ISSN 1479-2621

Plant Genetic Resources: Characterization and Utilization 6(2); 98–112
doi:10.1017/S1479262108993138

Approaches and constraints of using existing landrace and extant plant material to understand agricultural spread in prehistory

Huw Jones[1†], Diane L. Lister[2†], Mim A. Bower[2], Fiona J. Leigh[1], Lydia M. Smith[1]* and Martin K. Jones[3]

[1]*NIAB, Huntingdon Road, Cambridge CB3 0LE, UK,* [2]*McDonald Institute for Archaeological Research, University of Cambridge, Downing Street, Cambridge CB2 3ER, UK and* [3]*Department of Archaeology, University of Cambridge, Downing Street, Cambridge CB2 3DZ, UK*

Received 19 December 2006; Accepted 25 October 2007; First published online 23 May 2008

Abstract

The potential for the phylogeographical analysis of cereal landraces to determine the initial patterns of agricultural spread through Europe is discussed in relation to two of the first cereals to be domesticated, emmer wheat (*Triticum turgidum* subsp. *dicoccum*) and barley (*Hordeum vulgare*). Extant landraces available from germplasm collections have a patchy distribution, largely being confined to regions of rugged upland topography, and the phylogeographical patterns observed may be due to 'overstamping' by more recent crop movements. Phylogeographical studies of non-viable historical landrace material held in herbarium and old seed collections and found in historical buildings have the potential to fill in the gaps in time and space. We explore the importance of precise geographical provenance and the limitations of this in extant and historical material. Additionally, we consider the effect of various chemicals and the preservation of DNA in the historical material.

Keywords: ancient DNA; germplasm; herbarium specimens; barley; emmer wheat; landraces

Introduction

The study of plant genetics has the potential to play a key role in aiding our understanding of the domestication of cereal crops in Europe, not only with regards to the origin of European agriculture, but also to its spread and establishment. This paper considers the scope and quality of plant material that can be accessed and reviews potential constraints of its use in a phylogeographical study. A central resource for this is provided by collections of cereal landraces that conserve considerable genetic diversity as well as a great deal of biogeographical information. More recently, this important source of data has been complemented by older specimens, as a result of the development of ancient DNA techniques. These specimens potentially include: herbarium specimens and old seed collections, which may go back at least three centuries; crop products used in buildings, especially roofing and mudbrick, which may go back a millennium; archaeological specimens (Fig. 1). The extraction of genetic information from the historical and ancient material is more challenging and time consuming than from viable landrace accessions, but is essential in filling in the gaps left by the study of extant landraces. These gaps are most obviously not only in time, but also in space; the geographical distribution of extant landraces is patchy.

This paper explores the patchy distribution of extant landraces with reference to two of the most important cereals in the early development of agriculture, emmer wheat [*Triticum turgidum* subsp. *dicoccum* (Schrank ex Schlüb.) Thell.] and barley (*Hordeum vulgare* Koch),

* Corresponding author. E-mail: lydia.smith@niab.com
†These two authors are equally contributed to this paper.

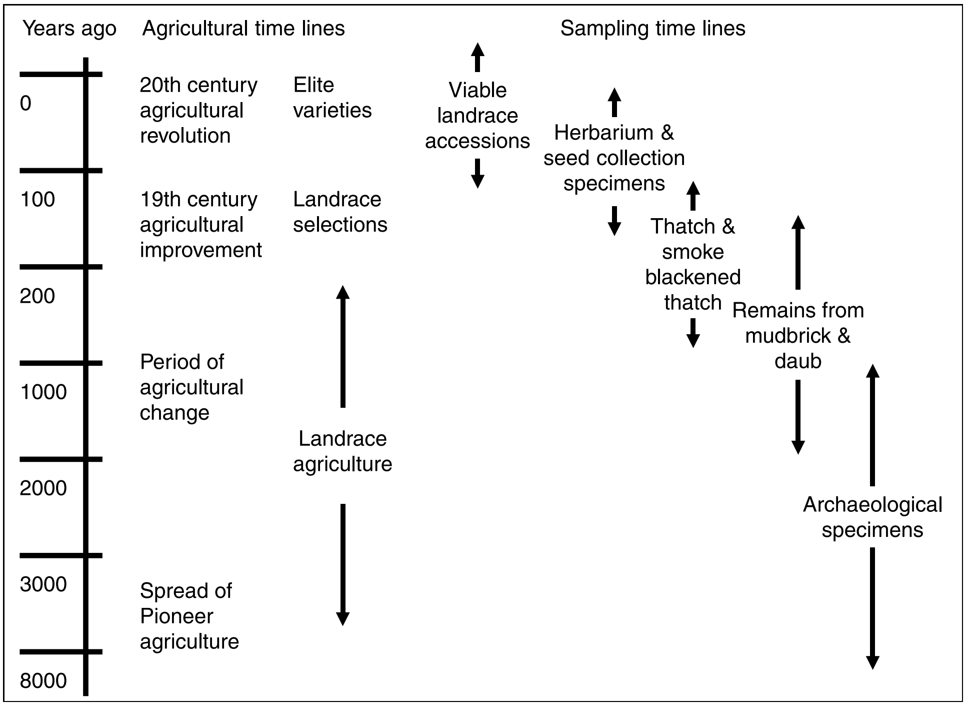

Fig. 1. Time-line of sources of cereal landraces and associated agrarian changes. Extant cereal landraces are found *in situ* and in germplasm collections. Desiccated historical material is found in herbarium collections, old seed collections and historical buildings (thatch, smoke-blackened thatch, daub and mudbricks). Material older than 1000 AD is generally found in archaeological sites and is mostly charred or waterlogged.

and considers the potential for filling the gaps through the analysis of historical and ancient specimens. Of these two founder crops, one (barley) remains a major world crop, while the other (emmer wheat) has long since diminished in importance, providing a valuable contrast for this exercise.

Background

There have been a number of successful genetic studies of wild progenitors of domesticated cereals looking at questions of their origin (Heun *et al.*, 1997; Badr *et al.*, 2000; Bertin *et al.*, 2001), lending support to the suggestion that it is also possible to use phylogeographical techniques to reveal details of the spread of cereals using landrace material, albeit over a shorter timescale. This requires that the geographical locations of the landraces have remained relatively static between their original establishment 8500–5500 years ago and their sampling during the last 100 years. Results presented by Allaby *et al.* (1999) and Brown *et al.* (2006) suggest that, in some parts of Europe at least, this is the case. These studies showed that allele distributions for two of the high-molecular-weight glutenin loci of wheat display geographical partitioning that may reflect differences between the genotypes of the plants that followed the central European and Mediterranean trajectories of agricultural spread into Europe. These results suggest that European wheat landraces may retain, at least to some extent, phylogeographical information that pertains to the events that took place during the domestication of cereal crops in Europe.

The potential of landraces

There is some debate in the research community as to the precise definition of a landrace, and opinions differ subtly from each other. However, for the purposes of this paper, we will use the following definition: a landrace is a 'dynamic population or populations 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).

Phylogeographical studies such as those described above require the genetic analysis of landraces rather than modern cultivars. This is because the landraces are historically associated with a specific geographical location, whereas the cultivars are bred remotely and subsequently cultivated in diverse locations. In some cases, it may be possible to include early selected varieties if pedigree data are known. The relationship between landraces, traditional varieties and early selected varieties

https://doi.org/10.1017/S1479262108993138 Published online by Cambridge University Press

presents interesting problems. The term variety, when applied to a population of an agricultural cultivar, implies a measure of seed improvement such as the activity in selecting from a wider population and some degree of homogeneity within all the populations bearing the same name. By contrast, 'traditional' varieties, such as the early German variety Bavaria, represent a subset of genotypes from regional landraces and may, in some cases, be used as proxies for them.

Use of viable landrace material

Modern cereal varieties are the product of two processes: selection from landrace populations by the 'seed improvers' from the 1850s onwards, followed by systematic breeding during the 20th century. The replacement of cereal landraces first with selections and then with more modern varieties started during the mid- to late 19th century. The original drive towards seed improvement focussed on the differing plant types that were present in the local landrace populations. Landrace varieties from other regions within Europe were collected, and the performance of isolated phenotypes was systematically compared. Favoured selections were multiplied and sold on to client farmers. These named selections replaced local landraces in time. During the second half of the 19th century, seed stocks from improved selections were distributed more widely across Europe by growers along with the knowledge of techniques used to produce them (Bonjean and Angus, 2000). Experimental or enterprising growers made further selections from the introduced stocks or from the local landraces and, on this empirical basis, the seed improvement and plant breeding industry came into existence. The economic benefits from the new methods were such that landrace cultivation had virtually disappeared from central and north-western Europe by the early decades of the 20th century (Bonjean and Angus, 2000).

Constraints and potential limitations to the use of viable accessions

Landraces indigenous to many areas have largely been lost because no systematic collections were made during the time before their demise. There are exceptions to these general trends within Europe, whereas in much of northern and central Europe, some regions did not industrialize so early or to the same extent. In these remote or agriculturally marginal regions, landraces survived into the later parts of the 20th century, for example, Iberia, the Balkans and large parts of European Russia (Bonjean and Angus, 2000). The efforts of collectors such as N. I. Vavilov have prevented the extinction of some of these local ecotypes. Wheat and barley landraces are also still cultivated on significant acres across other parts of the world, in inverse proportion to the extent of 20th-century agricultural modernization.

Conserved landraces offer an opportunity to sample the geographical distributions of genetic diversity that existed before the advent of elite varieties. Collections of landraces made in the 20th century must, however, be treated with caution since admixture with cultivars from neighbouring regions may have occurred. Although cross-pollination is minimal in inbreeding crops such as wheat and barley, accidental mixing by technicians could also occur, such as through reuse of bags still containing seeds. Similarly, viable landrace seeds held in germplasm collections need periodic regeneration; in this process, seeds are grown, the progeny harvested and the original accession either supplemented or replaced. The germplasm may be vulnerable to introgression from neighbouring plots of the same species during this process (Zeven, 1996). This risk of gene flow is minimized by maintaining a set separation distance between plots and bagging plants in outcrossing species. The integrity of regenerated material is assured by checking the growing plots and the resulting seeds against reference descriptions. These measures assure the genetic integrity of accessions (IPGRI, 1994; M. Ambrose, pers. commun.). Additionally, there is some duplication of accessions both within and between germplasm collections; this gives the investigator the opportunity to reacquire a landrace from an alternative regeneration cycle or alternative germplasm collection, should the data from an accession need checking.

The early selections and later crosses that produced elite cultivars have resulted in some landraces being widely represented in modern elite varieties, effectively giving them a far wider geographical distribution than that of the original landrace. The early German variety Bavaria, for example, was developed by selection from landraces in southern Germany during the late 1890s and appears in the pedigree of three varieties, including Isaria, which in turn appear in the pedigree of 25 varieties and so on. Some of these varieties were commercially successful and were consequently widely grown both in Germany and elsewhere in Europe (Baumer and Göppel, 1998, Baumer and Cais, 2000).

Non-viable historical landraces

The reasons for considering historical material extend beyond filling the gaps in space – they also concern disjunctions in time. Conclusions regarding prehistorical patterns of agricultural spread derived purely from genetic analysis of modern germplasm may be erroneous; phylogeographical patterns seen in modern landraces may not be a true reflection of the patterns formed during prehistory, but may be due to 'overstamping' by more recent events. These events, which could include later prehistorical movements of people and crops, documented periods of cross-continental trade and exchange of crops, connected to classical, mediaeval and early historical trading patterns; world trade and scientific breeding of the 19th and 20th centuries such as that described above. The addition of genetic data derived from historical and archaeological landrace specimens to that gained from modern landraces should reflect genetic patterns established further back in time.

Biogeographical studies of wild progenitors have often equated the current range of the plant with the domestication event, but this assumes that the biogeography of the species has remained the same since domestication. For example, Brown (1999b) argues that data derived from ancient DNA from crop remains is essential if we are to correctly infer the time depth of biogeographical patterns observed in modern landraces. Likewise, we propose that DNA derived from preserved non-viable historical landrace material can add important information to such phylogeographical studies.

Historical cereal specimens preserved in various collections across Europe often represent numerous landraces that are no longer extant and therefore are not represented in germplasm collections. Inclusion of such non-viable specimens in genetic studies gives a vital link with historical distributions of cereal landraces. Additionally, genetic data derived from non-extant historical landrace material can add important information, such as calculations of DNA mutation rates, for example, to phylogeographical studies.

Archaeological material would be optimal for this exercise; however, DNA preservation in very old samples is unpredictable (Reed *et al.*, 2003) and even where DNA persists, a considerable amount of genetic information is lost (Alonso *et al.*, 2004). Historical cereal landrace material, sourced from collections dating from early mediaeval times until the mid-20th century, is more readily available than archaeological material, and generally has better DNA preservation (Cota-Sanchez *et al.*, 2006; Li *et al.*, 2006; Lister *et al.*, 2008).

Landraces that are from distinct locations and have been preserved in historical collections are more likely to be genetically distinct than those that have been brought together and propagated in the interim in genebanks. Errors that may have occurred during planting out, harvesting and recording details will also contribute to the lack of genetic integrity of landrace germplasm.

Constraints to the use of non-viable historical landrace material

The limitations in the use of landrace accessions, in terms of geographically defined provenance, is particularly true for historical material, where original location information is often vague and, in many cases, non-existent. In older herbarium collections, it was apparent that the collector was primarily interested in the form of the plant, but not the precise location, environment and when or where it was grown. The information associated with the wheat specimens contained in the Percival Collection, for example, rarely contains more than the country of origin.

Labels on older herbarium specimens are usually very brief in terms of the accompanying information that they record, are often in illegible handwriting and are, in many cases, the only source of information that is available for that specimen. It is rarely specified whether the specimen is a landrace, though for older herbarium specimens they are more likely to be landraces than cultivars. Where collection information is given on the label, many record the name of a botanic garden or agricultural institution; it is possible that such specimens were collected from a distant location and brought back to a botanic garden where it was propagated. Any original geographical information was either never recorded or lost at a later date. Some labels also describe the specimen as being a 'ruderal'; a weedy form or growing in a location such as would cause it to be described as a weed, e.g. by a roadside or a port. Accessions labelled as ruderal are, in some cases, morphologically similar to old varieties, and may well represent landraces originating from countries far removed from where they were collected. Alternatively, they may represent backcrosses of cultivars with wild plants and should therefore not be used in phylogeographical studies.

Data and information on viable landraces collections are usually easily accessible, but access to non-viable historical material requires painstaking research to locate herbarium and museum collections and to determine whether the collection holds material of interest and its availability. Once suitable samples are located, many herbarium and museum curators are understandably hesitant about the destructive sampling of valuable historical material, and it is important to respect their decision to refuse sampling. However, at times, a curator's reluctance may be due to a lack of clarity about how much material needs to be taken; thus, it is very important for researchers using historical material for genetic studies to have thoroughly optimized DNA extraction and amplification techniques to minimize sample requirements.

If herbarium material is to be used for DNA studies, then the effects of various pesticides, which were often applied at the time of addition to the collection

https://doi.org/10.1017/S1479262108993138 Published online by Cambridge University Press

in order to deter fungal and insect damage during storage, must also be considered. Chemicals such as mercuric salts have been shown to damage DNA (Brown, 1999c). More recently, herbarium curators have realized the danger of using such chemicals and use other less damaging methods for pest control to the specimen, its DNA and the people working with the collection. Modern cabinets are designed to be insect-proof, and specimens coming into herbarium collections are often frozen in order to kill insect pests (G. Murrell, pers. commun.).

In addition to chemicals added after sampling, the samples themselves may include substances that damage the seeds' DNA or inhibit genetic analysis. The soot that encrusts smoke-blackened thatch (SBT) not only contains many inhibitors to the amplification of DNA, but the DNA itself may be considerably damaged by exposure to soot and high temperatures (Zhang and Wu, 2005). Daub material may also contain substances inhibitory to amplification reactions from clay, degraded plant material or animal dung (e.g. humic acid and phenolic compounds; Hänni *et al.*, 1995). Protocols suitable for the removal of inhibitory substances and genetic marker sets suitable for ancient DNA that is degraded to different extents must be considered carefully when doing genetic studies of historical materials.

Unfortunately, many herbarium collections throughout the world are in a very poor condition, damaged by water and ravaged by insects and mould. Such collections have often been neglected due to lack of funds, lack of recognition or misunderstanding of their historical heritage and scientific value, especially for the more recent advent of DNA analysis of herbarium materials (Morrison, 2001). We hope that papers such as this one will emphasize the great importance of these historical collections to the future study of biodiversity and sustainability of crop genetic resources.

Sources of landrace material

Much of the material useful in a study of the phylogeography of cereals is available as viable seeds. Considerable international effort is made to conserve the genetic diversity of economically important plant species, either in germplasm collections or *in situ*. The activities of these collections are coordinated by the Biodiversity International. Biodiversity International coordinates the information on conservation activities, including a searchable online database of germplasm collections (this can be found at http://www.biodiversityinternational.org/ see Tables 1 and 2).

The most abundant source of non-viable historical landraces is found in herbarium collections. Such collections are widespread and can be found in many botanic gardens, university botany departments and natural history museums. These collections date from as early as the 17th century, but the most numerous material is from the 19th and early 20th centuries. The New York Botanic Garden and the International Association for Plant Taxonomy jointly run Index Herbariorum, a detailed directory of the public herbaria of the world. The searchable online Index Herbariorum database has been a major source of information about the herbarium collections accessed for this project (see http://sciweb.nybg.org/science2/IndexHerbariorum.asp).

One of the most important herbarium collections that has provided samples for this study is the Percival Collection of wheat accessions from over 40 countries. It was assembled by John Percival, who was the Professor of Agricultural Botany at the University of Reading from 1907 to 1932. The Percival Collection of germplasm has virtually disappeared, as no doubt have many of the landraces in those parts of the world where Percival collected.

Non-viable historical seed collections gathered together during agricultural exhibitions are also valuable sources of genetic information. For example, the Vänersborg Museum in Vänersborg, Sweden, holds a collection of seeds collected from farms throughout the country in about 1880, brought together as a result of a serious famine which occurred in Sweden *c*. 1870 (Ann-Charlott Öberg, pers. commun.).

Historical buildings can also form an important source of material. Cereal material has been used in strengthening daub and mudbricks, as insulation under floors or in the ceiling and as thatch for roofs. A source from historical buildings in several countries, particularly England, is SBT. Traditional methods of thatching have preserved mediaeval plant material in some late mediaeval former open hall buildings. Smoke was directly vented into the roof space, blackening the base coats of thatch. The SBT that has remained in these buildings is remarkably well preserved and contains the remains of both cereals and an interesting selection of crop weeds from cultivation that was preherbicide. The lowest layer of SBT is thought to be contemporaneous with the building of the house. Historical records or dendrochronological dating of the structural timbers can give a date for when this thatch was first laid down. Dendrochronology is a long-established and relatively accurate method in archaeological dating, by which the felling date of a given timber is ascertained by comparing seasonal variations in climate-induced growth as reflected in the varying width of a series of measured annual rings with other, previously dated reference ring sequences to allow precise dates to be ascribed to each ring (Bridge, 1988). Archaeobotanical analysis of the accompanying ruderal weed flora

Table 1. Significant genebank repositories of European barley (*Hordeum vulgare*) and emmer wheat (*Triticum turgidum* subsp. *dicoccum*) landraces as listed by the International Plant Genetic Resources Institute (IPGRI) Directory of Germplasm Collections. Emmer is also known in databases as *Triticum dicoccum*, *Triticum dicoccon* or *Triticum diccocom*. Institutions that do not hold European material are omitted. Where the information is available, holdings of European landraces are shown in brackets. Web addresses are shown where available

Institutions with viable germplasm collections	Number of *H. vulgare* landraces	Number of *T. dicoccum* landraces
Albania		
Plant Breeding/Seed Production Section, Department of Agronomy, Agricultural University, Tiranë; http://www.ubt.edu.al/aut.htm	25	
Austria		
Austrian Agency of Health and Foodsafety/Lwvie, Institute of Agroecology, Vienna; http://www.ages.at	28	
Belgium		
Center for Applied Biology, Linter-Neerhespen	9	
Cyprus		
National (CYPARI) Genebank, Agricultural Research Institute, Nicosia; http://www.ari.gov.cy	26	
Czech Republic		
Genebank Department, Research Institute of Crop Production (RICP), Prague; http://genbank.vurv.cz/genetic/resources/	25	76 (43)
Agricultural Research Institute Kromeriz, Co. Ltd, Kromeriz; http://www.vukrom.cz	47 (24)	
France		
Institut National de la Recherche Agronomique (INRA)[a], Clermont-Ferrand; http://www.inra.fr	3700	
Station de Genetique et Amelioration des Plantes[a], INRA C.R. Montpellier; http://www.ensam.inra.fr/	1400	
Georgia		
Protection Society of Agrobiodiversity, DIKA, Tbilisi; http://www.itic.org.ge/dika/eng.htm	25	5
Germany		
Institut fuer Pflanzengenetik und Kulturpflanzen-forschung (IPK), Gatersleben; http://www.ipk-gatersleben.de	3198 (607)	137 (67)
Federal Centre for Breeding Research on Cultivated Plants (BAZ), Braunschweig; http://www.bafz.de	3198 (607)	13 (67)
Greece		
Greek Genebank, Agricultural Research Center of Makedonia and Thraki, NAGREF, Thessaloniki; http://www.nagref.gr/	99	
Hungary		
Institute for Agrobotany, Tápiószele; http://www.rcat.hu	14	
Italy		
Istituto Sperimentale per la Cerealicoltura, Sezione di Fiorenzuola d'Arda[a], Fiorenzuola d'Arda (PC); http://www.cerealicoltura.it	600	
Dipartimento di Biologia Vegetale e Biotecnologie, Agroambientali e Zootecniche, Università di Perugia, Perugia; http://www.agr.unipg.it/dbvba/		5
Lithuania		
Lithuanian Institute of Agriculture, Kedainiai raj; http://www.lzi.lt/	1	
The Netherlands		
Centre for Genetic Resources (CGN), Wageningen; http://www.cgn.wur.nl	1724 (55)	
Norway		
Safety Base Collection of NGB, Svalbard	401 (271)	
Planteforsk Holt, Tromso; http://www.planteforsk.no	1	

Table 1. *Continued*

Institutions with viable germplasm collections	Number of *H. vulgare* landraces	Number of *T. dicoccum* landraces
Portugal		
Banco Português de Germoplasma Vegetal (BPGV), Braga	42	
Romania		
Research Institute for Cereals and Technical Plants, Fundulea, Calarasi; http://www.ricic.ro/informatii_en.htm	7	8 (1)
Agricultural Research Station, Suceava	2	5
Banca de Resurse Genetice Vegetale Suceava, Suceava; http://www.svgenebank.ro/sections_collecting_ro.htm	33	
University of Agricultural Sciences and Veterinary Medicine Timisoara, Timisoara; http://www.univagro-iasi.ro/agricultura/	8	
Russian Federation		
N. I. Vavilov All-Russian Scientific Research Institute of Plant Industry, St Petersburg; http://www.vir.nw.ru	4128 (1840)	463 (170)
Slovakia		
Slovakia Research Institute of Plant Production, Piestany	50 (28)	
Spain		
Centro de Recursos Fitogeneticos, INIA[a], Madrid; http://www.inia.es	1828	
Centre UdL, IRTA[a], Lleida	950	
Compania Espanola de Cultivos Oleaginosos S.A. (CECOSA), Madrid	20	
Grupo Cruzcampo S.A. Dirección de Servicios Agrícolas, Sevilla	166	
Consejo Superior de Investigaciones Cientificas, Zaragoza; http://www.dicar.csic.es/del/	99	
Sweden		
Nordic Gene Bank, Alnarp; http://www.ngb.se	411 (278)	
Switzerland		
Agroscope RAC Changins, Nyon; http://www.racchangins.ch	791	58
Schweizer Bergheimat, Lucerne; http://www.schweizer-bergheimat.ch/	36	5
Syrian Arab Republic		
International Center for Agricultural Research in the Dry Areas (ICARDA), Aleppo; http://www.icarda.cgiar.org	16,866 (3351)	
Ukraine		
Institute of Plant Production n.a. V. J. Yurjev of UAAS, Kharkiv	350 (253)	
United Kingdom		
John Innes Centre, Norwich; http://www.jic.bbsrc.ac.uk	4453	
Welsh Plant Breeding Station, Institute of Grassland and Environmental Research, Ceredigion; http://www.iger.bbsrc.ac.uk	1	
Scottish Agricultural Science Agency, Edinburgh; http://www.sasa.gov.uk	1	
United States of America		
Department of Botany and Plant Sciences, University of California, Riverside, CA; http://www.plantbiology.ucr.edu/		431 (102)
National Small Grains Germplasm Research Facility, USDA-ARS, Aberdeen, Idaho; http://www.ars-grin.gov/ars/PacWest/Aberdeen/	13,057 (2307)	

[a] Biodiversity International listing does not give details on whether accessions are wild, weedy, landraces, traditional varieties, elite cultivars or breeders lines.

https://doi.org/10.1017/S1479262108993138 Published online by Cambridge University Press

Table 2. European barley (*Hordeum vulgare*) and emmer wheat (*Triticum turgidum* subsp. *dicoccum*) landrace accessions available from major germplasm collections, showing the number of accessions supplied with coordinate data (latitude and longitude), detailed collection site or accession name data and with information sufficient to assign coordinates by reference to geographical databases

Germplasm collection	Species	Accessions	Coordinates available	Collection site or accession name available	Accessions where co-ordinates data found by cross referencing to databases
Agroscope RAC Changins (RAC)	Barley	795	0	782	507
Institut fuer Pflanzengenetik und	Barley	1369	0	1369	1117
Kulturpflanzenforschung (IPK)[a]	*Emmer wheat*	*195*	*0*	*195*	*185*
Institute for Agrobotany (RCAT)	Barley	14	0	14	14
National Small Grains Germplasm	Barley	240	240	240	–
Research Facility (NCSG)	*Emmer wheat*	*32*	*32*	*32*	–
Nordic Gene Bank (NGB)[b]	Barley	59	12	59	21
Institut National de la Recherche Agronomique (INRA)	Barley	237	0	237	119
John Innes Centre (JIC)[c]	Barley	17	0	17	13
Private collectors (PC)	*Emmer wheat*	*19*	*0*	*19*[d]	*19*[d]
Collecting expedition to Asturias (NIAB)	*Emmer wheat*	*40*	*40*	*40*	*40*
Total	Barley	2731	252	2718	1791
	Emmer wheat	*286*	*72*	*286*	*244*

[a] European Barley Core collection.
[b] Only accessions from Scandinavia were considered for this analysis.
[c] A limited set of accessions from this collection were considered in order to avoid duplication with the European Barley Core collection at IPK. (European Barley Core collection: The European Barley Database (EBDB) is an inventory of European barley collections. In 1997, it contained information about barley from 35 European genebanks as well as information about the 1126 accessions held in the International Barley Core collection, established in 1989 as part of an ECP/GR initiative. A core collection is a limited set of accessions derived from existing collections that are chosen to represent the genetic spectrum of the gene pool.)
[d] Information sufficient to locate samples to a region or country.

communities can provide contextual corroboration for the dating of the thatch (Letts, 1999). Generally, for northern Europe, dendrochronological dating is far more accurate than C14 dating, which has a large standard deviation at certain parts of the radiocarbon curve, and may be no more accurate than ± 500 years. This is not, therefore, useful if the material is less than 500 years old. Dendro dating can provide the precise year that a tree was cut. Even if wood is left to season, this may add 10–30 years to the date. Thatching straw was available wherever rivet wheat or rye was grown, while barley and oat straw were used when these preferred materials were unavailable. The straw used for thatching was mainly left over from threshing grain and so would represent the key cereal food crops grown in that locality. Ancient, tall-stemmed and genetically diverse cereal landraces useful for thatching began to disappear from cultivation by 1900 (Letts, 1999).

Sample location information

Information that is obtained by a phylogeographical study of European landraces depends on the availability of samples of relevance to an area of interest, but the quality of data obtained in a study of these accessions is constrained by their availability in a chosen geographical region allied to corresponding provenance data describing collection and subsequent storage. In addition, a key factor in whether useable samples of viable seeds are available is whether sampling expeditions were made on behalf of germplasm collections before landrace agriculture ceased. These germplasm accessions are of value if data recording the collection site and collection date are available from the germplasm collection. The provenance of all collected materials, viable or non-viable, depends on the collectors' notes made at the time of collection and on good record keeping at the holding institutions.

Phylogeographical studies of landraces require accurate spatial and temporal information about the accessions in question. Both non-viable historical and viable germplasm accessions are associated with varying qualities of geographical information: some germplasm accessions are provided with precise geographical coordinates, some, especially in the case of historical material, have descriptions of the sampling location described in relation to communes, villages, towns or topographical

https://doi.org/10.1017/S1479262108993138 Published online by Cambridge University Press

features such as rivers, waterfalls or mountains, and others provide rarely more than the name of the province or country of origin. In the case of viable material, for example, many traditional variety names refer to place names or regions: the German varieties Almersfelder (commercialised in 1891), Franken (commercialised in 1895) and Pfälzer (commercialised in 1909) originate from near the town of Almerfeld, from the region of Franconia, northern Bavaria, and from the Palatinate of western Germany, respectively. Early varieties listed as 'Berrichonne sélection dans une population de pays du Berry' and 'Märzengerste, Landgerste aus Ostfriesland' originate in the Berry region of central France and on the German Friesian Islands, respectively.

In the absence of accurate locational information, latitude and longitude may be assigned to accessions where the sampling location is named by reference to maps, gazetteers and online databases. Public domain databases are available to download from the US National Geospatial Intelligence Agency (http://earth-info.nga.mil/gns/html/cntry_files.html), or online locational databases such as the Getty Thesaurus of Geographic Names (http://www.getty.edu/research/conducting_research/vocabularies/tgn/) can be used. Table 2 lists the samples and location information available for the landraces used in this study. Figure 2 illustrates the availability of geographical data for extant barley landraces in European germplasm collections.

Various problems arise with finding location information for landrace accessions that were collected in earlier decades and as the collectors were not always nationals of the countries in which they were collecting, consequently there are errors in the recording of place names. In addition, national borders may have moved, notably in the rearrangements following World War II, and many local place names have changed to reflect a new political and/or cultural identity; a notorious example being the town infamous with the German name of Auschwitz, which now has the Polish name Oswiecim.

Distribution of viable landrace material

The maps in Fig. 3a, b show viable landrace survival, drawn from the significant landrace depositories listed in Table 1. Figure 3a, showing the availability of barley, indicates a substantial trend towards landrace survival in the regions of rugged upland topography, in particular the Aegean, the Apennine spine of Italy, the Alpine forelands and the flanks of the western Carpathians. France and the Iberian Peninsular have a moderate coverage, but elsewhere the survival of landraces is very patchy, notably in much of northern Europe, the United Kingdom and Ireland, and parts of eastern Europe. This may be due to a number of reasons. The lack of extant landraces in the eastern Carpathians probably reflects the lack of material

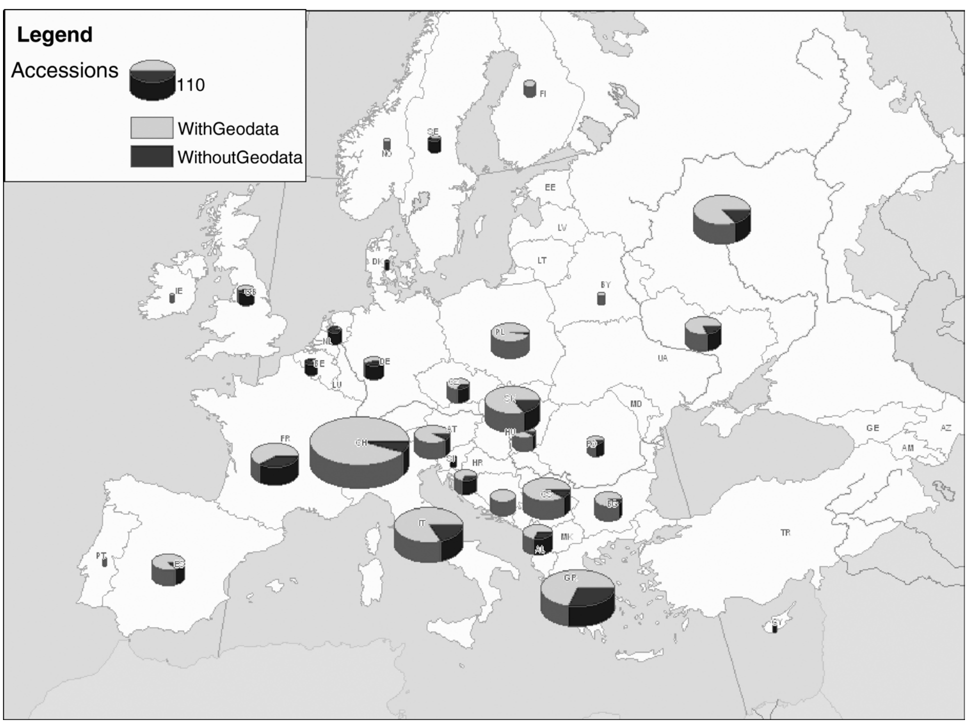

Fig. 2. European barley (*Hordeum vulgare)* accessions available from major germplasm collections. The number of accessions supplied with coordinate data (latitude and longitude) or with detailed collection site or accession name data is shown (see Table 2).

https://doi.org/10.1017/S1479262108993138 Published online by Cambridge University Press

Fig. 3. Collecting sites for (a) barley (*Hordeum vulgare*) and (b) emmer wheat (*Triticum turgidum* subsp. *dicoccum*) landrace accessions with coordinate data supplied by germplasm collections or where collection site and accession names have enabled coordinates to be assigned by cross-referencing to geographical databases. The germplasm collection supplying the accessions is indicated; the abbreviations for germplasm collection identity are explained in Table 1.

https://doi.org/10.1017/S1479262108993138 Published online by Cambridge University Press

stored in genebanks (see Fig. 2 and Table 1) and may not be due to an absence of landrace cultivation in these regions. Small-scale, subsistence-based, non-industrial farming communities have persisted in some of the more remote areas in these countries, and cereal landraces may still be cultivated *in situ*. However, the intensification and mechanization of eastern European agriculture over the past 50 years may also have led to a wider spread of 'improved' cereal varieties, even into remote areas. This could be confirmed only by extensive fieldwork.

Figure 3b shows the availability of emmer wheat, whose landrace survival partly echoes that of barley, but is considerably poorer: emmer is largely absent from northern Europe, the United Kingdom and Ireland, France and parts of central Europe. As with barley, the lack of extant landraces or the lack of fieldwork to discover persisting landraces in eastern Europe may reflect their lack of representation in genebank material.

In summary, the maps of viable barley and emmer landraces indicate their general restriction to remote areas such as mountainous regions or where growing conditions are suboptimal for modern elite varieties. This may be because farmers prefer to cultivate landraces in such environments as they are more successfully adapted to the local agroenvironmental conditions and have greater yield stability than modern cultivars (Harlan, 1992; Frankel *et al.*, 1998; Brown, 1999a). The landraces that survive probably do not represent the full genetic diversity of the wider populations now displaced by elite cultivars. Where they do survive, they may represent *in situ* conservation of historical genetic diversity (cf. Fischbeck, 2003).

Preserved landraces may indeed represent more ancient patterns of genetic diversity, or retain echoes of earlier patterns of distribution. If this is the case, the distribution of landraces shown in Fig. 3 should be appropriate to explore the spread of agriculture in prehistory because coverage is well placed to track the coastal route of barley expansion across Europe, and could quite feasibly generate detailed archaeogenetic patterns in its own right. It may be less suitable to track the inland spread of emmer across Europe, particularly north of the Alps and Carpathians. In this type of situation, particularly a crop such as emmer that has dwindled in importance, historical material may have the potential to fill the gaps.

Distribution of non-viable, historical landrace material

Table 3 and the maps shown in Fig. 4 indicate the sources and distribution of historical material to which access has been gained by the authors. It represents the results of making and following up requests for information across a significant part of Europe. While it is by no means exhaustive, it serves to illustrate some trends in regional accessibility.

It is immediately apparent that some regions are rich in both viable landraces and non-viable historical specimens. A clear example can be seen in the Alpine forelands (the historical landraces from this region were largely collected in the 1920s and 1930s; see Table 3). In this case, modern material is probably of greater value, but it will be interesting to compare what genetic changes have occurred in these barley landraces in the intervening 80 years.

It is also apparent that some regions, which although poorly represented in viable landraces, have retained non-viable historical landrace material and are therefore potential sources of genetic data. This is especially true of central and eastern Europe north of the Alps and Carpathians. Additionally, historical landrace repositories are particularly valuable for obtaining barley landrace accessions from Scandinavia, especially within the Arctic Circle.

There is a significant difference in the availability of the different species discussed in this paper: historical barley is considerably more accessible emmer wheat. The Percival Collection Herbarium has proved to be the major source of historical emmer accessions and, as stated above, rarely records more than the country of origin. Hence, geographical data for historical emmer wheat are poor, which may influence whether sufficient time depth can be added to the phylogeographical patterns seen in viable emmer landraces. One possibility for extending the genetic data derived from emmer is to include genetic data from other tetraploid wheats: rivet (*T. turgidum* subsp. *turgidum*) and durum (*T. turgidum* subsp. *durum*). These wheat species are not among the first founder crops, but they do share the AABB genome structure with emmer, and hence are closely related to each other. Rivet and durum wheat samples are often better represented in herbarium collections than emmer.

Conclusions

In this paper, we argue that phylogeographical analysis of viable and non-viable historical cereal landraces can allow us to determine the patterns of early agricultural spread. We consider the development of modern cultivars from landraces since the late 1800s, the sources of landrace germplasm still available today, and the distribution of these landraces where location information is available. We concluded that this data source, though rich, is geographically variable and lends itself to exploring particular aspects of early agricultural spread more than others. It should provide rich information for the spread of barley farming along the coastal routes of agricultural expansion, but much sparser information for the

Table 3. Sources of historical accessions of barley (*Hordeum vulgare*) and emmer wheat (*Triticum turgidum* subsp. *dicoccum*) that have been used in this study

Institution with historical barley or emmer wheat landrace accessions	Type of material (approximate dates)	No. of *H. vulgare* landraces	No. of *T. dicoccum* landraces
Percival Collection of Wheats, Natural History Museum, London, UK	Herbarium (1920s)	–	13
European Herbarium, Natural History Museum, London, UK	Herbarium (1830–1920)	11	–
National Herbarium of The Netherlands, Leiden, The Netherlands	Herbarium (1870–1940)	8	1
Landesdenkmalamt Baden – Württemberg, Esslingen am Neckar, Germany	Daub (1520–1700)	1	1
John Letts, Thatch expert, Oxford, UK	Thatch (mediaeval)	1	–
Centre for Economic Botany, Kew Gardens, London, UK	Herbarium (1920–1930)	2	–
Department of Plant Systematics and Geography, University of Warsaw, Warsaw, Poland	Herbarium (1880–1905)	5	–
Wroclaw Natural History Museum, Wroclaw University, Wroclaw, Poland	Herbarium (1900–1930)	2	–
LBBZ Plantahof, Landquart, Switzerland	Herbarium (1925–1935)	16	–
University Museum, University of Tromso, Tromso, Norway	Herbarium	2	–
	Old seed bank (1870–1890)	8	–
Adger Museum of Natural History, Kristiansand, Norway	Herbarium (1890–1950)	6	–
Herbarium, Institute of Botany Polish Academy of Sciences, Krakow, Poland	Herbarium (1870–1900)	6	–
Institute of Botany, University of Innsbruck, Innsbruck, Austria	Old seed bank (1920–1930)	55	–
Department of Plant and Environmental Sciences, Göteborg University, Göteborg, Sweden	Herbarium (1890–1935)	22	–
Botanical Museum, Lund University, Lund, Sweden	Herbarium (1880–1925)	24	–
Regional Herbarium, Oskarshamn, Sweden	Herbarium (1880–1930)	15	–
Vänersborg Museum, Vänersborg, Sweden	Old seed bank (1880)	29	
Botanical Institute, University of Copenhagen, Copenhagen, Denmark	Herbarium (1865–1935)	3	

https://doi.org/10.1017/S1479262108993138 Published online by Cambridge University Press

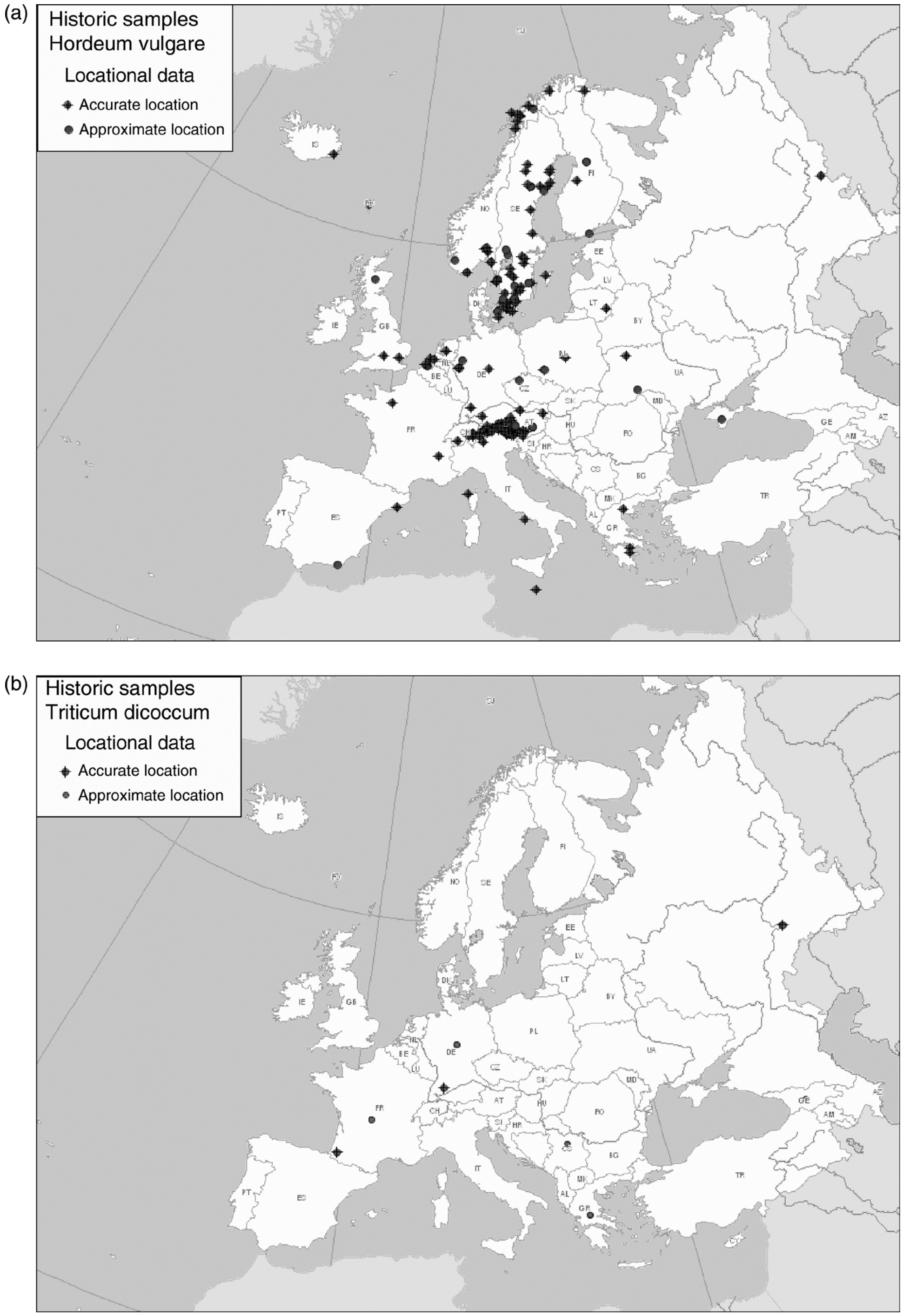

Fig. 4. The location of historical accessions of European (a) barley (*Hordeum vulgare*) and (b) emmer wheat (*T. turgidum* subsp. *dicoccum*) collected for this study. The sources of historical material are shown in Table 3. The map shows landrace accessions with more specific location information and accessions with only province or country of origin stated.

inland spread of founder crops, particularly those that have diminished in use in the historical period.

This survey of historical material suggests that some geographical gaps can be filled, particularly in the case of barley, whereas emmer wheat is more poorly represented in the historical landrace collections.

The inclusion of genetic data derived from non-viable historical landrace material can be usefully used alongside that from the modern landraces, not only to fill the geographical gaps, but also to allow time depth to be added to the phylogeographical data, allowing overstamping events that obscure older patterns of

https://doi.org/10.1017/S1479262108993138 Published online by Cambridge University Press

agricultural spread to be seen. Historical material is readily available, but often lacks good provenance and can sometimes contain substances that damage DNA or inhibit genetic analysis. Nonetheless, it remains an important source of genetic information that otherwise would not be available to help us understand the past patterns of cereal cultivation in Europe.

Acknowledgements

This work was funded by the Natural Environment Research Council as part of the consortium project 'The Domestication of Europe'. We thank the project partners at the University of Manchester, NIAB Cambridge and the University of Sheffield for their helpful advice, in particular Prof. T. A. Brown. We would like to thank the curators of all of the collections mentioned in this paper for their invaluable assistance in providing information and cereal material. The following individuals kindly assisted with the collection of historical material: Peter Stafford, Natural History Museum, London; Stephen Jury, Reading, MA; Hugh Cross, Leiden; Elske Fischer and Manfred Rösch, Baden–Württemberg; John Letts; Mark Nesbitt and Julia Steele, Kew Gardens; Krzysztof Spalik, Warsaw; Krzysztof Swierkosz, Wroclaw; Peer Schilperoord; Torbørn Alm; Tromsø; Per ArvidÅsen, Kristiansand; Aldona Bienniek, Krakow; Klaus Oeggl and Andreas Heiss, Innsbruck; Thorsten Elfstrom, Göteborg; Per Lassen, Lund;Åke Rühling, Oskarshamn; Ann-Charlott Öberg, Vänersborg; Gitte Petersen and Ole Seberg, Copenhagen.

References

Allaby RG, Banerjee M and Brown TA (1999) Evolution of the high molecular weight glutenin loci of the A, B, D, and G genomes of wheat. *Genome* 42: 296–307.

Alonso A, Martin P, Albarran C, Garcia P, Garcia O, de Simon LF, Garcia-Hirschfeld J, Sancho M, de La Rua C and Fernandez-Piqueras J (2004) Real-time PCR designs to estimate nuclear and mitochondrial DNA copy number in forensic and ancient DNA studies. *Forensic Science International* 139: 141–149.

Badr A, Müller K, Schäfer-Pregl R, El Rabey H, Effgen S, Ibrahim HH, Pozzi C, Rohde W and Salamini F (2000) On the origin and domestication history of barley (*Hordeum vulgare*). *Molecular Biology and Evolution* 17: 499–510.

Baumer M and Cais R (2000) Abstammung der Gerstensorten. Manuscript (unpublished) information available from www.lfl.bayern.de/ipz/gerste/09740/gerstenstamm.php

Baumer M and Göppel W (1998) Gerste, Sorten, Zuchter, Urspungland, Zulassungsjahr, Abstammung. Manuscript (unpublished) information available from www.lfl.bayern.de/ipz/gerste/09740/gerstenstamm.php

Bertin P, Gregoire D, Massart S and de Froidmont D (2001) Genetic diversity among European cultivated spelt revealed by microsatellites. *Theoretical and Applied Genetics* 1102: 148–156.

Bonjean AP and Angus WJ (2000) *The World Wheat Book: A History of Wheat Breeding*. Paris: Lavoisier Publishing.

Bridge MC (1988) The dendrochronological dating of buildings in southern England. *Medieval Archaeology* 32: 166–174.

Brown AHD (1999a) The genetic structure of crop landraces and the challenge to conserve them *in situ* on farms. In: Brush S (ed.) *Genes in the Field*. Rome: International Plant Genetic Resources Institute, pp. 29–48.

Brown TA (1999b) How ancient DNA may help in understanding the origin and spread of agriculture. *Philosophical Transactions of the Royal Society London B* 354: 89–98.

Brown TA (1999c) Genetic material. In: Carter D and Walker AK (eds) *Care and Conservation of Natural History Collections*. Oxford: Butterworth Heinemann, pp. 133–138.

Brown TA, Lindsay S and Allaby RG (2006) Using modern landraces of wheat to study the origins of European Agriculture. In: Motley TJ, Zeregra N and Cross H (eds) *Darwin's Harvest: New Approaches to the Origins, Evolution, and Conservation of Crops* (0231133162). New York: Columbia University Press, pp. 197–212.

Camacho Villa TC, Maxted N, Scholten MA and Ford-Lloyd BV (2005) Defining and identifying crop landraces. *Plant Genetic Resource: Characterization and Utilization*, 3(3): 373–384.

Cota-Sanchez JH, Remarchuk K and Ubayasena K (2006) Ready-to-use DNA extracted with a CTAB method adapted for herbarium specimens and mucilaginous plant tissue. *Plant Molecular Biology Reporter* 24: 161–167.

Fischbeck G (2003) Diversification through breeding. In: von Bothmer R, van Hintum T, Knüpffer H and Sato K (eds) *Diversity in Barley (*Hordeum vulgare*)*. London: Elsevier, pp. 29–52.

Frankel OH, Brown AHD and Burdon JJ (1998) *The Conservation of Plant Biodiversity*. 2nd edn. Cambridge: Cambridge University Press, pp. 56–78.

Hänni C, Brousseau T, Laudet V and Stehelin D (1995) Isopropanol precipitation removes PCR inhibitors from ancient bone extracts. *Nucleic Acids Research* 23: 881–882.

Harlan JR (1992) *Crops and Man*. 2nd edn. Madison, WI: American Society of Agronomy, pp. 147–148.

Heun M, Schäfer-Pregl R, Klawan D, Castagna R, Accerbi M, Borghi B and Salamini F (1997) Site of Einkorn wheat domestication identified by DNA fingerprinting. *Science* 278: 1312–1314.

IPGRI (1994) *Genebank Standards*. Rome: Food and Agriculture Organization of the United Nations/International Plant Genetic Resources Institute

Letts J (1999) *Smoke Blackened Thatch: A Unique Source of Late Mediaeval Plant Remains from Southern England*. London/ Reading, MA: English Heritage/University of Reading.

Li W, Brlansky R and Hartung JS (2006) Amplification of DNA of *Xanthomonas axonopodis* pv. *citri* from historic citrus canker herbarium specimens. *Journal of Microbiological Methods* 65: 237–246.

Lister DL, Bower MA, Howe CJ and Jones MK (2008) Extraction and amplification of nuclear DNA from herbarium specimens of emmer wheat: a method for assessing DNA preservation by maximum amplicon length recovery. *Taxon* 57: 254–258.

Morrison LA (2001) The Percival Herbarium and wheat taxonomy: yesterday, today and tomorrow. In: Caligari PDS and Brandham PE (eds) *Wheat Taxonomy: The Legacy of*

https://doi.org/10.1017/S1479262108993138 Published online by Cambridge University Press

John Percival. The Linnean Society of London Special Issue No. 3. London: Academic Press, pp. 65–80.

Reed FA, Kontanis EJ, Kennedy KA and Aquadro CF (2003) Ancient DNA prospects from Sri Lankan highland dry caves support an emerging global pattern. *American Journal of Physical Anthropology* 121: 112–116.

Zeven AC (1996) Results of activities to maintain landraces and other material in some European countries *in situ* before 1945 and what we may learn from them. *Genetic Resources and Crop Evolution* 43: 337–341.

Zhang L and Wu Q (2005) Single gene retrieval from thermally degraded DNA. *Journal of Bioscience* 30: 599–604.