horiz=T, xlim=c(1.5, 0), axes=F)
axis(1, at=seq(0, 2, by=.1))
dev.off()
```

4.4 RESULTS

The 274 samples were divided in 3 populations according to Structure Version 2.3.4 (Pritchard *et al.*, 2000; Falush *et al.*, 2003, 2007; Hubisz *et al.*, 2009) (Table 4.2.). Five samples were randomly duplicated (picked second time for analysis), all from the Azores archipelago.

Table 4.2.: Populations' size.

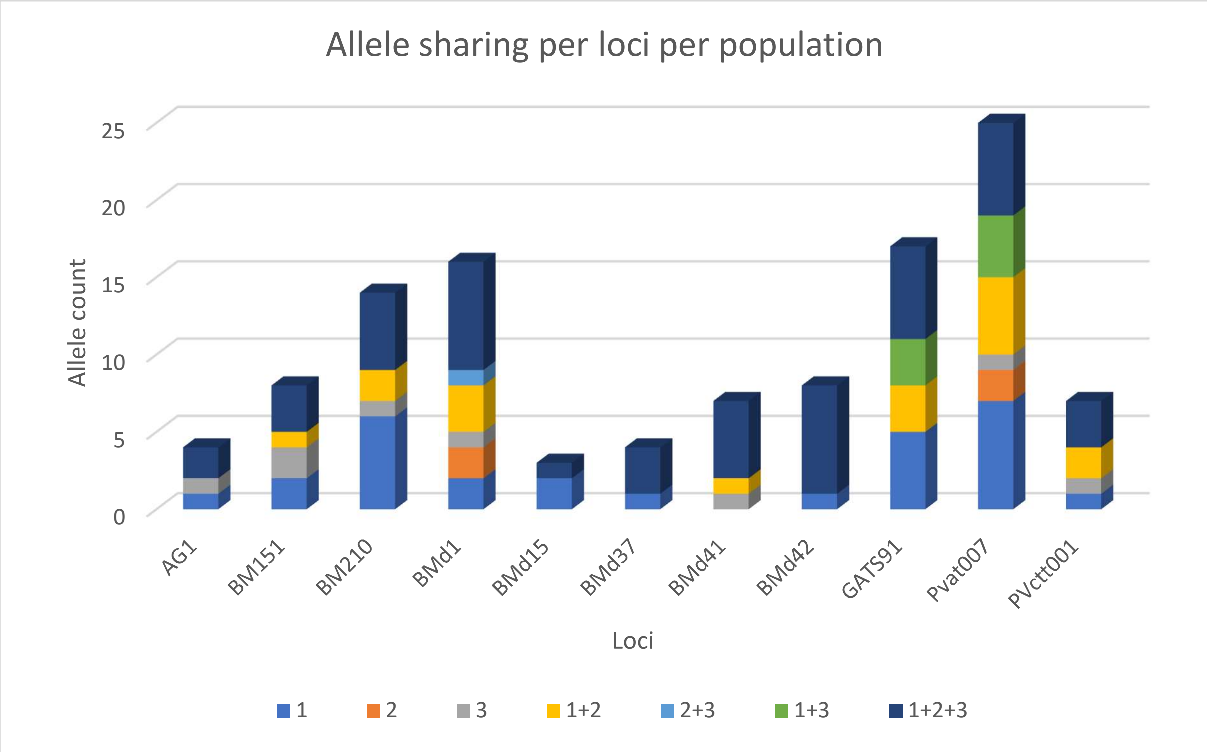
Population	**N**
1	154
2	62
3	58
Total	274

After fragment analysis with GeneMarker® v.2.6.7 software, markers BM53, BM152, BM187, PVat001 and PVcct001 were not used in the analysis due to poor amplification, leaving 11 markers for further analysis (Table 4.1.).

Markers AG1 and BMd1 are in linkage group (LG) three, and markers BMd15 and PVctt001 in LG four; whereas all other markers are in different LG (Table 4.1.).

The alleles' frequency and sample size per population is presented in Supplementary Table 4.3. (see also Supplementary Table 4.4.). Population one had 28 exclusive (private) alleles, while population two and population three LR presented 4 and 8 exclusive alleles, respectively (Supplementary Table 4.5.). The graph in Figure 4.6. represents the shared alleles per loci per population as the alleles present in only one or two LR populations. PVat007 was the locus with the highest number of shared alleles between populations whereas BMd15 had the lowest number of shared alleles (Figure 4.6.).

Figure 4.6.: Allele sharing per loci per population graph.



The mean of the total number of alleles per locus was 10 alleles, nevertheless it was a very variable number, ranging from 3 to 25 alleles per locus. Marker GATS91 had the highest PIC value in population one, while BMd1 and PVat007 had the highest PIC value in population two and three, respectively. The marker with the lowest PIC value was BMd15 in all populations (Table 4.6.). Population one had the highest mean PIC value.

Table 4.6.: Polymorphic information content (PIC) per locus per population (1, 2 and 3) and F-statistics and total number (n.) of alleles per locus.

Locus	PIC			Total n. alleles	All populations		
	1	2	3		Fis	Fit	Fst
PVctt001	0.609	0.379	0.452	7	0.907	0.910	0.030
BMd1	0.833	0.834	0.785	16	0.816	0.819	0.017
BMd37	0.530	0.529	0.444	4	0.906	0.912	0.064
BMd41	0.551	0.250	0.321	7	0.903	0.908	0.059
BM210	0.678	0.547	0.610	14	0.951	0.955	0.076
BM151	0.605	0.665	0.637	8	0.961	0.962	0.023
Pvat007	0.761	0.729	0.817	25	0.779	0.784	0.020
GATS91	0.875	0.687	0.773	18	0.847	0.856	0.056
AG1	0.312	0.498	0.404	4	0.758	0.774	0.066
BMd42	0.733	0.766	0.737	8	0.938	0.939	0.017
BMd15	0.088	0.000	0.000	3	0.920	0.922	0.024
Mean	0.598	0.535	0.544	10	0.881	0.885	0.041
SE					0.021	0.020	0.007

The mean total number of alleles per population was 78 with values ranging from 64 to 100. Population one LR had 100% polymorphic loci, while population two and three had 90,91% each, with a standard error of 3,03% and mean of 93,94% (Table 4.7.).

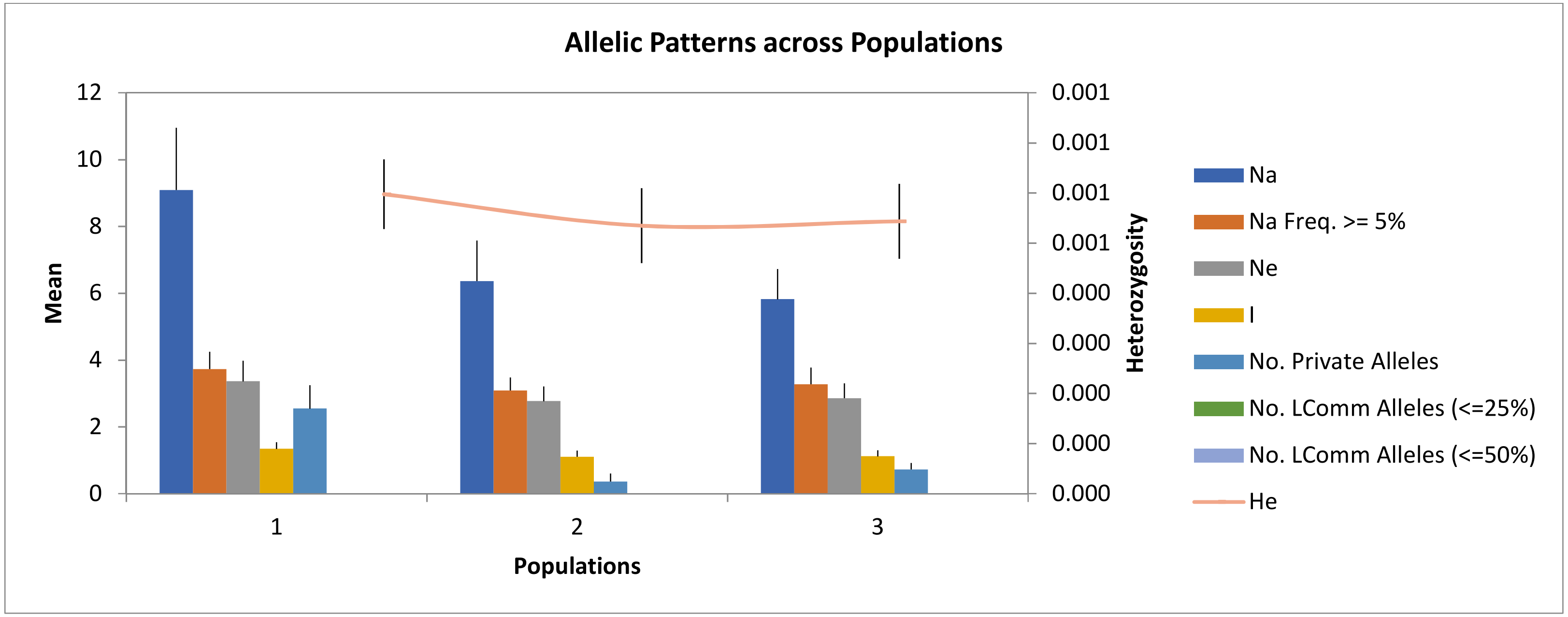
The mean of the number of different alleles (*Na*) with a frequency ≥ 5% was higher in LR from population one, followed by population three LR. The mean of the number of private alleles was also higher in population one LR (Figure 4.7. and Table 4.7.).

Table 4.7.: Mean allelic patterns and standard error (SE) across studied populations. Total number of alleles and percentage of polymorphic loci per population.

Population	**1**	**SE**	**2**	**SE**	**3**	**SE**
Na	9.091	1.866	6.364	1.216	5.818	0.913
Na Freq. >= 5%	3.727	0.524	3.091	0.392	3.273	0.506
Ne	3.362	0.618	2.768	0.446	2.856	0.450
I	1.341	0.202	1.106	0.186	1.119	0.181
No. Private Alleles	2.545	0.705	0.364	0.244	0.727	0.195
No. LComm Alleles (<=25%)	0.000	0.000	0.000	0.000	0.000	0.000
No. LComm Alleles (<=50%)	0.000	0.000	0.000	0.000	0.000	0.000
He	0.598	0.069	0.535	0.075	0.544	0.075
uHe	0.600	0.070	0.539	0.075	0.549	0.076
Percentage of polymorphic loci	100.000		90,91		90,91	
Total n. alleles	100		70		64	

Na	No. of Different Alleles
Na (Freq >= 5%)	No. of Different Alleles with a Frequency >= 5%
Ne	No. of Effective Alleles = 1 / (Sum pi^2)
I	Shannon's Information Index = -1* Sum (pi * Ln (pi))
No. Private Alleles	No. of Alleles Unique to a Single Population
No. LComm Alleles (<=25%)	No. of Locally Common Alleles (Freq. >= 5%) Found in 25% or Fewer Populations
No. LComm Alleles (<=50%)	No. of Locally Common Alleles (Freq. >= 5%) Found in 50% or Fewer Populations
He	Expected Heterozygosity = 1 - Sum pi^2
uHe	Unbiased Expected Heterozygosity = (2N / (2N-1)) * He

Figure 4.7.: Allelic patterns across the studied populations.



Na	No. of Different Alleles
Na (Freq >= 5%)	No. of Different Alleles with a Frequency >= 5%
Ne	No. of Effective Alleles = 1 / (Sum pi^2)
I	Shannon's Information Index = -1* Sum (pi * Ln (pi))
No. Private Alleles	No. of Alleles Unique to a Single Population
No. LComm Alleles (<=25%)	No. of Locally Common Alleles (Freq. >= 5%) Found in 25% or Fewer Populations
No. LComm Alleles (<=50%)	No. of Locally Common Alleles (Freq. >= 5%) Found in 50% or Fewer Populations
He	Expected Heterozygosity = 1 - Sum pi^2
uHe	Unbiased Expected Heterozygosity = (2N / (2N-1)) * He

The observed heterozygosity (H_o) was always lower than the expected heterozygosity (H_e) (Supplementary Table 4.8.) except for marker BMd15 in population two and three, which have both H_o and H_e equal to zero. The values of F-statistics are listed in Table 4.6. and pairwise population F_{ST} values in Table 4.9.. The mean value of F_{IS} is 0.881 which suggests high levels of inbreeding within population. The F_{ST} value over all populations (0,041; Table 4.6.) translated little genetic differentiation between populations, which indicates high genetic flow amongst populations. Table 4.9. contains pairwise populations F_{ST} values and according to those values, all populations presented little genetic differentiation (Wright, 1978). Nevertheless, populations one and two are the most genetic differentiated, while populations two and three are the less genetic differentiated. Nonetheless, and according to Meirmans (2006), the standardize measure of genetic variation F'_{ST} takes into account the markers mutation rates and different population effective dimensions. The F'_{ST} value obtained in the present work was 0,126. Thus, populations one, two and three may have moderate genetic differentiation.

Table 4.9.: Pairwise population F_{ST} values.

Populations 1	2	3	
0.000			**1**
0.043	0.000		**2**
0.033	0.015	0.000	**3**

Table 4.10.: Pairwise population matrix of Nei Genetic Distance.

Populations 1	2	3	
0.000			**1**
0.119	0.000		**2**
0.096	0.034	0.000	**3**

As presented in the pairwise population matrix of Nei´s genetic distance (D) (Table 4.10.), population one was the most genetically distant population from all other populations (population two D=0,119; population three D=0,096). Nonetheless, population one and two LR were the most dissimilar (D=0,119). Population two and three LR were the less dissimilar populations (D=0,034).

A PCoA on Nei's genetic distance was performed using GeneAlEx 6.5 in order to interpret the dissimilarities across data, which in our case supported the F-statistics findings: population one with the greater genetic differentiation from all populations (Figure 4.8.), and populations two and three LR with little genetic differentiation between them (Figure 4.8.). The first two coordinates correspond to 100% of the variation across data with the first coordinate responsible for 80,4% of the variation.

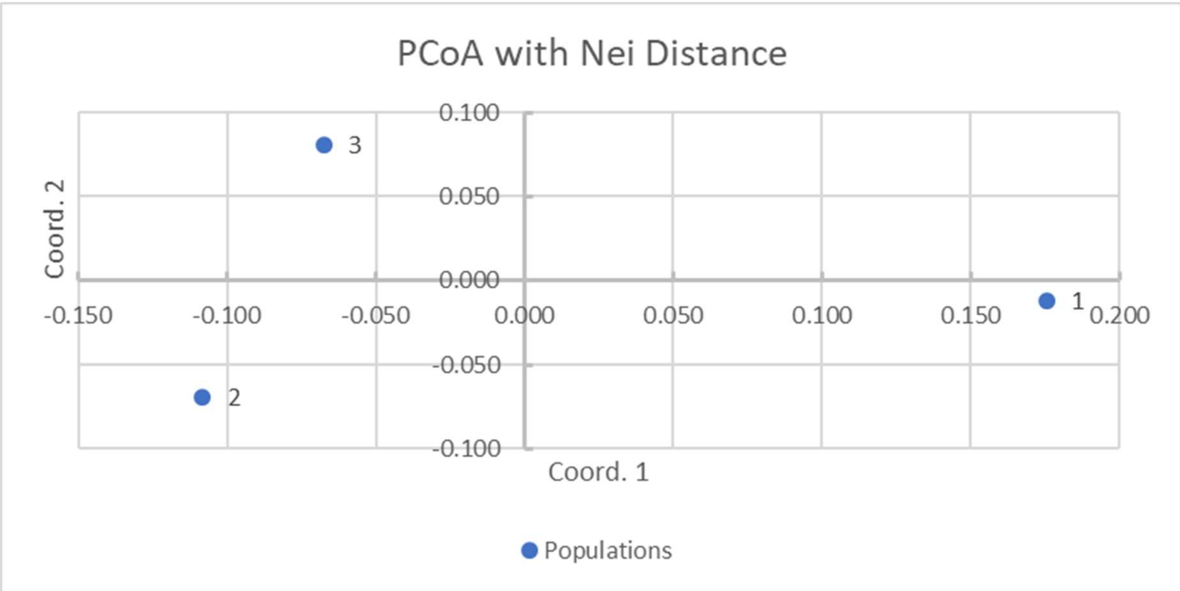


Figure 4.8.: Principal coordinate analysis (PCoA) of the 3 populations, developed using Nei's genetic distance values matrix.

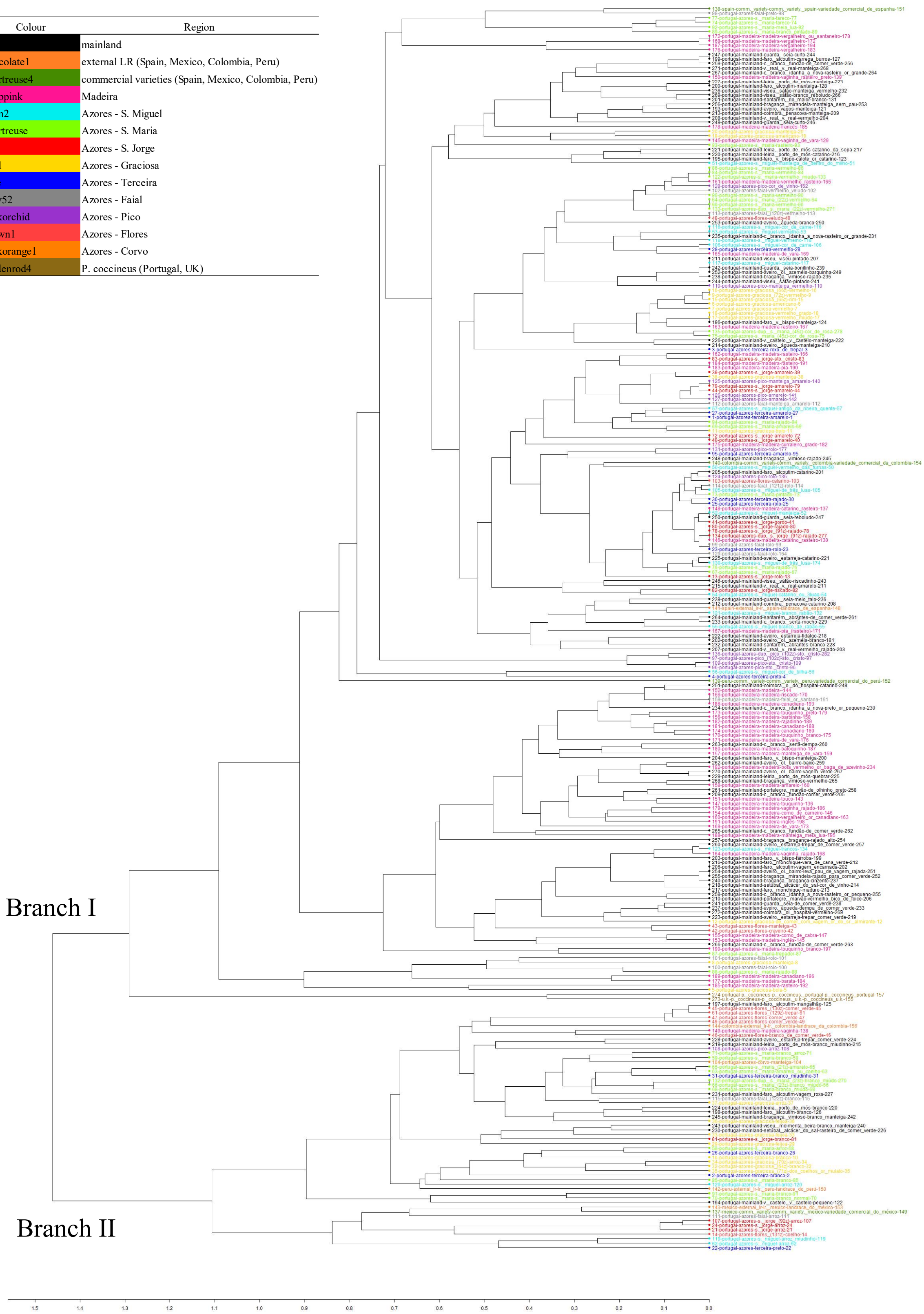
The UPGMA dendrogram allows us to observe the associations between the studied samples/accessions (Figure 4.9a.,b. and Supplementary Figure 4.10a.,b.). We used different colours for each of the 3 populations (Figure 4.9a), for the mainland and for each Azores and Madeira islands (Figure 4.9b), so the dendrogram would be more comprehensible.

Figure 4.9a.: UPGMA dendrogram based on Nei's genetic distance for the 274 studied samples (see CD at the end of this thesis). Different populations highlighted with different colours.

Colour	Population
deeppink	1
gold	2
blue	3

Figure 4.9b.: UPGMA dendrogram based on Nei's genetic distance for the 274 studied samples (see CD at the end of this thesis). Different regions' samples highlighted with different colours.

Colour	Region
	mainland
chocolate1	external LR (Spain, Mexico, Colombia, Peru)
chartreuse4	commercial varieties (Spain, Mexico, Colombia, Peru)
deeppink	Madeira
cyan2	Azores - S. Miguel
chartreuse	Azores - S. Maria
red	Azores - S. Jorge
gold	Azores - Graciosa
blue	Azores - Terceira
gray52	Azores - Faial
darkorchid	Azores - Pico
brown1	Azores - Flores
darkorange1	Azores - Corvo
goldenrod4	P. coccineus (Portugal, UK)



Population one comprises 154 samples/accessions; population two assembles 62, while population three has 58 samples/accessions.

Population one is formed by 60% of Azores LR samples, 27% mainland LR samples and 11% from Madeira. It also has four accessions requested from CIAT: one *P. vulgaris* LR from Spain, and three commercial varieties from Spain, Peru and Colombia.

Population two assembled 47% of mainland and 47% of Madeira LR samples. It only comprises 6% of samples from Azores.

Population three has 69% of LR samples from Azores, 17% from mainland and 3% from Madeira. It also presents three LR accessions from Mexico, Colombia and Peru, a commercial variety from Mexico and two *P. coccineus*, requested from CIAT.

There are two major branches in the dendrogram (Figure 4.9a.): branch I groups population one and two LR samples, as *P. coccineus* (UK and Portugal), whereas branch II groups population three samples. In total, branch two groups fewer number of samples. Branches shows groups of samples from the same region (e.g. Azores, Graciosa island) together however, we may not detect groups with all the samples from one region. As expected, *P. coccineus* from Portugal and the UK are in a separated branch in group I (Figure 4.9b. goldenrod4 colour). The majority of Azores' LR samples are in population one and three while population two has the larger number of mainland and Madeira samples.

Azores LR samples in population one are from all islands, except Corvo. They are mostly from S. Maria (20), S. Miguel (15), Graciosa (13) and S. Jorge (12) islands. In population three, Azores LR samples are also from all islands, essentially from S. Maria (11), Graciosa (8) and Flores (8).

Population one mainland LR samples are mainly from the North and Centre part of the country. The Southern LR samples are predominantly in population two.

Within branch I there are the majority of sampled Portuguese LR populations and the LR from Spain, whereas branch II present fewer number of sampled Portuguese LR and the LR from Mexico, Colombia and Peru. The duplicated samples were grouped with the original sample.

4.5 DISCUSSION

Analysing all populations together, the number of private alleles was higher in population one LR, followed by population three LR. This might be due to the fact that population one and three had higher number of samples collected from all Azorean islands, a wide area, which might have led to the gathering of more samples with more distinct, different genetic patterns caused by different farmers' selection and preferences (Asfaw *et al.*, 2009).

The mean value of F_{IS} indicated high levels of inbreeding within population and the pairwise F_{ST} values showed little genetic differentiation between populations, which may be caused by high genetic flow amongst populations. Nevertheless, the F'_{ST} value (0,126) indicates moderate genetic differentiation.

The observed heterozygosity was always lower than the expected heterozygosity (except for two markers which have both zero value) which is in agreement with the levels of inbreeding of the studied populations and as expected for a self-pollinating species as mainly common bean is (Blair *et al.*, 2009).

The number of alleles of the marker PVctt001, BMd15 and AG1 in population one, and BMd1 and AG1 in population three were similar to the work presented by Díaz *et al.*, (2011). The markers BMd41, BMd42, PVat007 and GATS91 had allele numbers comparable to the work of Blair *et al.*, (2006), whereas marker PVctt001 allele number in population one was similar to the work of Blair *et al.*, (2009).

The majority of markers were in different linkage groups (LG), which led to an assessment of diversity based on different genomic regions.

Migration is part of Portuguese culture for a very long time. These movements of people have structural influence in society and culture (Baganha *et al.*, 2005). With the calculations of Nei's genetic distance it was clear that population two LR were more closely related to population three than to population one LR. One reason for this dissimilarity may be related to the migration patterns, which may have caused different LR distribution thus, possible genetic differentiation.

Additionally, the result presented by the dendrogram shows the majority of the Portuguese LR samples and the Spanish LR in branch I (Figure 4.9b.) and the smaller amount of sampled Portuguese LR and the Mexican, Colombian and Peruvian LR in branch II (Figure 4.9b.). These results and considering the Iberian Peninsula as a second centre of genetic diversity of common bean (Santalla *et al.*, 2002), might be the cause the majority of samples being more dissimilar nowadays to the samples of the centre of origin. Thus, branch one with the higher number of LR and branch two with the smaller amount of Portuguese LR and the Mexican, Colombian and Peruvian LR.

Finally, the two major branches presented by the dendrogram are in accordance with other studies (e.g. Ligarreto & Ocampo, 2012), and might be attributable to the fact that the samples can be part of the Andean or the Mesoamerican genepool (e.g. Leitão *et al.*, 2017).

4.6 CONCLUSION

Microsatellite markers have been using as a tool in population genetic diversity and structure studies. Here, we used 11 markers to assess the genetic diversity of representative sample of common bean LR from Portugal. The heterozygosity parameter obtained with this analysis can be used when assessing threat of the studied species (see Chapter III *A Threat Assessment Methodology for Crop Landraces*). Despite the level of inbreeding, populations one and two and one and three still present a degree of genetic distance (D=0.119 and 0.096 respectively) thus, new collecting missions should be organized and systematically collect. Particularly in the archipelagos, so this genetic diversity would be sampled and represented in gene banks. At the same time, studies should be undertaken to assess the better sites to promote on-farm conservation based on the diversity found.

CHAPTER V. GENERAL DISCUSSION

5.1 SUMMARY

The need to conserve Plant Genetic Resources (PGR) specifically landraces (LR) to contribute to future food security, leads us to need to develop, improve and apply procedures that will help LR conservation. Procedures as standardized methodologies for LR inventories and threat assessment should be user-friendly and standardized, so knowledge and information can be exchanged and understood universally.

Portugal is signatory of various treaties that promote conservation and sustainable use of PGR such as LR. Comprehensive inventories should be developed as a first step to attain the treaties goals. Similarly, assessing the threat of national LR would help not only to prioritize them for conservation action but also to apply more efficient and targeted conservation measures. The molecular work undertaken in this research facilitated the gathering of *P. vulgaris* LR populations information and supported threat assessment and conservation planning. Therefore, our work is timely and helps to meet the goals and the practical application of the Portuguese National Strategic Action Plan, a document where guidelines for conservation and sustainable use of PGR are delineated in accordance with international treaty obligations.

5.2 INVENTORY

Developing a comprehensive inventory is a difficult task due to background information being scattered, sometimes scarce thus collation of data takes a substantial amount of time; LR and LR information are connected to human factors and their vicissitudes which may create obstacles to implementation of standardized methodologies of LR conservation; nomenclature and synonyms/homonyms are also a difficulty when establishing LR conservation strategies, due to the genetic distinction and attribution of LR accessions.

With the development of the Portuguese inventory we gathered information in a wide period range and developed a LR inventory for Portugal with 14,813 LR records for 123 crops. Nonetheless, implementing the distinct LR criteria (same crop LR that have the same name and collecting site are considered the same LR, and different LR that have different names or same name, but different collecting sites are considered different LR), the study identified 7,492 distinct LR within 36 families, 88 genera, and 130 taxa. The four families with the highest number of gathered LR' records were *Fabaceae*, *Vitaceae*, *Poaceae* and *Rosaceae*. Bragança (855), Faro (555), Aveiro (537) and Viseu (467) districts of Portugal presented the highest number of LR with the assembled data in mainland Portugal. On Madeira island 333 distinct LR were identified and on the Azores archipelago, Terceira island (101) had the highest number of distinct LR, followed by S. Miguel and Graciosa islands (both with 95) and Pico island (86). The most represented crop group were the leguminous crops followed by vegetables and melons group, according to FAO Indicative Crop Classification version 1.0 (FAO, 2005). Those records from undisclosed or with inaccurate origin (e.g. Portugal), thus not used in the analysis, were kept in the database.

This will help future collecting missions to be more thorough and systematic, to monitor LR diversity and possible threats as to promote the use of LR by farmers and breeders. However, the inventory should never be a static work not only because LR are continually evolving and changing (Negri, 2005) but also because conditions where they are maintained, as well as the maintainers, will also change over time.

5.3 THREAT ASSESSMENT

Threat assessment for wild species is well established with IUCN Red Listing and Criteria (IUCN, 2001). However, for LR, there is no standardized methodology used/applied globally, that could address the assessment of threats at infra-specific level. The field work carried out and questionnaires applied allowed the observation and register of threats to LR permitting the development of a standardized methodology that can be applied globally. The results of the application of the methodology, with the threat assessment of a sub-set of 26 common bean LR collected during field work, presented LR mostly with maximum threat risk level. The results are comparable with the work of Porfiri *et al*. (2009) although, presenting more detail.

The same pattern of abandonment and disappearing of LR reported in other works (e.g. Negri *et al.*, 2003, 2005) can be confirmed. Generally, the reasons for LR abandonment are due to (a) LR maintainers' ageing; (b) the young maintainers are not exclusively farmers, with their main source of income off-farm, which leads to lack of time to maintain seeds and LR cultivation traditions, and (c) lack of LR knowledge by young maintainers allied to the easy acquisition of modern varieties in stores. Nonetheless, the interaction between farmers and researchers does exist, particularly in the Madeira and Azores archipelagos and should be preserved and enhanced in all Portuguese territory due the importance of the relationship to promote and sustain conservation strategies.

Lastly, the LR threat assessment should be reviewed and performed regularly. Such iterative assessment should be repeated triennially to monitor conservation success or the development of novel threats, but this period might by extended or shortened depending on the results from previous assessments, discussion held with farmers and researchers and with the necessity of the country where the assessment is being carried out.

5.4 MOLECULAR WORK

For the molecular study of *P. vulgaris* LR, samples collected during field work were analysed and genetic diversity parameters, F-statistics and genetic distance were estimated.

High levels of inbreeding in the studied populations were observed and according to Nei's genetic distance, population two LR were more closely related to population three LR than to population one. Population two samples were predominantly from mainland and Madeira, while population three had a smaller number of samples mostly from Azores. The genetic results can help guide future collecting missions. With the genetic differences assessed in Portuguese LR populations and the LR entries in the inventory, researchers may establish where is most needed to collect material for *ex situ* conservation. Likewise, this work may improve LR threat assessment, with useful information about the studied LR, as the heterozygosity parameter.

5.5 RECOMMENDATIONS

Given the results obtained in this research and presented in the several chapters, here are advanced some recommendations:

- The districts with the highest number of LR with the gathered data from inventory development, namely Bragança (855), Faro (555), Aveiro (537) and Viseu (467) in mainland Portugal, should be assessed for possible on-farm conservation implementation. For example, farmers' socio-economic situation, willingness for LR maintenance, and ecogeographic characteristics (e.g. diverse geographical characteristics) of potential locations should be studied and weighed to establish the sites with better conditions for on-farm conservation of LR.
- The same should also be applied to Madeira and Azores archipelagos. Initially, the islands that present the highest number of LR with the gathered data from inventory development, specifically Madeira island and in Azores on Terceira, S. Miguel, Graciosa and Pico islands, should be assessed.

- The implementation of the threat assessment methodology should be applied firstly to the aforementioned districts and islands, so LR threat risk level can be assessed and priority LR for conservation can be identified. The most threatened LR should be collected and duplicated in the national gene bank, if they are not already included. Local action should be undertaken in order to mitigate the threat reviewed and if possible implemented. Thus, the threat assessment development could help the promotion of more effective conservation strategies.

5.6 FUTURE WORK PROPOSALS

The following future work recommendations might be some interesting and urgent ideas to develop in a near future:

- To undertake collecting missions or LR surveys, by the several Portuguese Agricultural Departments in partnership with BPGV so the inventory could be verified and updated, especially after the fires that occurred in the last few years;
- To continue to study and update the inventory as a basis to assess the areas with higher number of LR for posterior analysis by researchers for implementation of on-farm conservation;
- To apply the threat assessment methodology to LR of studied areas so more organized and suited conservation actions could be established for LR;
- To apply the threat assessment methodology to determined and quantify landrace erosion;
- To apply the threat assessment methodology to help on selection of sites to implement on-farm conservation;

- To organize workshops in each municipality with multidisciplinary groups from diverse fields such as social sciences, biological sciences, marketing and education, in order to share and record knowledge of LR;
- To promote "Tasting Fairs", contests to appraise the better LR, gatherings and discussion groups, as actions to encourage LR maintenance;
- To undertake a gap analysis to assess gaps in LR on-farm and *ex situ* conservation;
- To promote of ethnographic studies where data concerning history, gastronomical uses, ecogeographic adaptations, and other reasons for certain LR development are registered;
- To make the studies referred in the previous point accessible so morphological and molecular classification can be complemented, and the knowledge obtained used by plant breeding sector.
- To promote the benefits of using LR, how to grow them, and donate seeds to local population that uses the areas provided by some municipalities for cultivation in cities.

5.7 CONCLUSION

The value of LR, in the livelihood of farmers/maintainers and food security, should be better appreciated by farmers, policy makers and the general public. In the light of climate as well as social and economic changes, plant breeding sector requires the genetic diversity and relevant traits they hold. Nevertheless, it is still the case that a lack of knowledge about the deep environmental and economic benefits of growing LR exists. Plant genetic resources as LR should be included in school curricula and the ecological and social interactions between LR, humans and natural ecosystems explained since early age. The value of traditional product designations (e.g. PDO) and the possibility of a sustainable and productive economy based on products where LR are main ingredients, should be presented and explained in channels used by farmers (e.g. local newspapers and radios, TV broadcasting) and not only on internet. The added value of using LR in organic agriculture, and the creation of a sustainable agricultural system where the cost of using bought seeds and pesticides are cut, should also be described in the formerly spoken channels. The agricultural cooperatives and municipalities should also be institutions where information about plant genetic resources could be learned and clarified.

Some municipalities have events where farmers can sell their LR, and organic markets (e.g. Feijão.come in São Pedro do Sul municipality of mainland Portugal). Nonetheless, usually these events have little national visibility and thorough marketing campaigns could be carried out to raise general awareness. Portuguese municipalities already disseminate information about the various events related to traditional products, recipes and LR. For example, *Festival das Sopas de Sernancelhe* in Viseu district, where a variety of traditional soups can be savoured, as *Festival da Castanha* (chestnut festival); *Festival da Batata-Doce de Aljezur* (Aljezur sweet potato festival) a Protected Geographical Indication (PGI) LR celebrated in Faro district. In Évora district, the Montemor-o-Novo municipality present the *Ao sabor das Estações* festival where the use of local products is encouraged thus, favouring a healthy diet and the reduction of carbon footprint, and *Festival das Sopas* with traditional soup recipes and use of traditional products. However, these events require a stronger marketing campaign to promote LR value nationally and not only regionally. Moreover, municipalities should seize the opportunity of these festivals to promote and improve LR knowledge and sustainable use, as undertake projects specifically linked to the subject of LR conservation and sustainable use.

With this work a Portuguese crop LR inventory was created, a LR threat assessment methodology was developed, and a molecular study of sampled common bean, a Portuguese staple crop, was performed which may help to put into practice the Portuguese National Strategic Action Plan for Plant Genetic Conservation (*Plano Nacional para os Recursos*

Genéticos Vegetais) (INIAV *et al.*, 2015) thus, enabling a platform where strategies and actions to promote enduring use and conservation of LR are outlined.

REFERENCES

Alercia, A.; Diulgheroff, S.; Mackay, M. (2012). *FAO/Bioversity Multi-Crop Passport Descriptors V.2 [MCPD V.2]*. Food and Agriculture Organization of the United Nations (FAO), Bioversity International, Rome (Italy).

Alexandratos, N. & Bruinsma, J. (2012). *World Agriculture towards 2030/2050: the 2012 revision*. ESA working paper Nº12-03. Rome, FAO.

Angioi, S.A.; Rau, D.; Attene, G.; Nanni, L.; Bellucci, E.; Logozzo, G.; Negri, V.; Spagnoletti Zeuli, P.L. & Papa, R. (2010). Beans in Europe: origin and structure of the European landraces of Phaseolus vulgaris L.. *Theoretical and Applied Genetics* 121:829–843.

Angioi, S.A.; Rau, D.; Rodriguez M.; Logozzo, G.; Desiderio, F.; Papa, R. & Attene G. (2008). Nuclear and chloroplast microsatellite diversity in Phaseolus vulgaris L. from Sardinia (Italy). *Molecular Breeding*. DOI 10.1007/s11032-008-9245-8.

Antofie, M.-M.; Sand, M.; Ciotea, G. & Iagrăru, P. (2010). Data sheet model for developing a Red List regarding crop landraces in Romania. *Annals: Food Science and Technology* 11(1): 45-49. Available at http://www.postdoctorat.ro/articole/4AR01.pdf.

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

Baganha, M.I.; Góis, P. & Pereira, P.T. (2005). International migration from and to Portugal: what do we know and where are we going? In: K. Zimmermann (Ed.), *European migration: what do we know?* Oxford University Press.

Barata, A.; Bettencourt, E.; Santos, A.; Eiras-Dias, J.; Oliveira, M.; Sousa, R. & Matos, C. (2008). The State of *Ex Situ* Management. In: INIA (ed.). *State of Plant Genetic Resources for Food and Agriculture in Portugal. Second Portuguese National Report on Conservation and Sustainable Utilization of Plant Genetic Resources for Food and Agriculture*. Ministério da

Agricultura do Desenvolvimento Rural e das Pescas, Instituto Nacional dos Recursos Biológicos, IP, Oeiras.

Blair, M.W.; Díaz, L.; Buendía, H. & Duque, M.C. (2009). Genetic diversity, seed size associations and population structure of a core collection of common beans (Phaseolus vulgaris L.). *Theoretical and Applied Genetics* 119: 955-972.

Blair, M.W.; Giraldo, M.C.; Buendía, H.; Tovar, E.; Duque, M.C. & Beebe, S.E. (2006). Microsatellite marker diversity in common bean (Phaseolus vulgaris L.). *Theoretical and Applied Genetics* 113: 100-109.

Blair, M.W.; Pedraza, F.; Buendia, H.F.; Gaitán-Solís, E.; Beebe, S.E.; Gepts, P. & Tohme, J. (2003). Development of a genome-wide anchored microsatellite map for common bean (Phaseolus vulgaris L.). *Theoretical and Applied Genetics* 107: 1362–1374. DOI 10.1007/s00122-003-1398-6.

Broughton, W.J.; Hernández, G.; Blair, M.; Beebe, S.; Gepts, P. & Vanderleyden, J. (2003). Beans (Phaseolus spp.) – model food legumes. *Plant and Soil* 252: 55-128.

Brush, S. (2000). The issues of *in situ* conservation of crop genetic resources. In: Brush SB (edt.). *Genes in the field: on-farm conservation of crop diversity*. International Development Research Centre and International Plant Genetic Resources Institute. Lewis Publishers.

Brush, S. (1992). Ethnoecology, biodiversity, and modernization in Andean potato Agriculture. *Journal of Ethnobiology* 12 (2): 161-185.

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

Câmara, C.R.S.; Urrea, C.A.& Schlegel, V. (2013). Pinto beans (Phaseolus vulgaris L.) as a functional food: implications on human health. *Agriculture* 3, 90-111.

Cardoso, S. & Maxted, N. (2009). Regional and Crop-Specific Survey: Grapevine Landraces in Douro and Colares, Portugal. In: Veteläinen, M.; Negri, V. and Maxted N. (eds.). *European landraces on – farm conservation, management and use*. Biodiversity Technical Bulletin No. 15. Biodiversity International, Rome, Italy.

Carvalho, N. S. (2008). Introduction to the Country and the Agricultural Sector. In: INIA (ed.). *State of Plant Genetic Resources for Food and Agriculture in Portugal. Second Portuguese National Report on Conservation and Sustainable Utilization of Plant Genetic Resources for Food and Agriculture*. Ministério da Agricultura do Desenvolvimento Rural e das Pescas, Instituto Nacional dos Recursos Biológicos, IP, Oeiras.

CBD (1992). Convention on Biological Diversity. Secretariat of the Convention on Biological Diversity. Montreal, Quebec, Canada.

Ceccarelli, S. (2012). Landraces: Importance and use in breeding and Environmentally Friendly Agronomic systems. In: Maxted, N.; Dulloo, M. E.; Ford-Lloyd, B.; Frese, L.; Iriondo, J. & Pinheiro de Carvalho, M.A.A. (eds.). *Agrobiodiversity Conservation: Securing the Diversity of Crop Wild Relatives and Landraces*. CABI International.

Chase, M.W. & Hills, H.H. (1991). Silica gel: an ideal material for field preservation of leaf samples for DNA studies. *Taxon* 40(2): 215- 220.

CIAT (2006). Molphas: Phaseolus molecular diversity database in: http://isa.ciat.cgiar.org/molphas/ [Accessed 15-06-2015].

CNV (2016). *Catálogo Nacional de Variedades 2016*. Ministério da Agricultura, Florestas e Desenvolvimento Rural; Direção-Geral de Alimentação e Veterinária; Direção de Serviços de Sanidade Vegetal; Divisão de Variedades e Sementes, Lisboa.

CNVF (2016). *Catálogo Nacional de Variedades – Fruteiras 2016*. Direção Geral de Alimentação e Veterinária; Direção de Serviços de Sanidade Vegetal; Divisão de Inspeção Fitossanitária e de Materiais de Propagação Vegetativa, Lisboa.

de Haan, S.; Bonierbale, M.; Ghislain, M.; Núñez, J. & Trujillo, G. (2007). Indigenous biosystematics of Andean potatoes: folk taxonomy, descriptors and nomenclature. *ISHS Acta Horticulturae* 745: 89-134.

Devos, K.M.; Bryan, G.J.; Collins, A.J.; Stephenson, P. & Gale, M.D. (1995). Application of two microsatellites sequences in wheat storage proteins as molecular markers. *Theoretical and Applied Genetics* 90: 247-252.

Díaz, L.; Buendía, H. & Duque, M.C. & Blair, M.W. (2011). Genetic diversity of Colombian landraces of common bean as detected through the use of silver-stained and fluorescently labelled microsatellites. *Plant Genetic Resources: Characterization and Utilization* 9 (1): 86-96.

European Commission (2017). *Agriculture and Rural Development, DOOR "Database Of Origin & Registration"* online database. http://ec.europa.eu/agriculture/quality/door/list.html. [Accessed 07-07-2017].

Falush, D.; Stephens, M. & Pritchard, J.K. (2003). Inference of population structure: Extensions to linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587.

Falush, D.; Stephens, M. & Pritchard, J.K. (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* 7: 574–578.

FAO (2005). Appendix 3, Classification of Crops. In: *FAO. A system of integrated agricultural censuses and surveys. World Programme for the Census of Agriculture 2010. Vol.1*. FAO, Rome, Italy.

FAO (2005a). *Building on Gender, Agrobiodiversity and Local Knowledge – a training manual*. FAO, Rome, Italy.

FAO (2009). *International Treaty on Plant Genetic Resources for Food and Agriculture*. FAO, Rome, Italy.

FAO, IFAD and WFP (2015). *The State of Food Insecurity in the World 2015. Meeting the 2015 international hunger targets: taking stock of uneven progress*. Rome, FAO.

Fonseca, J. (2004). *Colher para Semear - Manual prático para a colheita e conservação de sementes*. Colher para Semear.

Frankel, O. & Hawkes, J. (1975). Genetic Resources – the past ten years and the next. In: Frankel, O.H. and Hawkes, J.G. (eds.). *Crop Genetic Resources for today and tomorrow*. University Press, Cambridge.

Frankel, O.; Brown, A. & Burdon, J. (1995). The genetic diversity of cultivated plants. In: *The conservation of plant biodiversity*. University Press, Cambridge.

FSIN (2017). *Global Report on Food Crises 2017*. Online Available at: http://www.fao.org/emergencies/resources/documents/resources-detail/en/c/876564/.

Gioia, T.; Logozzo, G.; Attene, G.; Bellucci, E.; Benedettelli, S.; Negri, V.; Papa, R & Zeuli, P.S. (2013). Evidence for Introduction Bottleneck and Extensive Inter-Gene Pool (Mesoamerica x Andes) Hybridization in the European Common Bean (Phaseolus vulgaris L.) Germplasm. *PLoS ONE* 8(10): e75974. DOI: 10.1371/journal.pone.0075974.

Gonçalves-Vidigal, M.C. & Rubiano, L.B. (2011). Development and application of microsatellites in plant breeding. *Crop Breeding and Applied Biotechnology* S1: 66-72.

Gu, Z. (2014). circlize implements and enhances circular visualization in R. *Bioinformatics*.

Guarino, L. (1995). Assessing the threat of genetic erosion. In: Guarino, L.; Rao, V.R. and Reid, R. (eds.). *Collecting Plant Genetic Diversity – Technical Guidelines*. IPGRI, FAO, IUCN, UNEP. University Press, Cambridge.

Hammer, K. & Khoshbakht, K. (2005). Towards a “red list” for crop plant species. *Genetic Resources and Crop Evolution* 52: 249-265.

Hammer, K.; Knupffer, H.; Xhuveli, L. & Perrino, P. (1996). Estimating genetic erosion in landraces – Two case studies. *Genetic Resources and Crop Evolution* 43:329–336.

Hammer, K. (1991). Checklist and germplasm collecting. FAO/IBPGR. *Plant Genetic Resources Newsletter* 85: 15-17.

Hanelt, P. & IPK (2001). *Mansfeld's Encyclopedia of Agricultural and Horticultural Crops. 6 vols*. Springer, Berlin.

Hannukkala, A. (2009). The origin and special cultivation features of the northern local potato strain "Puikula". In: Veteläinen, M.; Negri, V. and Maxted N. (eds.). *European landraces on–farm conservation, management and use. Biodiversity Technical Bulletin No. 15*. Biodiversity International, Italy.

Harlan, J.R. (1992). *Crops and man. Second Edition*. American Society of Agronomy, Madison, USA.

Harlan, J. (1975). Our Vanishing Genetic Resources. *Science* 188 (4188): 617-621.

Hawkes, J.G.; Maxted, N. & Ford-Lloyd, B.V. (2000a). The Plant Genetic Resources Conservation Movement. In: *The ex situ conservation of plant genetic resources*. Kluwer Academic Publishers.

Hawkes, J.G.; Maxted, N. & Ford-Lloyd, B.V. (2000b). The Genetic Resources of Plants and their Value. In: *The ex situ conservation of plant genetic resources*. Kluwer Academic Publishers.

Hayward, A.C.; Tollenaere, R.; Dalton-Morgan, J. & Batley, J. (2015). Molecular marker applications in plants. *Methods Molecular Biology* 1245: 13-27.

Hubisz, M.; Falush, D.; Stephens, M. & Pritchard, J. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, In Press.

IBPGR Executive Secretariat (1982). *Descriptors for Phaseolus vulgaris*. IBPGR, Rome, Italy.

INE (2014). *Inquérito à estrutura das explorações agrícolas 2013. Edição 2014*. Instituto Nacional de Estatística, Lisboa.

INE (2017). *Estatísticas Agrícolas 2016. Edição 2017*. Instituto Nacional de Estatística, Lisboa.

INIAV, DGADR & DGAV (2015). *Plano Nacional para os Recursos Genéticos Vegetais*. INIAV, DGADR, DGAV. Ministério da Agricultura e do Mar, Lisboa.

IPCC (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (Core Writing Team, R. K. Pachauri and L.A. Meyer (eds.))*. IPCC, Geneva, Switzerland, 151pp.

IUCN (2001). *IUCN Red List Categories and Criteria: Version 3.1*. IUCN Species Survival Commission. IUCN, Gland, Switzerland and Cambridge, UK.

IVV (2016). *Instituto da Vinha e do Vinho IP*. http://www.ivv.gov.pt. [Accessed March 2016].

Jarvis, D., Hodgkin, T., Sthapit, B., Fadda, C. & Lopez-Noriega, I. (2011). An Heuristic Framework for Identifying Multiple Ways of Supporting the Conservation and Use of Traditional Crop Varieties within the Agricultural Production System. *Critical Reviews in Plant Sciences* 30:1-2, 125-176.

Jarvis, A.; Upadhyaya, H.; Gowda, C.L.L.; Aggarwal, P.K.; Fujisaka, S.; Anderson, B. (2010). *Climate Change and its effect on conservation and use of plant genetic resources for food and agriculture and associated biodiversity for food security. ICRISAT/FAO. Thematic background study for the Second Report on The State of the World's Plant Genetic*. Resources for Food and Agriculture. Rome.

Jombart, T. & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*. DOI: 10.1093/bioinformatics/btr521.

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24: 1403-1405. DOI: 10.1093/bioinformatics/btn129.

Joshi, B.; Upadhyay, M.P.; Gauchan, D.; Sthapit, B.R. & Joshi, K.D. (2004). Red Listing of Agricultural Crop Species, Varieties and Landraces. *Nepal Agricultural Research Journal* vol.5.

Kamvar, Z.N.; Brooks, J.C. & Grünwald, N.J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetic*. 6:208. DOI: 10.3389/fgene.2015.00208.

Kamvar, Z.N.; Tabima, J.F. & Grünwald, N.J. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2:e281. DOI: 10.7717/peerj.281.

Kell, S.; Maxted, N.; Allender, C.; Astley, D.; Ford-Lloyd, B.; and contributors (2009). *Vegetable landrace inventory of England and Wales*. The University of Birmingham, UK. 117 pp.

Leitão, S.T.; Dinis M.; Veloso M.M.; Šatovic, Z. & Vaz Patto, M.C. (2017). Establishing the Bases for Introducing the Unexplored Portuguese Common Bean Germplasm into the Breeding World. *Frontiers in Plant Science*. 8:1296. DOI: 10.3389/fpls.2017.01296.

Leitão, S.T.; Almeida, N.F.; Moral, A.; Rubiales, D. & Vaz Patto, M.C. (2013). Identification of resistance to rust (Uromyces appendiculatus) and powdery mildew (Erysiphe diffusa) in Portuguese common bean germplasm. *Plant Breeding* 132, 654–657.

Lemon, J. (2006). Plotrix: a package in the red light district of R. *R-News*, 6(4): 8-12.

Ligarreto, G. & Ocampo, C. (2012). Genetic diversity in a Colombian bean (Phaseolus vulgaris L.) collection as assessed by phaseolin patterns and isoenzymatic markers. *Agronomía Colombiana* 30(2): 179-187.

Litrico, I. & Violle, C. (2015). Diversity in Plant Breeding: A new conceptual framework. *Trends in Plant Science* 20(10): 604-613.

Lopes, V.; Barata, A.; Reis, A.; Rocha, F.; Bettencourt, E. & Nunes, E. (2007). Avaliação preliminar da biodiversidade de populações de feijão de áreas protegidas do Minho e Galiza com recurso a marcadores moleculares. Conference: Workshop "Biodiversidade em sistemas agrícolas e florestais". Estação Agronómica Nacional Oeiras, 25 e 26 October 2007. Available at: https://www.researchgate.net/publication/268816146 [Accessed 03-06-2015].

Lorenzetti, F. & Negri, V. (2009). The European Seed Legislation on Conservation Varieties. In: Veteläinen, M.; Negri, V. and Maxted N. (eds.). *European landraces on – farm conservation, management and use. Biodiversity Technical Bulletin No. 15*. Biodiversity International, Rome, Italy.

Maechler, M.; Rousseeuw, P.; Struyf, A.; Hubert, M.& Hornik, K. (2016). cluster: Cluster Analysis Basics and Extensions. R package version 2.0.4.

Martins, S.R.; Vences, F.J.; Sáenz de Miera, L.E.; Barroso, M.R. & Carnide, V. (2006). RAPD analysis of genetic diversity among and within Portuguese landraces of common white bean (Phaseolus vulgaris L.). *Scientia Horticulturae* 108: 133-142.

Maxted, N.; Magos Brehm, J. & Kell, S. (2013). *Resource book for preparation of national conservation plans for crop wild relatives and landraces*. University of Birmingham, UK.

Maxted, N.; Kell, S. & Magos Brehm, J. (2011). *Options to promote food security: on-farm management and in situ conservation of plant genetic resources for food and agriculture. Background study paper nº51*. Commission on genetic resources for food and agriculture. FAO.

Maxted, N.; Veteläinen, M. & Negri, V. (2009). Landraces Inventories: Needs and Methodologies. In: Veteläinen, M.; Negri, V. and Maxted N. (eds.). *European landraces on – farm conservation, management and use. Biodiversity Technical Bulletin No. 15*. Biodiversity International, Rome, Italy.

Maxted, N. & Guarino, L. (2006). Genetic erosion and genetic pollution of crop wild relatives. In: Ford-Lloyd, B.V.; Dias, S.R. and Bettencourt E. (eds). *Genetic Erosion and Pollution Assessment Methodologies. Proceedings of PGR Forum Workshop 5, Terceira Island,* *Autonomous Region of the Azores, Portugal, 8–11 September 2004*. Published on behalf of the European Crop Wild Relative Diversity Assessment and Conservation Forum, by Biodiversity International, Rome, Italy. 100 p.

Maxted, N. (2000). Genetic reserve conservation of PGRFA in Europe. In: Laliberté, B.; Maggioni, L.; Maxted, N. & Negri, V. compilers. *ECP/GR In situ and On-farm Conservation Network. Report of a joint meeting of a Task Force on Wild Species Conservation in Genetic Reserves and a Task Force on On-farm Conservation and Management, 18-20 May 2000, Isola Polvese, Italy*. International Plant Genetic Resources Institute, Rome, Italy.

Maxted, N.; Ford-Lloyd, B.V. & Hawkes, J.G. (1997). Complementary conservation strategies. In: Maxted, N.; Ford-Lloyd, B.V. & Hawkes, J.G. (eds.). *Plant Genetic Conservation – the in situ approach*. Chapman & Hall, London, U.K.

Maxted, N.; Hawkes, J.G.; Ford-Lloyd, B.V. & Williams, J. T. (1997a). A practical model for *in situ* genetic conservation. In: Maxted, N.; Ford-Lloyd, B.V. and Hawkes, J.G. (eds.). *Plant genetic conservation: the in situ approach*. Chapman & Hall, London, UK. Pp. 339-367.

Maxted, N.; Hawkes, J.; Guarino, L. & Sawkins, M. (1997b). Towards the selection of taxa for plant genetic conservation. *Genetic Resources and Crop Evolution* 44: 337-348.

Meirmans, P.G. (2006). Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution* 60(11): 2399-2402.

Mendes Moreira, P.; Pêgo, S. & Vaz Patto, M.C. (2009). On-farm Conservation Portugal – Vaso Project – a Long-term Conservation Programme. In: Veteläinen, M.; Negri, V. and Maxted N. (eds.). *European landraces on – farm conservation, management and use. Biodiversity Technical Bulletin No. 15*. Biodiversity International, Rome, Italy.

Mendes-Moreira, P. & Vaz-Patto, M.C. (2012). Maize traditional landraces in Portugal: Past, Present or Future? In: Negri, V.; Torricelli, R. & Maxted, N. (eds.). *Landraces issue 1 – Conserving Europe's Plant Genetic Resources for use now and in the future*. University of Perugia, Italy.

Moore, S.S.; Sargenat, L.L.; King, T.J.; Mattick, J.S.; Georges, M. & Metzel, D.J.S. (1991). The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. *Genomics* 10: 654-660.

Morgante, M. & Olivieri, A.M. (1993). PCR-amplified microsatellites as markers in plant genetics. *The Plant Journal* 3: 175-182.

Negri, V. (2009). 'Fagiolina' (Vigna unguiculata subsp. unguiculata (L.) Walp.) from Trasimeno Lake (Umbria Region, Italy). In: Veteläinen, M.; Negri, V. and Maxted N. (eds.). *European landraces on – farm conservation, management and use. Biodiversity Technical Bulletin No. 15*. Biodiversity International, Rome, Italy.

Negri, V.; Maxted, N. & Veteläinen, M. (2009). European Landrace Conservation: an Introduction. In: Veteläinen, M.; Negri, V. and Maxted N. (eds.). *European landraces on – farm conservation, management and use. Biodiversity Technical Bulletin No. 15*. Biodiversity International, Rome, Italy.

Negri, V. (2005). Agro-biodiversity conservation in Europe: Ethical Issues. *Journal of Agricultural and Environmental Ethics* 18: 3-25.

Negri, V. (2003). Landraces in central Italy: where and why they are conserved and perspectives for their on-farm conservation. *Genetic Resources and Crop Evolution* 50: 871-885.

Newton, A.C.; Akar, T.; Baresel, J.P.; Bebeli, P.J.; Bettencourt, E.; Bladenopoulos, K.V.; Czembor, J.H.; Fasoula, D.A.; Katsiotis, A.; Koutis, K.; Koutsika-Sotiriou, M.; Kovacs, G.; Larsson, H.; Pinheiro de Carvalho, M.A.A.; Rubiales, D.; Russell, J.; Dos Santos, T.M.M. & Vaz Patto, M.C. (2010). Cereal landraces for sustainable agriculture. A review. *Agronomy for Sustainable Development* 30: 237-269.

Okii, D.; Tukamuhabwa, P.; Kami, J.; Namayanja, A.; Paparu, P.; Ugen, M. & Gepts, P. (2014). The genetic diversity and population structure of common bean (Phaseolus vulgaris L.) germplasm in Uganda. *African Journal of Biotechnology* 13(29): 2935-2949 DOI:10.5897/AJB2014.13916.

Padulosi, S. & Dulloo, E. (2012). Towards a viable system for monitoring agrobiodiversity on-farm: A proposed new approach for Red Listing of cultivated plant species. In: Padulosi, S.; Bergamini, N. and Lawrence, T. (eds.). *On farm conservation of neglected and underutilized species: status, trends and novel approaches to cope with climate change. Proceedings of an International Conference, Frankfurt, 14-16 June 2011*. Biodiversity International, Rome. p 171-187.

Paradis, E.; Claude, J. & Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289-290.

Parra-Quijano, M.; Lamas, E. T.; Alegría J. M. I.; López, F. & Pertínez, A. (2016). *CAPFITOGEN Tools. User Manual version 2.0*. International Treaty on Plant Genetic Resources for Food and Agriculture. FAO. Rome. Available online at: http://www.capfitogen.net/en/. [Accessed 06-12-2017].

Parra-Quijano, M.; Lamas, E.T.; Alegría, J.M.I. & López, F. (2014). *Capfitogen. Programme to Strengthen National Plant Genetic Resources Capacities in Latin America version 1.2*. FAO.

Parra-Quijano, M.; Iriondo, J. M. & Torres, E. (2012). Improving representativeness of genebank collections through species distribution models, gap analysis and ecogeographic maps. *Biodiversity and Conservation* 21: 79-96.

Peakall, R. & Smouse P.E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28 (19): 2537–2539. doi:10.1093/bioinformatics/bts460.

Peakall, R. & Smouse P.E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6, 288-295.

Pritchard, J.K.; Stephens, M. & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

Porfiri, O.; Costanza, M. & Negri, V. (2009). Landrace Inventories in Italy and the Lazio Region Case Study. In: Veteläinen, M.; Negri, V. and Maxted, N. (eds.). *European landraces on-farm conservation, management and use. Biodiversity Technical Bulletin No. 15*. Biodiversity International, Rome, Italy.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Saghai-Maroof, M.A.; Biyashev, R.M.; Yang, G.P.; Zhang, Q. & Allard, R.W. (1994). Extraordinarily polymorphic microsatellites DNA in barley: Species diversity, chromosomal locations and population dynamics. *Proceedings of the National Academy of Sciences U.S.A*. 91: 5466-5470.

Santalla, M.; Rodiño, A.P.& De Ron, A.M. (2002). Allozyme evidence supporting southwestern Europe as a secondary centre of genetic diversity for the common bean. *Theoretical and Applied Genetics* 104: 934-944.

Senior, M.L. & Heun, M. (1993). Mapping maize microsatellites and polymerase chain reaction confirmation of targeted repeats using a CT primer. *Genome* 36: 884-889.

Tal Galili (2015). dendextend: an R package for visualizing, adjusting, and comparing trees of hierarchical clustering. *Bioinformatics*. DOI: 10.1093/bioinformatics/btv428.

Torricelli, R. & Negri, V. (2015). Successful examples of landrace on-farm conservation and use in Italy. In: Negri, V.; Torricelli, R. & Maxted, N. (eds.). *Landraces issue 3 – Conserving Europe's Plant Genetic Resources for use now and in the future*. University of Perugia, Italy.

United Nations (2015). *World Population Prospects: The 2015 Revision, Key Findings and Advance Tables. Working Paper No. ESA/P/WP.241*. Department of Economic and Social Affairs, Population Division.

USDA, ARS, GRIN (2017). *Germplasm Resources Information Network* [Internet]. Beltsville (MD): United States Department of Agriculture, Agricultural Research Service. Available at: https://www.ars-grin.gov/. [Accessed 16-03-2017].

Veloso, M. (2008). The State of Diversity. In: INIA (ed.) *State of Plant Genetic Resources for Food and Agriculture in Portugal. Second Portuguese National Report on Conservation and Sustainable utilization of Plant Genetic Resources for Food and Agriculture*. Ministério da Agricultura do Desenvolvimento Rural e Pescas, Instituto Nacional dos Recursos Biológicos IP, Oeiras.

Veloso, M.; Duarte, M. & Moreira, P. (2008). The State of *In Situ* Management. In: INIA (ed.) *State of Plant Genetic Resources for Food and Agriculture in Portugal. Second Portuguese National Report on Conservation and Sustainable utilization of Plant Genetic Resources for Food and Agriculture*. Ministério da Agricultura do Desenvolvimento Rural e Pescas, Instituto Nacional dos Recursos Biológicos IP, Oeiras.

Veloso, M.M.; Passarinho, J.A.; Matos, C.R.; Carneiro, L.C.; Leitão, S.T. & Vaz-Patto, M.C. (2015a). Common bean (Phaseolus vulgaris L.) landraces in Portugal: morphological diversity of the seed. In: Negri, V.; Torricelli, R. & Maxted, N. (eds.). *Landraces issue 3 – Conserving Europe's Plant Genetic Resources for use now and in the future*. University of Perugia, Italy.

Veloso, M.M.; Reis, P. & Machado, D. (2015b). On-farm conservation of the olive tree (Olea europaea) landraces at Vila Verde de Ficalho (Portugal). In: Negri, V.; Torricelli, R. & Maxted, N. (eds.). *Landraces issue 3 – Conserving Europe's Plant Genetic Resources for use now and in the future*. University of Perugia, Italy.

Veteläinen, M.; Negri, V. & Maxted, N. (2009). A European Strategic Approach to Conserving Crop Landraces. In: Veteläinen, M.; Negri, V. and Maxted, N. (eds.) *European landraces on-farm conservation, management and use. Biodiversity Technical Bulletin No. 15*. Biodiversity International, Rome, Italy.

Voegel, R. (2012). Red List for crops – a tool for monitoring genetic erosion, supporting re-introduction into cultivation and guiding conservation efforts. In: Padulosi, S.; Bergamini, N.

and Lawrence, T. (eds). *On farm conservation of neglected and underutilized species: status, trends and novel approaches to cope with climate change: Proceedings of an International Conference, Frankfurt, 14-16 June, 2011*. Biodiversity International, Rome.

Vosman, B. & Arens, P. (1997). Molecular characterization of GATA/GACA microsatellite repeats in tomato. *Genome* 40: 25-33.

Walter, K. & Gillett, H. (1998) (eds.). *1997 IUCN Red List of Threatned Plants compile by the World Conservation Monitoring Centre*. IUCN - The World Conservation Union, Gland, Switzerland and Cambridge, U.K. Pp. 862.

Wright, S. (1978). *Evolution and the Genetics of Populations. Variability within and among natural populations. Vol 4*. The University of Chicago Press, Chicago.

Yu, K.; Park, S. J.; Poysa, V. & Gepts, P. (2000). Integration of Simple Sequence Repeat (SSR) markers into a molecular linkage map of common bean (Phaseolus vulgaris L.). *The American Genetic Association* 91: 429-434.

Yu, K.; Park, S.J. & Poysa, V. (1999). Abundance and variation of microsatellite DNA sequences in beans (Phaseolus and Vigna). *Genome* 42: 27-34.

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

SUPPLEMENTARY MATERIAL

Supplementary Figure 2.1.: Questionnaire applied to farmers during field work.

Questionnaire to farmers

1. Farmers' information

Date: ____________ Altitude: ____________________
Latitude: ___________ Longitude:___________________
Solar exposition: N / S / E / W / NE / SE / NW / SW

ID: ________ Age: _______ Contact later: y/ n
Name: ____
____
Address: ______________________________________
District: ______________ Municipality: ______________
Parish: ______________ Locality: ___________________
Teleph.: _____________ Mobile: ___________________
Fax: __________ Email: __________________________
Member of association? _________________________
Are you interested in maintaining LR long term/ years that been maintaining LR? __________________________

2. Social-economics information

Do yo work in other activity besides agriculture?
____
Farms´area: <0,5ha 0,5–2ha 2,5–5ha >5ha
Nr. people working on the farm: _____ Relation: ________
Are family members willing to maintain LR?
____

3. Crop data

Genus	Species	Local name

Collecting point: field / warehouse / other
____

Purpose: self-consumption / locally sold / sold outside community / partially sold/ other ____________________

Used part of plant: stem / leaves / seed / fruit

Plant uses: human consumption/ animal feeding/ other
____

No. of people in the locality that still maintain LR?____

Picture(s): y/ n ______________________________

Value of LR in the farm: main crop / harvest before main crop/ harvested after main crop / backyard small productions/ other_______________________________

Sowing period	Harvest period

4. Cultural practices

Irrigation / rotation / organic fertilizers / inorganic fertilizers / animal traction / mechanization

5. Qualities of LR

market request / tradition / flavor/ diseases and plagues resistance / high yield / good storage / other
____

6. Local physiography

Plain / valley / mountain / hill / other ____________________

7. Local topography

Plan / slop / mountainous / other
____

8. Soil texture: sandy/ clay/ balanced

9. Seed descriptors

Seed coat patterns:

1 / 2 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10 / 99 (see descript. for *P. vulgaris* 4.3.1)

Seed coat darker colour:

1 / 2 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10 / 11 / 12 / 13 / 14 / 99 (see descript. for *P. vulgaris* 4.3.2)

Seed coat lighter colour:

1 / 2 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10 / 11 / 12 / 13 / 14 / 99 (see descript. for *P. vulgaris* 4.3.3)

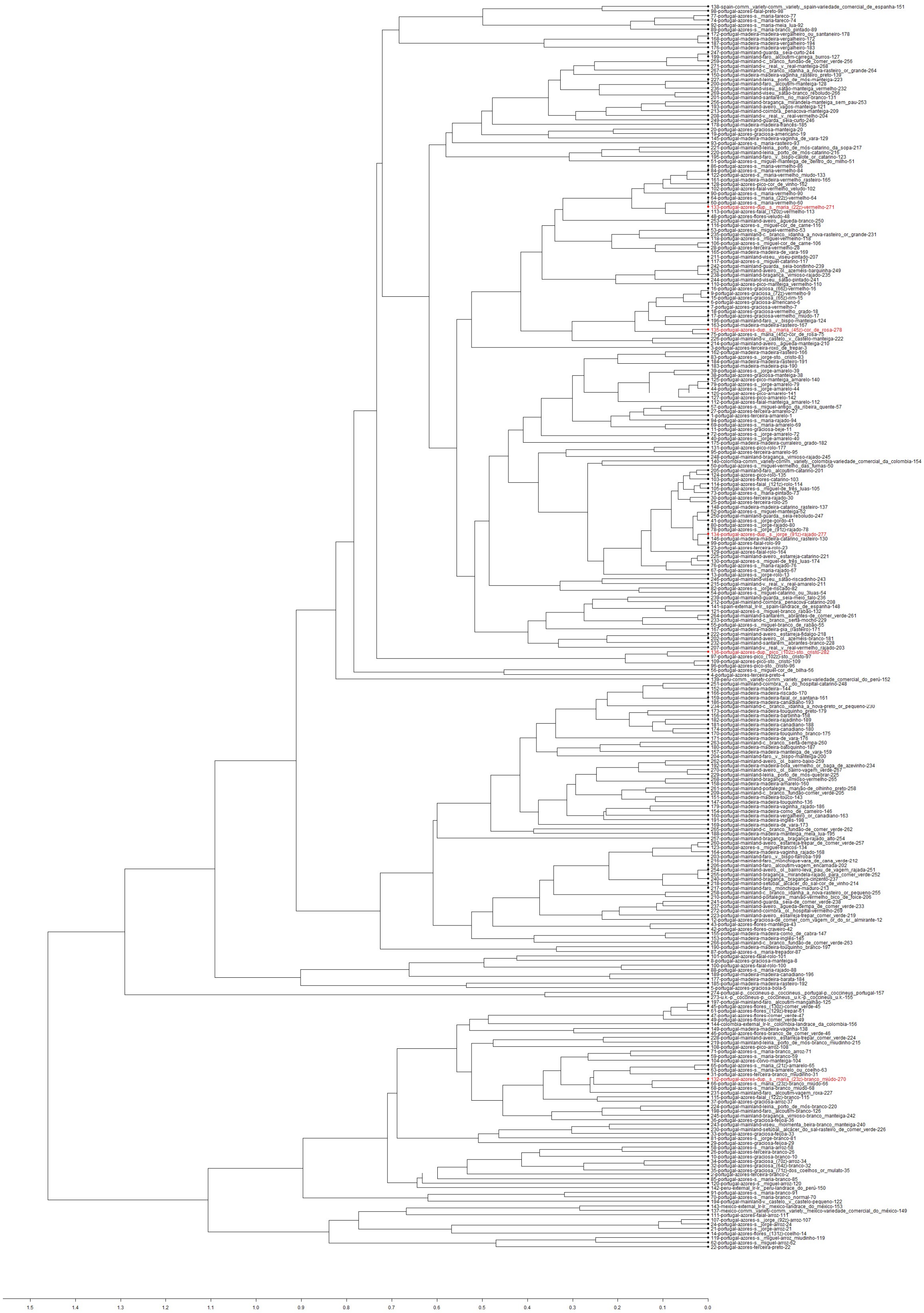
Brilliance of seed: Matt/ medium/ shiny

Seed shape:1 / 2 / 3 / 4 / 5 (see descript. for *P. vulgaris* 4.3.5)

Observation:_____________________________________
____
____
____
____
____
____
____
____

Supplementary Figure 4.10a.: UPGMA dendrogram based on Nei's genetic distance for the 274 studied samples (272 *P. vulgaris* L. and 2 *P. coccineus* L.) with external LR in red and commercial varieties in blue (see CD at the end of this thesis).

Supplementary Figure 4.10b.: UPGMA dendrogram based on Nei's genetic distance for the 274 studied samples (272 *P. vulgaris* L. and 2 *P. coccineus* L.) with duplicates samples in red (see CD at the end of this thesis).



Supplementary Table 2.6.: Number of landraces per taxa per district and per island in the Azores and Madeira archipelagos.

District	Taxa* (number of landraces)	Genera (number of landraces)
Aveiro (529)	*Vitis vinifera* (94), *Phaseolus vulgaris (88), Malus domestica (65), Zea mays subsp. mays (46), Pyrus communis (34), Brassica rapa (16), Allium cepa (11), Ficus carica (11), Prunus avium (11), Brassica oleracea var. acephala (9), Cucurbita maxima (8), Cucurbita moschata (8), Cucurbita pepo (8), Lactuca sativa (8), Prunus domestica (8), Brassica oleracea var. costata (7), Allium sativum (5), Pisum sativum (5), Prunus persica (5), Solanum lycopersicum var. lycopersicum (10), Avena sativa (4), Capsicum annuum (4), Lagenaria siceraria (4), Lolium multiflorum (4), Secale cereale (4), Cucumis sativus (3), Phaseolus coccineus (3), Brassica napus (2), Brassica rapa subsp. rapa (2), Citrus sinensis (2), Corylus avellana (2), Cucurbita ficifolia (2), Cydonia oblonga (2), Linum usitatissimum (2), Lupinus albus (2), Olea europeae (2), Triticum aestivum subsp. aestivum (2), Vicia faba (2), Vigna unguiculata subsp. sesquipedalis (2), Allium ampeloprasum (1), Arachis hypogaea (1), Arbutus unedo (1), Avena strigosa (1), Capsicum frutescens (1), Castanea sativa (1), Citrullus lanatus (1), Citrus limon (1), Citrus reticulata (1), Cucumis melo (1), Diospyrus kaki (1), Eriobotrya japonica (1), Glycine max (1), Hordeum vulgare subsp. vulgare (1), Lupinus luteus (1), Melissa officinalis (1), Oryza sativa (1), Petroselinum crispum (1), Punica granatum (1), Raphanus sativus (1), Setaria italica (1), Vigna unguiculata (1)*	*Cucurbita sp. (5), Ocimum sp. (2), Raphanus sp. (1)*
Corvo Island (Azores Autonomous Region) (10)	*Citrus sinensis (3), Allium cepa (1), Colocasia esculenta (1), Ipomoea batatas (1), Malus domestica (1), Phaseolus vulgaris (1), Pyrus communis (1), Zea mays subsp. mays (1)*	*Brassica sp. (2), Cucurbita sp. (2)*
Faial Island (Azores Autonomous Region) (67)	*Phaseolus vulgaris (10), Ipomoea batatas (8), Malus domestica (5), Solanum lycopersicum var. lycopersicum (5), Vicia faba (5), Zea mays susp. mays (5), Allium sativum (4), Allium cepa (3), Colocasia esculenta (3), Pyrus communis (3), Allium schoenoprasum (2), Cucumis sativus (2), Capsicum annuum (1), Citrullus lanatus (1), Citrus limon (1), Citrus sinensis (1), Cucumis melo (1), Ficus carica (1), Hordeum vulgare subsp. vulgare (1), Lactuca sativa (1), Lepidium sativum (1), Physalis peruviana (1), Pisum sativum (1), Secale cereale (1)*	*Cucurbita sp. (7), Brassica sp. (2), Capsicum sp. (2), Avena sp. (1), Musa sp. (1)*

District	Taxa* (number of landraces)	Genera (number of landraces)
Flores Island (Azores Autonomous Region) (64)	*Phaseolus vulgaris (15), Zea mays subsp. mays (12), Citrus sinensis (7), Ipomoea batatas (5), Brassica rapa (2), Colocasia esculenta (2), Pisum sativum (2), Pyrus communis (2), Solanum tuberosum (2), Vicia faba (2), Allium cepa (1), Citrus limon (1), Cucumis sativus (1), Eriobotrya japonica (1), Hordeum vulgare subsp. vulgare (1), Lupinus albus (1), Lupinus luteus (1), Malus domestica (1), Ornithopus sativus (1), Prunus domestica (1), Prunus persica (1), Solanum lycopersicum var. lycopersicum (1), Spinacia oleracea (1)*	*Brassica sp. (8), Cucurbita sp. (5), Avena sp. (1), Musa sp. (1)*
Graciosa Island (Azores Autonomous Region) (81)	*Phaseolus vulgaris (22), Citrus sinensis (7), Ficus carica (5), Allium cepa (3), Allium sativum (3), Pisum sativum (3), Psidium cattleyanum (3), Pyrus communis (3), Vicia faba (3), Zea mays subsp. mays (3), Capsicum annuum (2), Ipomoea batatas (2), Malus domestica (2), Prunus domestica (2), Solanum lycopersicum var. lycopersicum (2), Solanum tuberosum (2), Allium ampeloprasum (1), Brassica rapa (1), Cicer arietinum (1), Citrus limon (1), Colocasia esculenta (1), Cucumis sativus (1), Eriobotrya japonica (1), Lactuca sativa (1), Lathyrus sativus (1), Lathyrus tingitanus (1), Lupinus albus (1), Passiflora edulis (1), Solanum betaceum (1), Sorghum bicolor (1)*	*Brassica sp. (9), Cucurbita sp. (3), Musa sp. (2)*
Pico Island (Azores Autonomous Region) (72)	*Phaseolus vulgaris (19), Citrus sinensis (6), Solanum lycopersicum var. lycopersicum (6), Allium cepa (5), Citrus limon (4), Ficus carica (4), Vicia faba (4), Zea mays subsp. mays (4), Allium sativum (3), Colocasia esculenta (2), Cucumis sativus (2), Ipomoea batatas (2), Arachis hypogaea (1), Capsicum annuum (1), Citrus reticulata (1), Daucus carota (1), Lactuca sativa (1), Lupinus albus (1), Malus domestica (1), Pisum sativum (1), Prunus armeniaca (1), Prunus domestica (1), Pyrus communis (1)*	*Cucurbita sp. (7), Brassica sp. (7)*
S. Jorge Island (Azores Autonomous Region) (58)	*Phaseolus vulgaris (14), Ipomoea batatas (6), Zea mays subsp. mays (5), Allium cepa (4), Solanum lycopersicum var. lycopersicum (4), Allium sativum (3), Citrus sinensis (3), Ficus carica (3), Capsicum annuum (2), Cucumis sativus (2), Pisum sativum (2), Vicia faba (2), Brassica rapa (1), Colocasia esculenta (1), Eriobotrya japonica (1), Lepidium sativum (1), Malus domestica (1), Prunus domestica (1), Psidium cattleyanum (1), Solanum tuberosum (1)*	*Brassica sp. (4), Cucurbita sp. (4), Coffea (1), Musa sp. (1)*

District	Taxa* (number of landraces)	Genera (number of landraces)
S. Miguel Island (Azores Autonomous Region) (75)	*Phaseolus vulgaris (25), Allium sativum (5), Zea mays subsp. mays (4), Solanum lycopersicum var. lycopersicum (4), Vicia faba (4), Ficus carica (4), Allium cepa (4), Colocasia esculenta (3), Malus domestica (3), Ipomoea batatas (3), Capsicum annuum (2), Cucumis sativus (2), Lactuca sativa (2), Spinacia oleracea (1), Pyrus communis (1), Lepidium sativum (1), Sorghum bicolor (1), Citrus sinensis (1), Camellia sinensis (1), Eriobotrya japonica (1), Ananas comosu var. cayene (1), Citrullus lanatus (1), Physalis peruviana (1)*	*Cucurbita sp. (9), Brassica sp. (6), Capsicum sp. (3), Musa sp. (2)*
Sta. Maria Island (Azores Autonomous Region) (83)	*Phaseolus vulgaris (24), Ipomoea batatas (7), Pisum sativum (6), Allium cepa (5), Citrus sinensis (5), Vicia faba (5), Zea mays subsp. mays (4), Brassica rapa (3), Colocasia esculenta (3), Lathyrus sativus (3), Prunus domestica (3), Solanum lycopersicum var. lycopersicum (3), Allium sativum (2), Lepidium sativum (2), Malus domestica (2), Capsicum annuum (1), Citrus reticulata (1), Cucumis sativus (1), Psidium cattleyanum (1), Solanum betaceum (1), Vitis vinifera (1)*	*Brassica sp. (4), Cucurbita sp. (4), Capsicum sp. (1), Musa sp. (1)*
Terceira Island (Azores Autonomous Region) (87)	*Phaseolus vulgaris (35), Colocasia esculenta (7), Zea mays subsp. mays (6), Ipomoea batatas (5), Solanum lycopersicum var. lycopersicum (5), Ficus carica (4), Allium sativum (3), Vicia faba (3), Allium cepa (2), Brassica rapa (2), Arachis hypogaea (1), Beta vulgaris subsp. vulgaris (2), Castanea sativa (1), Cicer arietinum (1), Citrullus lanatus (1), Citrus sinensis (1), Cucumis melo (1), Cucumis sativus (1), Lolium multiflorum (1), Malus domestica (1), Nicotiana tabacum (1), Pisum sativum (1), Sorghum bicolor (1), Spinacia oleracea (1)*	*Brassica sp. (7), Cucurbita sp. (6), Musa sp. (1)*
Beja (168)	*Vitis vinifera (35), Zea mays subsp. mays (15), Solanum lycopersicum var. lycopersicum (14), Phaseolus vulgaris (10), Citrullus lanatus (9), Capsicum annuum (8), Cucumis melo (8), Allium sativum (7), Olea europeae (7), Lactuca sativa (6), Triticum aestivum (5), Cicer arietinum (4), Cucumis sativus (4), Lathyrus sativus (4), Coriandrum sativum (3), Hordeum vulgare subsp. vulgare (3), Vigna unguiculata (3), Allium cepa (2), Avena sativa (2), Brassica oleracea var. acephala (2), Cucurbita pepo (2), Petroselinum crispum (2), Vicia faba (2), Allium ascalonicum (1), Cucurbita maxima (1), Daucus carota (1), Lagenaria siceraria (1), Lens culinaris (1), Lepidium sativum (1), Lupinus albus (1), Pisum sativum (1), Raphanus sativus (1), Spinacia oleracea (1), Triticum turgidum subsp. durum (1)*	*Cucurbita sp. (1), Medicago sp. (1), Mentha sp. (1)*

District	Taxa* (number of landraces)	Genera (number of landraces)
Braga (393)	*Phaseolus vulgaris (74), Zea mays subsp. mays (72), Vitis vinifera (45), Malus domestica (19), Allium sativum (17), Vigna unguiculata (16), Brassica rapa subsp. rapa (13), Brassica rapa (12), Pisum sativum (12), Cucurbita pepo (10), Secale cereale (10), Allium cepa (8), Cucumis melo (6), Lolium multiflorum (6), Phaseolus coccineus (6), Cucurbita maxima (5), Brassica oleracea var. acephala (4), Brassica oleracea var. costata (4), Capsicum annuum (4), Lactuca sativa (4), Avena sativa (3), Brassica napus (3), Citrullus lanatus (3), Lagenaria siceraria (3), Linum usitatissimum (3), Pyrus communis (3), Solanum lycopersicum var. lycopersicum (6), Coriandrum sativum (2), Corylus avellana (2), Cucumis sativus (2), Triticum aestivum subsp. aestivum (2), Vicia faba (2), Allium ampeloprasum (1), Avena strigosa (1), Capsicum frutescens (1), Cucurbita ficifolia (1), Helianthus annuus (1), Holcus lanatus (1), Lolium perenne (1), Lupinus albus (1), Ornithopus compressus (1), Ornithopus sativus (1), Panicum miliaceum (1), Petroselinum crispum (1)*	*Avena sp. (1), Lolium sp. (1), Ocimum sp. (1)*
Bragança (849)	*Phaseolus vulgaris (152), Vitis vinifera (151), Capsicum annuum (39), Pyrus communis (38), Solanum lycopersicum var. lycopersicum (35), Cucumis melo (27), Zea mays subsp. mays (27), Lactuca sativa (25), Brassica rapa subsp. rapa (22), Brassica rapa (21), Solanum lycopersicum var. lycopersicum (20), Vigna unguiculata (19), Olea europeae (18), Cicer arietinum (14), Citrullus lanatus (14), Malus domestica (14), Triticum aestivum subsp. aestivum (14), Brassica oleracea var. costata (11), Prunus avium (11), Cucumis sativus (10), Cucurbita pepo (10), Pisum sativum (10), Allium cepa (9), Avena sativa (9), Brassica oleracea var. acephala (8), Castanea sativa (8), Ficus carica (8), Allium sativum (7), Cucurbita maxima (7), Lupinus albus (7), Juglans regia (6), Secale cereale (6), Vicia faba (6), Lathyrus sativus (5), Prunus dulcis (5), Sorghum bicolor (5), Capsicum frutescens (4), Cucurbita moschata (4), Raphanus sativus (4), Vigna unguiculata subsp. sesquipedalis (4), Brassica napus (3), Linum usitatissimum (3), Phaseolus coccineus (3), Beta vulgaris subsp. vulgaris (2), Hordeum vulgare subsp. vulgare (2), Lens culinaris (2), Punica granatum (2), Allium ampeloprasum (1), Avena strigosa (1), Brassica oleracea var. capitata (1), Brassica oleracea var. sabauda (1), Capsicum chinense (1), Coriandrum sativum (1), Cucurbita ficifolia (1), Cydonia oblonga (1), Helianthus annuus (1), Lagenaria siceraria (1), Morus nigra (1), Ocimum basilicum (1), Petroselinum crispum (1), Prunus domestica (1), Setaria italica (1), Solanum*	*Vicia sp. (3), Avena sp. (1), Brassica sp. (1), Lupinus sp. (1)*

District	Taxa* (number of landraces)	Genera (number of landraces)
	melongena (1), Triticum turgidum subsp. durum (1), Vicia articulata (1)	
Castelo Branco (134)	*Malus domestica (24), Phaseolus vulgaris (24), Zea mays subsp. mays (19), Olea europeae (9), Vigna unguiculata (9), Allium cepa (3), Brassica oleracea var. acephala (3), Linum usitatissimum (3), Solanum lycopersicum var. lycopersicum (5), Allium sativum (2), Brassica oleracea var. costata (2), Brassica rapa subsp. rapa (2), Cicer arietinum (2), Cucumis melo (2), Cucurbita pepo (2), Diospyrus kaki (2), Brassica napus (1), Castanea sativa (1), Citrullus lanatus (1), Citrus sinensis (1), Coriandrum sativum (1), Cucurbita moschata (1), Cydonia oblonga (1), Ficus carica (1), Juglans regia (1), Lactuca sativa (1), Lupinus luteus (1), Origanum vulgare subsp. virens (1), Phaseolus coccineus (1), Pisum sativum (1), Prunus avium (1), Pyrus communis (1), Raphanus sativus (1), Secale cereale (1), Triticum aestivum subsp. aestivum (1), Vicia faba (1), Vicia sativa (1)*	*Brassica sp. (2), Lolium sp. (1), Mentha sp. (1)*
Coimbra (302)	*Phaseolus vulgaris (83), Vitis vinifera (62), Zea mays subsp. mays (37), Brassica rapa (17), Brassica oleracea var. acephala (9), Pisum sativum (8), Pyrus communis (8), Secale cereale (7), Brassica oleracea var. costata (6), Lactuca sativa (6), Triticum aestivum subsp. aestivum (6), Brassica rapa subsp. rapa (5), Vicia faba (5), Solanum lycopersicum var. lycopersicum (6), Vigna unguiculata (4), Allium sativum (3), Avena sativa (3), Cucurbita pepo (3), Triticum turgidum subsp. durum (3), Lupinus albus (2), Malus domestica (2), Allium cepa (1), Beta vulgaris subsp. vulgaris (1), Brassica napus (1), Capsicum annuum (1), Cicer arietinum (1), Citrullus lanatus (1), Coriandrum sativum*	*Brassica sp. (1), Cucurbita sp. (1), Lupinus sp. (1)*

District	Taxa* (number of landraces)	Genera (number of landraces)
	(1), Cucumis melo (1), Cucurbita ficifolia (1), Cucurbita maxima (1), Hordeum vulgare subsp. vulgare (1), Lathyrus sativus (1), Lupinus luteus (1), Petroselinum crispum (1), Phaseolus coccineus (1), Prunus domestica (1), Prunus persica (1)	
Évora (387)	*Phaseolus vulgaris (53), Vitis vinifera (51), Pyrus communis (33), Lactuca sativa (19), Prunus domestica (17), Prunus persica (16), Cucumis melo (15), Ficus carica (15), Olea europeae (13), Allium sativum (12), Capsicum annuum (12), Zea mays subsp. mays (11), Citrus sinensis (7), Cucurbita pepo (7), Solanum lycopersicum var. lycopersicum (7), Vigna unguiculata (7), Vicia faba (6), Allium cepa (5), Cucumis sativus (5), Punica granatum (5), Vigna unguiculata subsp. sesquipedalis (5), Cicer arietinum (4), Malus domestica (4), Brassica rapa (3), Coriandrum sativum (3), Nasturtium officinale (3), Brassica oleracea var. acephala (2), Citrullus lanatus (2), Citrus limon (2), Citrus reticulata (2), Cucurbita maxima (2), Cucurbita moschata (2), Diospyrus kaki (2), Hordeum vulgare subsp. vulgare (2), Ipomoea batatas (2), Lathyrus sativus (2), Petroselinum crispum (2), Pisum sativum (2), Portulaca oleracea (2), Spinacia oleracea (2), Allium ascalonicum (1), Arbutus unedo (1), Avena sativa (1), Beta vulgaris subsp. vulgaris (2), Brassica oleracea var. costata (1), Brassica oleracea var. italica (1), Brassica rapa subsp. rapa (1), Ceratonia siliqua (1), Cucurbita ficifolia (1), Cydonia oblonga (1), Eriobotrya japonica (1), Lagenaria siceraria (1), Laurus nobilis (1), Lepidium sativum (1), Prunus cerasus (1), Raphanus sativus (1), Secale cereale (1), Solanum lycopersicum (1), Triticum aestivum subsp. aestivum (1), Vicia sativa (1)*	*Avena sp. (1), Lathyrus sp. (1)*

District	Taxa* (number of landraces)	Genera (number of landraces)
Faro (553)	*Phaseolus vulgaris (99), Ficus carica (54), Zea mays subsp. mays (50), Vitis vinifera (44), Pyrus communis (42), Prunus domestica (34), Solanum lycopersicum var. lycopersicum (33), Prunus dulcis (18), Citrus sinensis (15), Pisum sativum (15), Capsicum annuum (12), Cucurbita pepo (11), Avena sativa (10), Olea europeae (8), Cucumis sativus (6), Vicia faba (6), Citrus reticulata (5), Lactuca sativa (5), Lathyrus sativus (5), Lupinus albus (5), Triticum aestivum (5), Citrullus lanatus (4), Ipomoea batatas (4), Brassica oleracea var. acephala (3), Cicer arietinum (3), Citrus limon (3), Cucumis melo (3), Cucurbita ficifolia (3), Cucurbita maxima (3), Cucurbita moschata (3), Prunus persica (3), Punica granatum (3), Allium sativum (2), Coriandrum sativum (2), Daucus carota (2), Hordeum vulgare subsp. vulgare (2), Lagenaria siceraria (2), Malus domestica (2), Prunus armeniaca (2), Raphanus sativus (2), Secale cereale (2), Vigna unguiculata (2), Allium cepa (1), Arachis hypogaea (1), Brassica oleracea var. capitata (1), Cajanus cajan (1), Ceratonia siliqua (1), Cydonia oblonga (1), Eriobotrya japonica (1), Lepidium sativum (1), Linum usitatissimum (1), Luffa aegyptiaca (1), Lupinus luteus (1), Petroselinum crispum (1), Phaseolus coccineus (1), Triticum turgidum subsp. durum (1), Vicia ervilia (1), Vigna unguiculata subsp. sesquipedalis (1)*	*Cucurbita sp. (1), Brassica sp. (1)*
Guarda (367)	*Vitis vinifera (129), Phaseolus vulgaris (55), Zea mays subsp. mays (24), Brassica rapa subsp. rapa (13), Olea europeae (12), Brassica rapa (10), Vigna unguiculata (9), Pisum sativum (8), Lactuca sativa (7), Malus domestica (7), Allium cepa (6), Brassica oleracea var. acephala (6), Solanum lycopersicum var. lycopersicum (6), Capsicum annuum (5), Cucurbita pepo (5), Prunus dulcis (5), Triticum aestivum subsp. aestivum (5), Brassica oleracea var. costata (4), Raphanus sativus (4), Allium sativum (3), Brassica napus (3), Cicer arietinum (3), Avena sativa (2), Capsicum frutescens (2), Coriandrum sativum (2), Cucumis melo (2), Cucumis sativus (2), Cucurbita maxima (2), Cucurbita moschata (2), Hordeum vulgare subsp. vulgare (3), Secale cereale (2), Vicia faba (2), Avena strigosa (1), Beta vulgaris subsp. vulgaris (1), Citrullus lanatus (1), Lagenaria siceraria (1), Lathyrus cicera (1), Linum usitatissimum (1), Lolium multiflorum (1), Lupinus albus (1), Lupinus luteus (1), Nasturtium officinale (1), Petroselinum crispum (1), Phaseolus coccineus (1), Setaria italica (1), Spinacia oleracea (1), Triticum turgidum subsp. durum (1), Vicia ervilia (1), Vigna unguiculata subsp. sesquipedalis (1)*	*Cucurbita sp. (4), Brassica sp. (3), Lolium sp. (2), Lathyrus sp. (1), Ornithopus sp. (1), Vicia sp. (1)*

District	Taxa* (number of landraces)	Genera (number of landraces)
Leiria (200)	*Vitis vinifera (83), Phaseolus vulgaris (22), Pyrus communis (20), Zea mays subsp. mays (12), Brassica rapa (8), Vigna unguiculata (7), Brassica oleracea var. acephala (4), Lactuca sativa (4), Lathyrus sativus (4), Cucurbita pepo (3), Malus domestica (3), Solanum lycopersicum var. lycopersicum (4), Vicia faba (3), Allium cepa (2), Avena sativa (2), Brassica oleracea var. italica (2), Cucurbita ficifolia (2), Olea europeae (2), Pisum sativum (2), Solanum melongena (2), Beta vulgaris subsp. vulgaris (1), Brassica oleracea var. costata (1), Cicer arietinum (1), Coriandrum sativum (1), Cucurbita maxima (1), Linum usitatissimum (1), Lupinus albus (1), Nasturtium officinale (1), Secale cereale (1)*	*Cucurbita sp. (3)*
Lisbon (97)	*Vitis vinifera (56), Phaseolus vulgaris (7), Capsicum frutescens (3), Cicer arietinum (3), Vicia faba (3), Capsicum annuum (2), Cucumis melo (2), Allium cepa (1), Brassica rapa subsp. rapa (1), Capsicum chinense (1), Cucumis sativus (1), Cucurbita ficifolia (1), Cucurbita moschata (1), Cucurbita pepo (1), Lathyrus tingitanus (1), Lupinus albus (1), Satureja hortensis (1), Solanum lycopersicum var. lycopersicum (9), Vicia sativa (1), Zea mays subsp. mays (1)*	*Triticum sp. (49), Medicago sp. (1)*
Madeira Autonomous Region (274)	*Phaseolus vulgaris (72), Zea mays subsp. mays (34), Triticum aestivum subsp. aestivum (22), Vitis vinifera (13), Vicia faba (11), Ficus carica (8), Pisum sativum (8), Solanum tuberosum (8), Colocasia esculenta (7), Triticum aestivum subsp. compactum (6), Triticum turgidum subsp. durum (6), Brassica oleracea (5), Cucurbita pepo (5), Allium cepa (4), Capsicum baccatum (4), Ipomoea batatas (4), Phaseolus coccineus (4), Capsicum annuum (3), Cucurbita ficifolia (3), Hordeum vulgare subsp. vulgare (4), Linum usitatissimum (3), Lupinus albus (3), Saccharum officinarum (3), Solanum lycopersicum var. lycopersicum (5), Annona cherimola (2), Capsicum frutescens (2), Carica papaya (2), Cucumis sativus (2), Lactuca sativa (2), Malus domestica (2), Passiflora edulis (2), Beta vulgaris subsp. vulgaris (1), Brassica oleracea var. acephala (1), Citrus limon (1), Citrus sinensis (1), Cucurbita moschata (1), Foeniculum vulgare (1), Lathyrus sativus (1), Nasturtium officinale (1), Passiflora ligularis (1), Persea americana (1), Petroselinum crispum (1), Pyrus communis (1), Secale cereale (1), Triticum turgidum subsp. turgidum (1), Vicia sativa (1)*	*Cucurbita sp. (26), Triticum sp. (12), Brassica sp. (8), Musa sp. (3), Passiflora sp. (3), Allium sp. (2), Capsicum sp. (2), Avena sp. (1), Phaseolus sp. (1), Physalis sp. (1)*

District	Taxa* (number of landraces)	Genera (number of landraces)
Portalegre (215)	*Phaseolus vulgaris (44), Vitis vinifera (35), Zea mays subsp. mays (12), Solanum lycopersicum var. lycopersicum (12), Cucurbita moschata (9), Citrullus lanatus (8), Olea europeae (8), Allium sativum (7), Cucurbita pepo (7), Brassica oleracea var. acephala (5), Vicia faba (5), Capsicum annuum (4), Capsicum frutescens (4), Cucumis melo (4), Cucurbita maxima (4), Triticum aestivum subsp. aestivum (4), Allium cepa (3), Brassica rapa (3), Vigna unguiculata (3), Brassica oleracea var. costata (2), Capsicum chinense (2), Cicer arietinum (2), Coriandrum sativum (2), Lactuca sativa (2), Lagenaria siceraria (2), Pisum sativum (2), Raphanus sativus (2), Vigna unguiculata subsp. sesquipedalis (2), Allium ampeloprasum (1), Avena sativa (1), Citrus sinensis (1), Corylus avellana (1), Cucumis sativus (1), Cydonia oblonga (1), Lupinus albus (1), Malus domestica (1), Nasturtium officinale (1), Petroselinum crispum (1), Portulaca oleracea (1), Prunus domestica (1), Punica granatum (1), Pyrus communis (1), Satureja hortensis (1), Secale cereale (1)*	*Brassica sp. (3), Avena sp. (1), Cucurbita sp. (1)*
Porto (181)	*Vitis vinifera (45), Zea mays subsp. mays (35), Phaseolus vulgaris (22), Allium cepa (7), Lolium multiflorum (7), Lactuca sativa (6), Brassica rapa subsp. rapa (5), Allium sativum (4), Brassica oleracea var. costata (4), Capsicum annuum (4), Cucurbita pepo (4), Linum usitatissimum (4), Pisum sativum (4), Corylus avellana (3), Solanum lycopersicum var. lycopersicum (4), Brassica oleracea var. acephala (2), Cucumis melo (2), Phaseolus coccineus (2), Vigna unguiculata (2), Avena sativa (1), Avena strigosa (1), Brassica napus (1), Capsicum frutescens (1), Citrullus lanatus (1), Cucumis sativus (1), Cucurbita ficifolia (1), Cucurbita maxima (1), Malus domestica (1), Ornithopus sativus (1), Petroselinum crispum (1), Pyrus communis (1), Secale cereale (1), Solanum melongena (1), Vicia faba (1)*	
Santarém (369)	*Phaseolus vulgaris (89), Vitis vinifera (56), Lactuca sativa (21), Zea mays subsp. mays (21), Brassica rapa (18), Brassica oleracea var. acephala (14), Solanum lycopersicum var. lycopersicum (19), Cucumis melo (10), Allium cepa (8), Vicia faba (8), Cucurbita maxima (7), Raphanus sativus (6), Vigna unguiculata (6), Capsicum annuum (5), Capsicum frutescens (5), Cucurbita moschata (5), Cucurbita pepo (5), Allium sativum (4), Brassica oleracea var. costata (4), Citrullus lanatus (4), Lupinus albus (4), Olea europeae (4), Cucurbita ficifolia (3), Lagenaria siceraria (3), Phaseolus coccineus (3), Allium ampeloprasum (2), Capsicum chinense (2), Cucumis melo var. reticulatus (2),*	*Cucurbita sp. (4), Lathyrus sp. (1), Vicia sp. (1)*

District	Taxa* (number of landraces)	Genera (number of landraces)
	Cucumis sativus (2), Ficus carica (2), Pisum sativum (2), Avena sativa (1), Beta vulgaris subsp. vulgaris (2), Brassica oleracea var. capitata (1), Brassica oleracea var. sabauda (1), Cicer arietinum (1), Cichorium endivia (1), Citrus reticulata (1), Citrus sinensis (1), Coriandrum sativum (1), Daucus carota (1), Helianthus annuus (1), Hordeum vulgare subsp. vulgare (1), Lathyrus sativus (1), Lepidium sativum (1), Lupinus luteus (1), Malus domestica (1), Nasturtium officinale (1), Ocimum basilicum (1), Petroselinum crispum (1), Satureja hortensis (1), Setaria italica (1), Sorghum bicolor (1), Spinacia oleracea (1), Triticum turgidum subsp. durum (1)	
Setúbal (360)	*Vitis vinifera (223), Phaseolus vulgaris (33), Malus domestica (11), Zea mays subsp. mays (8), Pyrus communis (6), Allium cepa (5), Citrus sinensis (5), Ficus carica (4), Solanum lycopersicum var. lycopersicum (6), Avena sativa (3), Brassica oleracea var. acephala (3), Citrullus lanatus (3), Olea europeae (3), Prunus persica (3), Allium sativum (2), Capsicum annuum (2), Cicer arietinum (2), Coriandrum sativum (2), Cucurbita pepo (2), Lathyrus sativus (2), Prunus domestica (2), Punica granatum (2), Apium graveolens (1), Brassica oleracea var. costata (1), Brassica rapa (1), Capsicum frutescens (1), Castanea sativa (1), Citrus reticulata (1), Cucumis melo (1), Cucurbita ficifolia (1), Cucurbita maxima (1), Cucurbita moschata (1), Cydonia oblonga (1), Helianthus annuus (1), Lactuca sativa (1), Lagenaria siceraria (1), Lepidium sativum (1), Lupinus albus (1), Lupinus luteus (1), Morus nigra (1), Nasturtium officinale (1), Petroselinum crispum (1), Pisum sativum (1), Prunus avium (1), Prunus cerasus (1), Raphanus sativus (1), Spinacia oleracea (1), Triticum turgidum subsp. durum (1), Vicia faba (1), Vigna unguiculata (1)*	*Lathyrus sp. (1)*
Viana do Castelo (330)	*Phaseolus vulgaris (109), Zea mays subsp. mays (57), Vitis vinifera (45), Lactuca sativa (12), Pisum sativum (11), Phaseolus coccineus (8), Allium sativum (7), Malus domestica (7), Secale cereale (7), Allium cepa (6), Lolium multiflorum (6), Brassica rapa (5), Cucurbita pepo (5), Vicia faba (5), Solanum lycopersicum var. lycopersicum (6), Brassica rapa subsp. rapa (3), Linum usitatissimum (3), Avena sativa (2), Brassica oleracea var. acephala (2), Corylus avellana (2), Cucurbita maxima (2), Pyrus communis (2), Vigna unguiculata (2), Avena strigosa (1), Brassica napus (1), Brassica oleracea var. costata (1), Capsicum baccatum (1), Citrus sinensis (1), Cucumis melo (1), Cucumis sativus (1), Cucurbita ficifolia (1), Holcus lanatus (1), Hordeum vulgare subsp. vulgare (1),*	*Brassica sp. (1), Cucurbita sp. (1)*

District	Taxa* (number of landraces)	Genera (number of landraces)
	Lagenaria siceraria (1), Lupinus albus (1), Lupinus luteus (1), Ornithopus compressus (1), Ornithopus sativus (1), Petroselinum crispum (1)	
Vila Real (370)	*Vitis vinifera (132), Phaseolus vulgaris (66), Zea mays subsp. mays (23), Brassica rapa subsp. rapa (16), Pyrus communis (16), Allium cepa (9), Olea europeae (8), Phaseolus coccineus (8), Malus domestica (7), Pisum sativum (7), Brassica oleracea var. costata (6), Cucurbita pepo (6), Brassica rapa (5), Lactuca sativa (5), Prunus dulcis (5), Vigna unguiculata (5), Linum usitatissimum (4), Corylus avellana (3), Secale cereale (3), Triticum aestivum subsp. aestivum (3), Allium sativum (2), Brassica napus (2), Brassica oleracea var. capitata (2), Capsicum annuum (2), Cucumis sativus (2), Cucurbita maxima (2), Holcus lanatus (2), Lolium multiflorum (2), Solanum lycopersicum var. lycopersicum (3), Vicia faba (2), Avena sativa (1), Brassica oleracea var. acephala (1), Brassica oleracea var. sabauda (1), Capsicum frutescens (1), Cicer arietinum (1), Cucurbita ficifolia (1), Daucus carota (1), Lupinus luteus (1), Petroselinum crispum (1), Plantago lanceolata (1), Setaria italica (1), Spinacia oleracea (1)*	*Lathyrus sp. (1), Lolium sp. (1), Lupinus sp. (1)*
Viseu (462)	*Vitis vinifera (148), Phaseolus vulgaris (75), Malus domestica (68), Zea mays subsp. mays (33), Cucurbita pepo (10), Pyrus communis (10), Brassica rapa (9), Pisum sativum (9), Brassica oleracea var. acephala (8), Brassica rapa subsp. rapa (8), Allium sativum (7), Phaseolus coccineus (6), Allium cepa (5), Brassica oleracea var. costata (5), Prunus dulcis (5), Linum usitatissimum (4), Capsicum annuum (3), Corylus avellana (3), Lactuca sativa (3), Lupinus albus (3), Solanum lycopersicum var. lycopersicum (5), Triticum aestivum subsp. aestivum (3), Vigna unguiculata (3), Avena sativa (2), Cicer arietinum (2), Cucumis sativus (2), Vicia faba (2), Asparagus officinalis (1), Beta vulgaris subsp. vulgaris (1), Brassica napus (1), Brassica oleracea var. capitata (1), Brassica oleracea var. sabauda (1), Capsicum frutescens (1), Cichorium intybus (1), Citrullus lanatus (1), Coriandrum sativum (1), Cucumis melo (1), Cucurbita maxima (1), Daucus carota (1), Hordeum vulgare subsp. vulgare (1), Lagenaria siceraria (1), Olea europeae (1), Raphanus sativus (1), Secale cereale (1), Setaria italica (1), Sinapis alba (1), Spinacia oleracea (1), Vigna unguiculata subsp. sesquipedalis (1)*	*Lolium sp. (3), Cucurbita sp. (2)*

*Includes entries at both species and sub-specific level.

Supplementary Table 2.7.: Number of landraces per taxa per undefined district.

Undefined district	Taxa (number of landraces)	Genus (number of landraces)
Azores Autonomous Region (53)	*Vitis vinifera (22), Phaseolus vulgaris (15), Malus domestica (9), Zea mays subsp. mays (2), Capsicum annuum (1), Lupinus albus (1), Passiflora edulis (1), Pisum sativum (1), Vicia faba (1)*	
NA (non available) (1)	*Corylus avellana (1)*	
Portugal (9)	*Zea mays subsp. mays (2), Capsicum annuum (1), Cicer arietinum (1), Cucumis melo (1), Daucus carota (1), Phaseolus vulgaris (1), Vicia faba (1), Vigna unguiculata (1)*	*Triticum sp. (1)*

Supplementary Table 2.8.: Number of landraces per crop group per district.

District	Group (number of landraces)	Number of landraces
Aveiro	Fruit and nuts (149), vegetables and melons (112), other crops (110), leguminous crops (106), cereals (60)	537
Azores Autonomous Region – Corvo island	Vegetables and melons (5), fruit and nuts (5), root/tuber crops (2), cereals (1), leguminous crops (1)	14
Azores Autonomous Region – Faial island	Vegetables and melons (26), leguminous crops (16), fruit and nuts (13), root/tuber crops (11), cereals (8), other crops (6)	80

District	Group (number of landraces)	Number of landraces
Azores Autonomous Region – Flores Island	Leguminous crops (21), vegetables and melons (19), fruit and nuts (15), cereals (14), root/tuber crops (9), other crops (1)	79
Azores Autonomous Region – Graciosa island	Leguminous crops (31), fruit and nuts (27), vegetables and melons (23), root/tuber crops (5), other crops (5), cereals (4)	95
Azores Autonomous Region – Pico island	Vegetables and melons (30), leguminous crops (26), fruit and nuts (19), root/tuber crops (4), cereals (4), other crops (3)	86
Azores Autonomous Region – S. Jorge island	Vegetables and melons (22), leguminous crops (18), fruit and nuts (11), root/tuber crops (8), cereals (5), other crops (3), beverage and spice crops (1)	68
Azores Autonomous Region – S. Miguel island	Vegetables and melons (35), leguminous crops (29), fruit and nuts (14), root/tuber crops (6), cereals (5), other crops (5), beverage and spice crops (1)	95
Azores Autonomous Region – Sta. Maria island	Leguminous crops (38), vegetables and melons (25), fruit and nuts (13), root/tuber crops (10), cereals (4), other crops (3)	93

District	Group (number of landraces)	Number of landraces
Azores Autonomous Region – Terceira island	Leguminous crops (41), vegetables and melons (28), root/tuber crops (12), fruit and nuts (8), cereals (7), other crops (5)	101
Beja	Vegetables and melons (62), other crops (49), cereals (26), leguminous crops (26), fruit and nuts (7), NA (non available) (1)	171
Braga	Leguminous crops (111), cereals (90), vegetables and melons (89), other crops (82), fruit and nuts (24)	396
Bragança	Vegetables and melons (282), leguminous crops (223), other crops (170), Fruit and nuts (113), Cereals (67)	855
Castelo Branco	Fruit and nuts (42), leguminous crops (39), vegetables and melons (26), cereals (21), other crops (10)	138
Coimbra	Leguminous crops (107), other crops (67), vegetables and melons (62), cereals (57), fruit and nuts (12)	305
Évora	Fruit and nuts (121), vegetables and melons (96), leguminous crops (79), other crops (74), cereals (17), root/tuber crops (2)	389
Faro	Fruit and nuts (192), leguminous crops (140), vegetables and melons (97), cereals (70), other crops (52), root/tuber crops (4)	555
Guarda	Other crops (144), vegetables and melons (91), leguminous crops (81), cereals (39), fruit and nuts (24)	379
Leiria	Other crops (85), leguminous crops (40), vegetables and melons (38), fruit and nuts (25), cereals (15)	203
Lisboa	Other crops (59), cereals (50), vegetables and melons (23), leguminous crops (14), NA (non available) (1)	147
Madeira Autonomous Region	Leguminous crops (100), cereals (87), vegetables and melons (75), fruit and nuts (28), other crops (19), root/tuber crops (19), beverage and spice crops (3), NA (non available) (2)	333

District	Group (number of landraces)	Number of landraces
Portalegre	Vegetables and melons (79), leguminous crops (59), other crops (48), cereals (19), fruit and nuts (15)	220
Porto	Other crops (62), vegetables and melons (45), cereals (38), leguminous crops (31), fruit and nuts (5)	181
Santarém	Vegetables and melons (156), leguminous crops (115), other crops (69), cereals (26), fruit and nuts (9)	375
Setúbal	Other crops (231), fruit and nuts (42), leguminous crops (42), vegetables and melons (34), cereals (12)	361
Viana do Castelo	Leguminous crops (137), cereals (68), other crops (65), vegetables and melons (50), fruit and nuts (12)	332
Vila Real	Other crops (146), leguminous crops (91), vegetables and melons (66), fruit and nuts (39), cereals (31)	373
Viseu	Other crops (164), leguminous crops (101), fruit and nuts (87), vegetables and melons (74), cereals (41)	467
Undefined districts		
Azores Autonomous Region	Other crops (22), Leguminous crops (18), Fruit and nuts (10), Cereals (2), Vegetables and melons (1)	53
NA (non available)	Fruit and nuts (1)	1
Portugal	Leguminous crops (4), Vegetables and melons (3), Cereals (3)	10
Total		**7,492**

Supplementary Table 4.3.: Allele frequencies, sample size (N) and total number of alleles per locus (TL) and per populations (TP).

Locus	Allele/n	1	2	3
PVctt001	**N**	153	62	58
	152	0.000	0.000	0.017
	153	0.278	0.129	0.267
	156	0.010	0.016	0.000
	159	0.049	0.016	0.026
	162	0.552	0.774	0.690
	163	0.065	0.065	0.000
	165	0.046	0.000	0.000
	TL	6	5	4
BMd1	**N**	139	61	53
	164	0.061	0.025	0.038
	165	0.040	0.049	0.019
	166	0.230	0.279	0.377
	167	0.266	0.098	0.179
	168	0.000	0.016	0.000
	172	0.000	0.000	0.038
	176	0.000	0.016	0.047
	178	0.025	0.074	0.047
	179	0.000	0.016	0.000
	180	0.169	0.221	0.123
	181	0.032	0.049	0.000
	182	0.061	0.131	0.132
	184	0.029	0.008	0.000
	186	0.047	0.016	0.000
	188	0.025	0.000	0.000
	192	0.014	0.000	0.000
	TL	12	13	9
BMd37	**N**	132	60	54
	126	0.386	0.567	0.704
	130	0.564	0.383	0.241
	132	0.045	0.050	0.056
	140	0.004	0.000	0.000
	TL	4	3	3
BMd41	**N**	154	62	58
	227	0.000	0.000	0.034
	236	0.032	0.016	0.043
	237	0.052	0.032	0.034
	240	0.253	0.016	0.060
	246	0.019	0.008	0.009
	249	0.617	0.863	0.819
	250	0.026	0.065	0.000
	TL	6	6	6
BM210	**N**	151	62	56
	159	0.000	0.000	0.018

	165	0.132	0.581	0.518
	166	0.007	0.016	0.000
	167	0.010	0.000	0.000
	169	0.533	0.339	0.339
	170	0.013	0.016	0.000
	173	0.036	0.000	0.000
	177	0.013	0.000	0.000
	179	0.079	0.016	0.063
	180	0.020	0.000	0.000
	181	0.099	0.016	0.036
	182	0.040	0.000	0.000
	184	0.007	0.000	0.000
	185	0.010	0.016	0.027
	TL	13	7	6
BM151	**N**	154	62	58
	138	0.000	0.000	0.017
	141	0.000	0.000	0.017
	143	0.088	0.000	0.000
	145	0.575	0.395	0.491
	147	0.032	0.016	0.000
	149	0.227	0.347	0.302
	150	0.065	0.242	0.172
	152	0.013	0.000	0.000
	TL	6	4	5
Pvat007	**N**	142	49	54
	190	0.000	0.000	0.019
	192	0.461	0.480	0.296
	194	0.070	0.143	0.204
	196	0.007	0.010	0.000
	197	0.018	0.000	0.000
	198	0.035	0.061	0.065
	199	0.000	0.020	0.000
	200	0.081	0.102	0.194
	201	0.007	0.000	0.000
	202	0.077	0.041	0.000
	203	0.018	0.000	0.000
	204	0.039	0.020	0.000
	206	0.025	0.000	0.046
	207	0.007	0.000	0.000
	208	0.014	0.000	0.019
	209	0.021	0.000	0.009
	210	0.011	0.000	0.000
	211	0.060	0.041	0.037
	212	0.004	0.000	0.000
	213	0.025	0.020	0.083
	215	0.007	0.000	0.028
	216	0.007	0.000	0.000
	217	0.004	0.020	0.000

	219	0.004	0.020	0.000
	220	0.000	0.020	0.000
	TL	22	13	11
GATS91	**N**	153	61	56
	218	0.105	0.115	0.152
	226	0.062	0.066	0.054
	228	0.003	0.016	0.000
	230	0.072	0.516	0.402
	232	0.010	0.000	0.000
	234	0.222	0.156	0.107
	236	0.176	0.049	0.063
	238	0.007	0.000	0.009
	240	0.007	0.000	0.000
	252	0.098	0.000	0.063
	254	0.033	0.016	0.000
	256	0.046	0.033	0.000
	257	0.013	0.000	0.000
	258	0.023	0.000	0.000
	259	0.105	0.033	0.143
	261	0.007	0.000	0.009
	263	0.013	0.000	0.000
	TL	17	9	9
AG1	**N**	154	62	58
	120	0.185	0.468	0.241
	134	0.808	0.532	0.733
	136	0.006	0.000	0.000
	138	0.000	0.000	0.026
	TL	3	2	3
BMd42	**N**	154	62	56
	149	0.279	0.371	0.411
	155	0.013	0.000	0.000
	157	0.127	0.081	0.134
	158	0.075	0.097	0.036
	159	0.403	0.258	0.241
	160	0.068	0.048	0.018
	161	0.010	0.048	0.125
	162	0.026	0.097	0.036
	TL	8	7	7
BMd15	**N**	142	59	55
	165	0.954	1.000	1.000
	200	0.028	0.000	0.000
	201	0.018	0.000	0.000
	TL	3	1	1
	TP	100	70	64

Supplementary Table 4.4.: Number of samples per allele per population and number of alleles per locus per population.

Locus	Allele	1	2	3	Average
PVctt001	152	0	0	1	
	153	43	8	15	
	156	1	1	0	
	159	8	1	1	
	162	85	48	40	
	163	10	4	0	
	165	7	0	0	
	No alleles	6	5	4	5
BMd1	164	8	1	2	
	165	6	3	1	
	166	32	17	20	
	167	37	6	10	
	168	0	1	0	
	172	0	0	2	
	176	0	1	3	
	178	4	4	3	
	179	0	1	0	
	180	23	13	6	
	181	4	3	0	
	182	8	8	7	
	184	4	1	0	
	186	6	1	0	
	188	4	0	0	
	192	2	0	0	
	No alleles	12	13	9	11
BMd37	126	51	34	38	
	130	74	23	13	
	132	6	3	3	
	140	1	0	0	
	No alleles	4	3	3	3
BMd41	227	0	0	2	
	236	5	1	2	
	237	8	2	2	
	240	39	1	3	
	246	3	0	0	
	249	95	53	48	
	250	4	4	0	
	No alleles	6	6	6	6
BM210	159	0	0	1	
	165	20	36	29	
	166	1	1	0	
	167	1	0	0	

	169	80	21	19	
	170	2	1	0	
	173	5	0	0	
	177	2	0	0	
	179	12	1	4	
	180	3	0	0	
	181	15	1	2	
	182	6	0	0	
	184	1	0	0	
	185	1	1	2	
	No alleles	13	7	6	9
BM151	138	0	0	1	
	141	0	0	1	
	143	13	0	0	
	145	89	24	29	
	147	5	1	0	
	149	35	21	17	
	150	10	15	10	
	152	2	0	0	
	No alleles	6	4	5	5
Pvat007	190	0	0	1	
	192	66	23	16	
	194	10	7	11	
	196	1	0	0	
	197	3	0	0	
	198	5	3	3	
	199	0	1	0	
	200	12	5	10	
	201	1	0	0	
	202	11	2	0	
	203	3	0	0	
	204	5	1	0	
	206	4	0	3	
	207	1	0	0	
	208	2	0	1	
	209	3	0	1	
	210	1	0	0	
	211	9	2	2	
	212	0	0	0	
	213	4	1	4	
	215	1	0	2	
	216	1	0	0	
	217	0	1	0	
	219	0	1	0	
	220	0	1	0	
	No alleles	22	13	11	15
GATS91	218	16	7	9	

	226	9	4	3	
	228	1	1	0	
	230	11	31	23	
	232	1	0	0	
	234	34	10	6	
	236	27	3	4	
	238	1	0	1	
	240	1	0	0	
	252	15	0	4	
	254	5	1	0	
	256	7	2	0	
	257	2	0	0	
	258	4	0	0	
	259	16	2	8	
	261	1	0	1	
	263	2	0	0	
	No alleles	17	9	9	12
AG1	120	28	29	14	
	134	124	33	43	
	136	1	0	0	
	138	0	0	1	
	No alleles	3	2	3	3
BMd42	149	43	23	23	
	155	2	0	0	
	157	19	5	8	
	158	12	6	2	
	159	62	16	13	
	160	10	3	1	
	161	1	3	7	
	162	4	6	2	
	No alleles	8	7	7	7
BMd15	165	136	59	55	
	200	4	0	0	
	201	3	0	0	
	No alleles	3	1	1	2
Total No alleles per population		100	70	64	78

No alleles Number of alleles per locus per population

Average Average number of alleles per locus per population

Supplementary Table 4.5.: Summary of private alleles per population.

Population	Locus	Allele	Frequency
1	PVctt001	165	0.046
1	BMd1	188	0.025
1	BMd1	192	0.014
1	BMd37	140	0.004
1	BM210	167	0.010
1	BM210	173	0.036
1	BM210	177	0.013
1	BM210	180	0.020
1	BM210	182	0.040
1	BM210	184	0.007
1	BM151	143	0.088
1	BM151	152	0.013
1	Pvat007	197	0.018
1	Pvat007	201	0.007
1	Pvat007	203	0.018
1	Pvat007	207	0.007
1	Pvat007	210	0.011
1	Pvat007	212	0.004
1	Pvat007	216	0.007
1	GATS91	232	0.010
1	GATS91	240	0.007
1	GATS91	257	0.013
1	GATS91	258	0.023
1	GATS91	263	0.013
1	AG1	136	0.006
1	BMd42	155	0.013
1	BMd15	200	0.028
1	BMd15	201	0.018
2	BMd1	168	0.016
2	BMd1	179	0.016
2	Pvat007	199	0.020
2	Pvat007	220	0.020
3	PVctt001	152	0.017
3	BMd1	172	0.038
3	BMd41	227	0.034
3	BM210	159	0.018
3	BM151	138	0.017
3	BM151	141	0.017
3	Pvat007	190	0.019
3	AG1	138	0.026

Supplementary Table 4.8.: Sample Size (N), No. Alleles (Na), No. Effective Alleles (Ne), Information Index (I), Observed Heterozygosity (Ho), Expected and Unbiased Expected Heterozygosity (He, uHe), and Fixation Index (F).

Population	Locus	N	Na	Ne	I	Ho	He	uHe	F
1	**PVctt001**	153	6	2.557	1.196	0.052	0.609	0.611	0.914
	BMd1	139	12	5.987	2.063	0.173	0.833	0.836	0.793
	BMd37	132	4	2.128	0.852	0.038	0.530	0.532	0.929
	BMd41	154	6	2.225	1.082	0.058	0.551	0.552	0.894
	BM210	151	13	3.107	1.633	0.040	0.678	0.680	0.941
	BM151	154	6	2.531	1.214	0.006	0.605	0.607	0.989
	Pvat007	142	22	4.176	2.106	0.183	0.761	0.763	0.759
	GATS91	153	17	7.981	2.318	0.183	0.875	0.878	0.791
	AG1	154	3	1.454	0.517	0.110	0.312	0.313	0.646
	BMd42	154	8	3.742	1.557	0.071	0.733	0.735	0.903
	BMd15	142	3	1.097	0.216	0.007	0.088	0.089	0.920
2	**PVctt001**	62	5	1.611	0.772	0.065	0.379	0.382	0.830
	BMd1	61	13	6.041	2.073	0.148	0.834	0.841	0.823
	BMd37	60	3	2.125	0.839	0.067	0.529	0.534	0.874
	BMd41	62	6	1.333	0.587	0.016	0.250	0.252	0.935
	BM210	62	7	2.207	1.015	0.032	0.547	0.551	0.941
	BM151	62	4	2.983	1.144	0.016	0.665	0.670	0.976
	Pvat007	49	13	3.697	1.819	0.122	0.729	0.737	0.832
	GATS91	61	9	3.190	1.565	0.049	0.687	0.692	0.928
	AG1	62	2	1.992	0.691	0.097	0.498	0.502	0.806
	BMd42	62	7	4.271	1.666	0.032	0.766	0.772	0.958
	BMd15	59	1	1.000	0.000	0.000	0.000	0.000	#N/D
3	**PVctt001**	58	4	1.825	0.773	0.017	0.452	0.456	0.962
	BMd1	53	9	4.658	1.811	0.132	0.785	0.793	0.832
	BMd37	54	3	1.798	0.751	0.037	0.444	0.448	0.917
	BMd41	58	6	1.473	0.742	0.034	0.321	0.324	0.893

BM210	56	6	2.567	1.169	0.018	0.610	0.616	0.971
BM151	58	5	2.756	1.154	0.052	0.637	0.643	0.919
Pvat007	54	11	5.456	1.942	0.204	0.817	0.824	0.751
GATS91	56	9	4.401	1.757	0.125	0.773	0.780	0.838
AG1	58	3	1.678	0.665	0.086	0.404	0.408	0.787
BMd42	56	7	3.799	1.548	0.036	0.737	0.743	0.952
BMd15	55	1	1.000	0.000	0.000	0.000	0.000	#N/D

Na No. of Different Alleles

Ne No. of Effective Alleles = 1 / (Sum pi^2)

I Shannon's Information Index = -1* Sum (pi * Ln (pi))

Ho Observed Heterozygosity = No. of Hets / N

He Expected Heterozygosity = 1 - Sum pi^2

uHe Unbiased Expected Heterozygosity = (2N / (2N-1)) * He

F Fixation Index = (He - Ho) / He = 1 - (Ho / He)

Where pi is the frequency of the ith allele for the population & Sum pi^2 is the sum of the squared population allele frequencies.