# The spread of Neolithic plant economies from the Near East to northwest Europe: a phylogenetic analysis

Fiona Coward [a], Stephen Shennan [b,c,*], Sue Colledge [b], James Conolly [b,d], Mark Collard [c,e]

[a] Department of Geography, Royal Holloway University London, Egham, Surrey TW20 0EX, UK
[b] Institute of Archaeology, University College London, 31−34 Gordon Square, London WC1H 0PY, UK
[c] AHRC Centre for the Evolution of Cultural Diversity, University College London, 31−34 Gordon Square, London WC1H 0PY, UK
[d] Department of Anthropology, Trent University, 1600 West Bank Drive, Peterborough, ON, Canada K9J 7B8
[e] Department of Anthropology, University of British Columbia, 6303 NW Marine Drive, Vancouver, British Columbia, Canada V6T 1Z1

## Abstract

Phylogenetic techniques are used to analyse the spread of Neolithic plant economies from the Near East to northwest Europe as a branching process from a founding ancestor. The analyses are based on a database of c. 7500 records of plant taxa from 250 sites dated to the early Neolithic of the region in which they occur, aggregated into a number of regional groups. The analysis demonstrates that a phylogenetic signal exists in the data but it is complicated by the fact that in comparison with the changes that occurred when the crop agriculture complex expanded out of the Near East, once it arrived in Europe it underwent only limited further changes. On the basis of the analysis it has been possible to identify the species losses and gains that occurred as the complex of crops and associated weeds spread and to show the influence of geographical location and cultural affinity on the pattern of losses and gains. This has led to consideration of the processes producing that history, including some reasons why the dispersal process did not produce a perfect tree phylogeny, as well as to the identification of some specific anomalies, such as the unusual nature of the Bulgarian pattern, which raise further questions for the future.
© 2007 Elsevier Ltd. All rights reserved.

*Keywords:* Early Neolithic plant economies; Archaeobotany; Phylogenetic methods; Europe; Near East

## 1. Introduction

The nature and geographical patterning of the first Neolithic plant economies in the different regions of Europe and the Near East and the farming practices associated with them have become clearer in recent years (e.g. Bogaard, 2004; Willcox, 2002, 2005; Colledge et al., 2004, 2005; Weiss et al., 2006; papers in Colledge and Conolly, in press). In an earlier paper (Colledge et al., 2004), we suggested that the composition of archaeobotanical assemblages from the Levant, Turkey, Cyprus and Greece could be used to make inferences concerning the routes of spread of early crop agriculture practices and the factors that affected the practices themselves. In this paper we take this suggestion further by examining the distribution of the suites of plants at Early Neolithic sites relating to the initial local appearance of crop agriculture from southwest Asia to northwest Europe as a whole. We do so with the aid of methods developed by evolutionary biologists to reconstruct descent or ''phylogenetic'' relationships.

In what follows we will first discuss the rationale for the use of phylogenetic techniques to analyse the spread of Neolithic plant economies and specify the questions to be addressed. The data used will then be presented, followed by a description of the analytical methods and their results. The last section of the paper will discuss the implications of the results for what we can say about the processes and patterns involved in the spread of crop agriculture, and for the usefulness of phylogenetic techniques in addressing such issues.

* Corresponding author. AHRC Centre for the Evolution of Cultural Diversity, University College London, 31−34 Gordon Square, London WC1H 0PY, UK.

*E-mail address:* s.shennan@ucl.ac.uk (S. Shennan).

## 2. Phylogenetic methods and Neolithic plant economies

Most phylogenetic techniques reconstruct relationships among a set of taxa (species, genera, families etc) on the basis that the taxa are linked by a process of "descent with modification" involving successive branching. New taxa emerge through the splitting of existing ones and the new taxa are reproductively isolated from one another. When new character states appear in a particular taxon they will be inherited only by descendants of that taxon and no other. It is the pattern of distribution of shared derived character states or "synapomorphies" that provides the basis for reconstructing the tree of relationships among the focal taxa. Character states shared by all members of a focal group ("symplesiomorphies") are not considered informative regarding their phylogenetic relationships, nor are characters that are exhibited by only one member of the focal group ("autapomorphies").

In the case of the origin and spread into Europe of crop production systems based on domestic cereals and pulses, i.e. the "founder crops" of Neolithic agriculture, there are several reasons to think that a branching model of descent with modification is an appropriate way of conceptualising the relationships among the different regional packages of crops and associated weeds found at the earliest farming sites. Most obviously, the spread of domesticated crops involved the transmission of the crops themselves, genetic descendants of the ancestral species of the Near East. Second, ethnographic work on traditional agricultural systems indicates that farming practices are usually both relatively conservative and transmitted vertically between generations (e.g. Netting, 1993). It is thus highly probable that Early Neolithic farming practices were characterised by the same features (Bogaard, 2004).Third, while it is increasingly clear that a variety of cultivation systems based on different plants have existed in many parts of the world, including the Near East (see e.g. Smith, 2001; Harris, 2006; Weiss et al., 2006), those cereal and pulse founder crops that spread into Europe had essentially a single origin. At this scale the debate about the exact location in the northern or southern Levant of the domestication events that gave rise to the different crop species is immaterial, as the analyses described below make clear.

Finally, the spread of agricultural systems based on those crops, whether or not it involved demic as opposed to cultural diffusion, was a dispersal/expansion process. Expansion processes are inherently likely to produce branching patterns of change as successively modified sets of features spread from one place to another, although the branching process and the identification of its signal can be complicated by the existence of features that produce a conflicting signal. In the present case this could occur if an agricultural innovation that occurred in a given region spread back to the adjacent region from which crop cultivation had initially arrived as a result of continuing contacts between the regions (cf. Greenhill and Gray, 2005, p. 36–37 and figure 3.2).

There are thus strong a priori reasons for modelling the spread of agriculture as a phylogenetic branching process starting from a founding ancestor, and accordingly for using phylogenetic methods. This paper will therefore address a series of questions:

1. To what extent does analysis of the data support the claim that there should be a strong phylogenetic signal?
2. What species disappeared from or were added to the assemblages of crops and weeds as Neolithic plant economies spread? The losses and gains on each branch of the tree can be specified: what can this tell us about local adaptations or other processes producing variation?
3. How far does the branching pattern resulting from the analyses correspond to the geographical proximity of the different regions to the founding region? If the process is as claimed, the archaeobotanical assemblages from areas close to the source should be less "derived"—should have undergone less evolutionary change—than those from further away. Is this the case?
4. How closely do the patterns revealed by the analyses correspond to what we know from other archaeological sources about the routes along which crop agriculture spread (in particular the Mediterranean vs. the Balkans and Central Europe routes) and more generally to the established cultural patterns based on artefacts such as pottery?

## 3. Data

Analyses were based on a database of c. 7500 records of plant taxa recovered from 250 sites dated to the early Neolithic of the region in which they occur and therefore associated with the earliest local appearance of crop agriculture, from the Near East to northwest Europe, compiled by Colledge; each species found in a given early Neolithic site/phase represents a separate record in the database (Colledge et al., 2004, appendix B). The archaeobotanical information in the database was taken from published reports in which identified taxa were recorded, and comprise primarily plant remains that were preserved by charring, with very few mineralised or waterlogged specimens, or identifications based on impressions in pottery and daub. Our justification for the validity of using published data in this way is presented in the supplementary material. The remains were assigned a level of identification on a scale of 1 to 4 (e.g. 1: highest level of identification, i.e. to species; 4: lowest level of identification, i.e. to family or indeterminate category) according to the original classifications reported in the source publications. Only records of level 1 identifications (i.e. those identified to species) were included in this analysis. In addition, chaff (e.g. spikelet forks, glume bases and rachis fragments) and grains of the same cereals were amalgamated so that there was a single record for each species, as the analysis was designed to explore the presence of the domestic species rather than the individual elements of the ears of each crop *per se* (see Jones, 1991, p. 68). The frequency of occurrence of species in each site or phase was recorded in the original database, but simple presence (1) and absence (0) scores per site or phase were used in this analysis.

Wild progenitor species were excluded from the datasets because they were considered to be uninformative about the dispersal of domestic crops from their origins (i.e. wild einkorn, *Triticum boeoticum*; wild barley, *Hordeum spontaneum*; wild emmer, *Triticum dicoccoides*; wild rye, *Secale montanum/vavilovii*; wild faba bean, *Vicia narbonensis*; wild pea, *Pisum elatius*; wild grass pea, *Lathyrus cicera* and wild flax, *Linum bienne*). In line with a previous study of cultivation practices at early Neolithic sites in western Europe using the weed component of the archaeobotanical assemblages (Bogaard, 2002), trees and shrubs were also omitted from the analysis. All other wild species were included in the analysis. These included weeds (and potential weeds) that would have grown in cultivated fields with the crops.

Initially, individual early Neolithic site/phases were the taxa analysed, described in terms of the presence/absence of plant species. However, it rapidly became clear that chance sampling and preservation effects made these data too ''noisy'' for any patterns to emerge from a phylogenetic study. Accordingly, to reduce noise the sites were grouped into 22 geographical regions, which became the taxa analysed, each characterised by all the species present in any site/phase within that region. Species that were present in only one regional group, i.e. in less than 5% of the entire dataset, were omitted on the basis that exclusion of rare taxa also reduces sampling ''noise''. The site locations and region definitions are shown in the map (Fig. 1); the cultural definitions of the regions are listed in Table 1. The complete data matrix contained 115 freely-reversible binary characters recording the absence (0) or presence (1) of a particular species. Region 1, comprising the southern Levant earliest Neolithic sites (see Table 1) was

used as the outgroup for the phylogenetic analyses, the unit hypothetically ancestral to all the others and therefore the ''root'' of the evolutionary tree.

## 4. Methods (for a full description, see supplementary material)

The dataset was initially analysed with NeighborNet (Bryant and Moulton, 2004; Bryant et al., 2005), a program that constructs networks rather than trees and is therefore useful for assessing the degree to which a tree structure provides an adequate representation of a dataset. This was followed by the use of maximum parsimony analysis, using PAUP*4 (Swofford and Begle, 1993; Swofford, 1998), which attempts to find the optimal set of trees that will minimise the number of evolutionary ''steps'' required to explain the data. The fit of the data to the optimal trees was assessed by evaluation of the Consistency Index and the Retention Index, and by bootstrapping. Finally, the MrBayes program (Ronquist and Huelsenbeck, 2005) was used to carry out a Bayesian phylogenetic analysis. Bayesian analyses search for the set of most probable trees given the data and a particular model of evolution. They are therefore more realistic than optimising methods such as maximum parsimony.

## 5. Results

### 5.1. NeighbourNet

Fig. 2 shows the unrooted network resulting from the NeighborNet analysis. It is apparent from the box structures that there
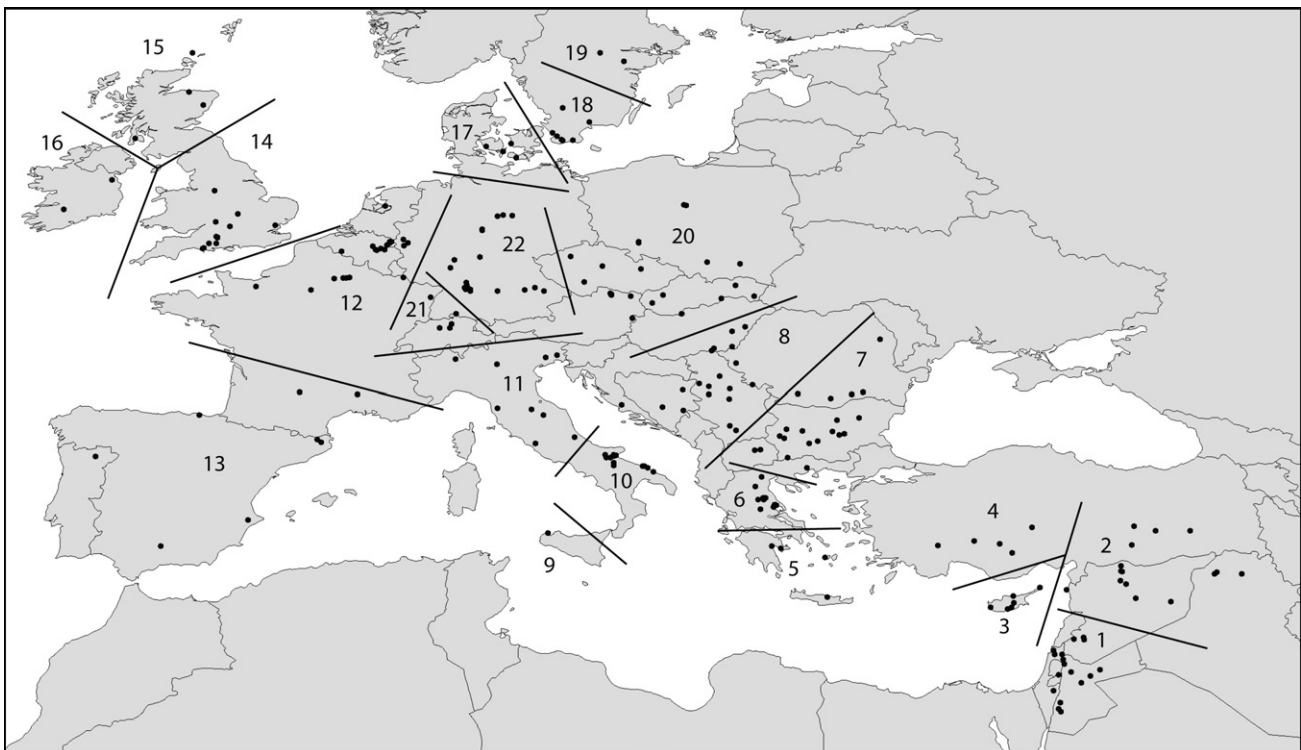


Fig. 1. Map of the distribution of the sample sites, showing the boundaries of the regions used in the analysis.

Table 1
Composition of the region groupings used in the dataset (see also Fig. 1)

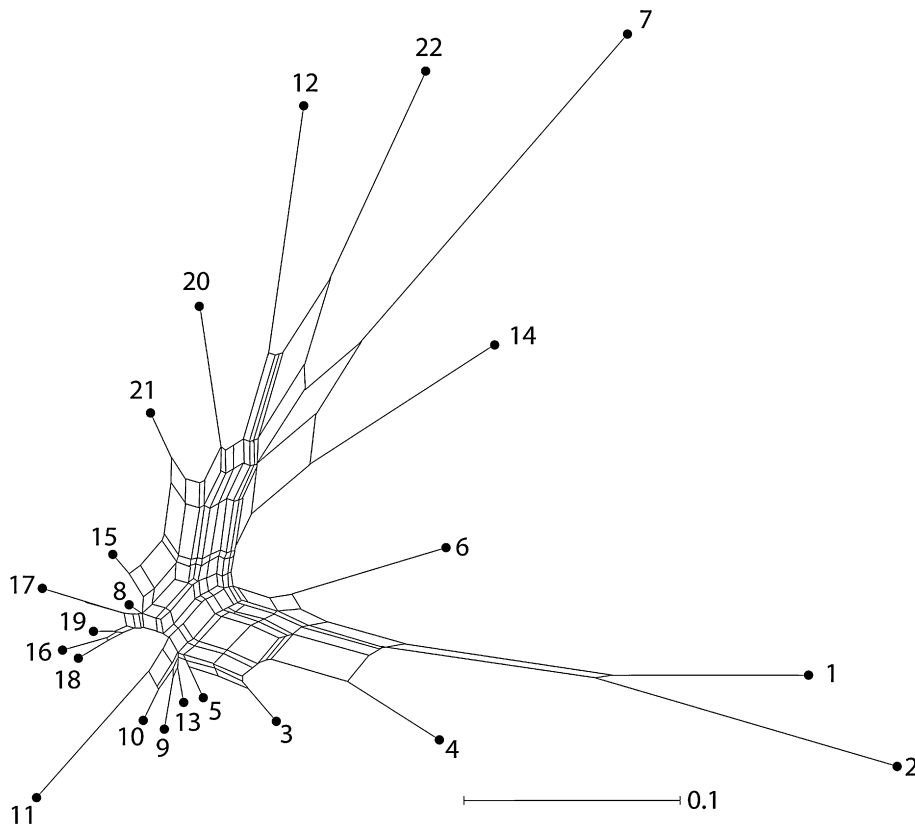| Regional Grouping | Area covered | Cultures | No. of sites |
|---|---|---|---|
| Region 1 | Jordan, Israel and the Syrian Damascus Basin | PPNA/B | 23 |
| Region 2 | Central Syria, the Euphrates Valley and southeastern Turkey | PPNA/B/C | 22 |
| Region 3 | Cyprus | Cypro-PPNB & Khirokitian | 10 |
| Region 4 | Central Turkey | PPNB & Aceramic Neolithic | 6 |
| Region 5 | Crete and southernmost Greece | Aceramic & Early Neolithic | 5 |
| Region 6 | Thessalian Greece | Aceramic, Early and Middle Neolithic (Pre- and Proto-Sesklo and Sesklo) | 15 |
| Region 7 | Bulgaria and Macedonia | Karanovo and Starčevo | 15 |
| Region 8 | Hungary and the former Yugoslavia | Körös and Starčevo | 8 |
| Region 9 | Sicily | Impressed ware/Stentinello | 1 |
| Region 10 | Southern Italy | Impressed ware, Serra d'Alto, Diana, Lagnano da Piede, Masseria la Quercia, Guadone, Impressed ware, red-band painted pottery, Passo di Corvo | 13 |
| Region 11 | Central and Northern Italy | Square-mouthed pottery, Fagnigola, Impressed ware, Catignano | 6 |
| Region 12 | The Netherlands, Belgium, northern France and western Germany (Nordrhein-Westfalen) | LBK, Swifterbant, Rubane, Group de Blicquy, Cerny | 27 |
| Region 13 | Portugal, Spain and southern France | Impressed ware, Cardial/Postcardial/Epicardial, Chassey | 11 |
| Region 14 | Southern and central Britain | Early Neolithic | 19 |
| Region 15 | Northern Britain and Scotland | Early Neolithic | 4 |
| Region 16 | Ireland | Western Neolithic/Decorated Pottery Complex | 3 |
| Region 17 | Denmark | TRB | 3 |
| Region 18 | Southern Sweden | TRB | 5 |
| Region 19 | Northern Sweden | TRB | 2 |
| Region 20 | Austria, Poland, Slovakia and the Czech Republic | LBK, eastern LBK | 22 |
| Region 21 | Western Germany | LBK | 2 |
| Region 22 | Central Germany | LBK | 19 |



Fig. 2. Results of a NeighbourNet analysis of the regional groupings dataset.

is extensive evidence of signals conflicting with a pure tree hypothesis. However, tree structure is also apparent and it is clear that the patterns in the network structure correspond to our initial evolutionary hypothesis. Regions 1 and 2, the two competing southern and northern Levant founder regions, are at one end of the network. Linked to them in turn are Region 4 (Central Anatolia), followed by Cyprus (Region 3) and Thessaly (Region 6). Going round the diagram clockwise from Region 3 it is apparent that Regions 5 to 15 (Crete, Iberia, Sicily, southern Italy, northern Italy, southern Sweden, Ireland, northern Sweden, Denmark, former Yugoslavia and Scotland) are not individually very distinctive, though they are markedly different from the ancestral areas. The remaining Regions, 21 round to 14 (western Germany, Central Europe, the Netherlands and adjacent areas, Central Germany, Bulgaria and southern Britain), are distinctly different from the rest and further removed in terms of the composition of their archaeobotanical assemblages from the founder regions.

## 5.2. Parsimony analysis

An initial branch-and bound search found 36 optimal trees of length 248. Characters were then re-weighted according to their individual rescaled consistency indices, effectively penalising those characters demonstrating more homoplasy and preferentially weighting the search towards those fitting the tree(s) well, and thus providing a better phylogenetic signal (Swofford and Begle, 1993, p. 61). This process was repeated, after which further reweighting did not produce a reduction of the three optimal trees of length 38.90278 found (one of the three is shown in Fig. 3 as an exemplar; differences between them were minimal and related to subtle differences in branch length). For a sample of trees derived from randomly permuted data the mean length was 66.03, standard deviation 2.74, indicating that the optimal tree found was extremely unlikely to have been produced by chance alone and hence that the real dataset demonstrated a highly significant phylogenetic signal.

For the initial "raw" trees produced from the data the Consistency Index (CI) was 0.439; the Retention Index (RI) was 0.457. The range of RI values presented by Collard et al. (2006) for 21 different cultural datasets produced a mean of 0.59 and a range from 0.42 to 0.78. It is thus apparent that the figure for the archaeobotanical dataset is at the bottom end of this range and does not indicate the strong phylogenetic signal expected. It will be suggested below that the reason for this is the remarkable uniformity or lack of evolutionary change in the majority of the European datasets, which means that different regions have multiple cross-cutting connections with one another.

The phylograms produced by the maximum parsimony and bootstrap analyses (Figs. 3 and 4) demonstrate considerable structure. This structure is underpinned by a significant number of different species (Table 2). The first and most obvious division within the phylogram in Fig. 3 is between regions 1 and 2 (southern Levant and northern Levant respectively) on the one hand, and regions 3–22 on the other. This division occurred in 100% of the bootstrap phylograms. At least 15
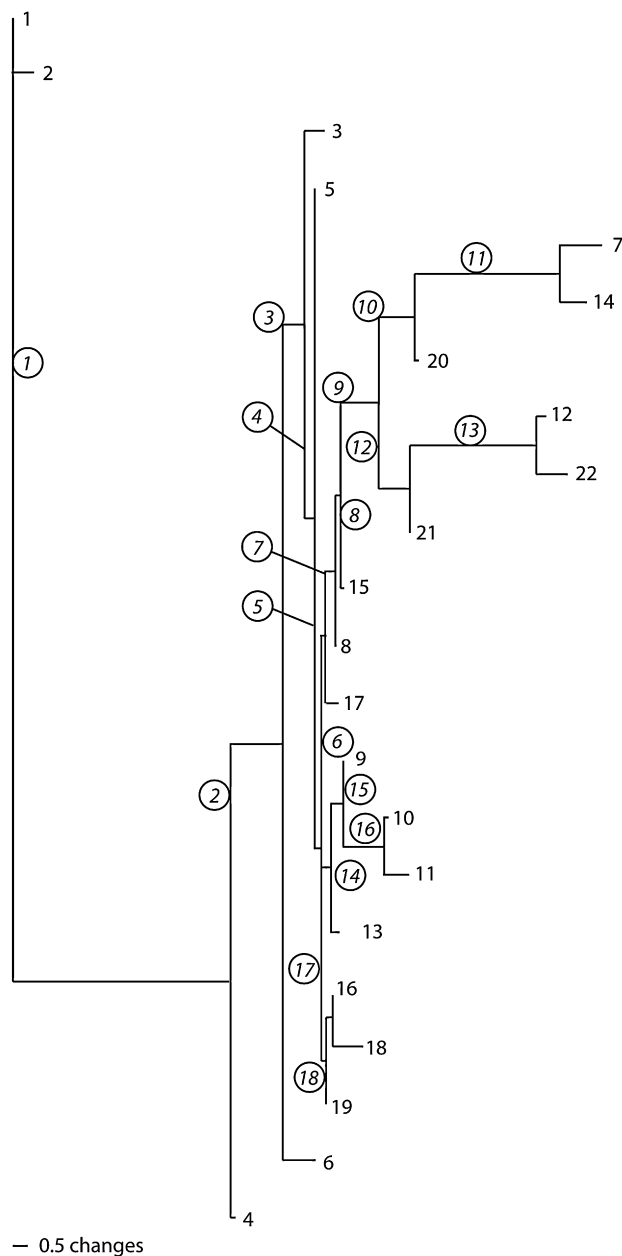


Fig. 3. Optimal phylogram derived from the regional groupings dataset. Tree length: 38.90278; CI: 0.7496; RI: 0.8166; RC: 0.6121. Numbers attached to internal branches refer to the apomorphies listed in Table 2.

species (including the domestic species *Vicia faba*, faba bean, and *Linum usitatissimum*, flax) are lost on branch 1, which divides regions 1 and 2 from the rest. The extent of the differences between the two Levantine regions and the other regions, and the fact that the vast majority of the losses comprise species that were only recorded on sites in the Near East, obviously contribute to the high bootstrap support value (Fig. 4). Making the northern Levant group (region 2) the outgroup rather than the southern Levant (region 1) made no difference to these results.

Bootstrap support for a branch consisting of region 4 (the central Turkish sites), dividing it from Thessaly (region 6) and the rest is also high, with a value of 77%; the remaining
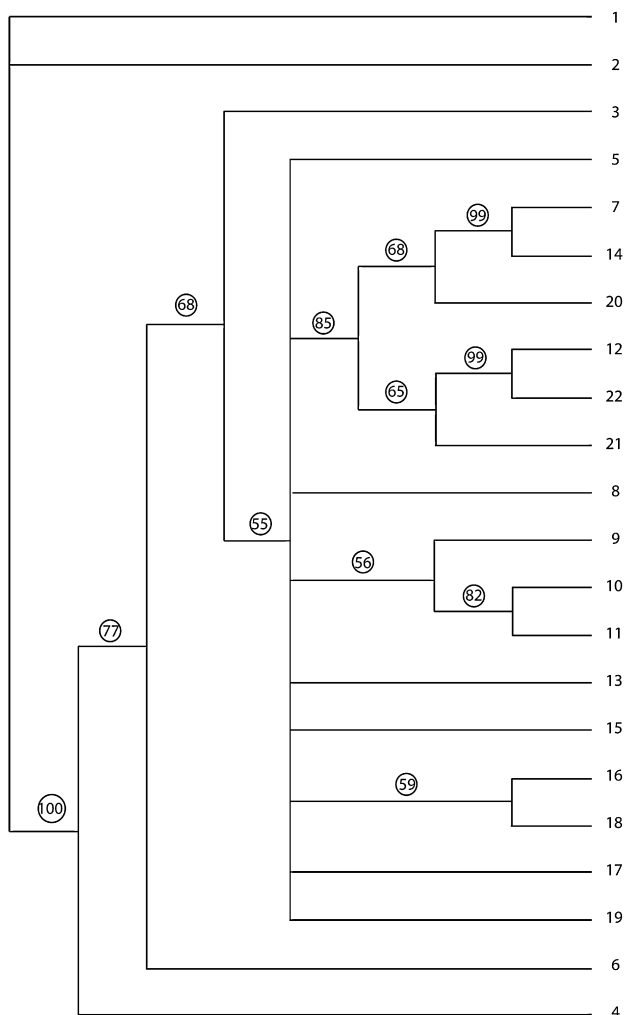
Fig. 4. Bootstrap consensus tree for the regional groupings dataset.

groups are linked by the gain of the domestic species *Lathyrus sativus* (grass pea) along branch 2 and also by the shared loss of three wild species. The loss of *Cicer arietinum* (chick pea) and a single wild species along branch 3 divides Thessaly (region 6) from Cyprus and all the other regions, with a bootstrap support value of 68%. Cyprus itself (region 3) is distinct from the remaining regions. All the other European regions are linked by the loss of the three crops *Lathyrus sativus*, *Pisum sativum* (pea) and *Vicia ervilia* (bitter vetch) along branch 4; the clade has a bootstrap support value of 55%.

Perhaps the most striking feature revealed by the rest of the most parsimonious phylogram (Fig. 3), with some exceptions to be discussed below, is how remarkably similar most of the European regions are. This is confirmed by the lack of structure in the bootstrap consensus tree (Fig. 4), beyond the branch separating off Cyprus (region 3). Within the general European clade three groups have some support from the bootstrap analysis. The first of these (at the bottom of Figs. 3 and 4) links regions 16 (Ireland) and 18 (southern Sweden) through the shared gain of a single wild species, *Galium aparine*, which links the Irish and southern Swedish sites; it is found in

53% of bootstrap replicates. However, it is possible that this link may be an artefact of the data (see below).

The second group is made up of the Italian regions (10 and 11) and Sicily (region 9), with the addition of two species at branch 15 (*Vicia ervilia* and a wild species), although the bootstrap support value is only 58%. A clade comprising regions 10 and 11 (the southern and central/northern Italian sites, respectively) is more strongly supported by three shared gains, which include the domestic species *Avena sativa* (oats) and two wild species, and was recovered in 82% of bootstrap replicates.

Five additions (including *Pisum sativum* and four wild species along branch 9, Fig. 3) underlie the clade comprising the remaining regions: 7, 12, 14, 20, 21 and 22, which was recovered in 85% of bootstrap replicates. Two distinct sub-groups are apparent in this clade; one of these (with a bootstrap value of 65%) includes regions 12 (the north-west European mainland), 21 (western Germany) and 22 (central Germany), which are linked by three shared gains (all wild species) along branch 12 of the phylogram. Within this sub-group regions 12 and 22 are more closely related than either are to region 21 and have a total of eight species in common (added at branch 13) including *Papaver somniferum* (opium poppy), which contribute to a bootstrap value of 99%.

The other sub-group in this clade comprises regions 7 (the eastern Balkan sites), 14 (southern and central Britain) and 20 (the eastern LBK sites), which are linked by five shared gains and one loss (all wild species) on branch 10 of the phylogram; this sub-group was recovered in 68% of bootstrap replicates. Within this group there is an unexpected sub-clade, which includes the eastern Balkan sites (region 7) and the southern and central British sites (region 14), linked along branch 11 by the addition of eight species (all wild species, four of which are unique to these two regions: *Aphanes arvensis*, *Atropa bella-donna*, *Chenopodium murinum* and *Poa annua*) and the loss of one species. This unusual sub-clade was strongly supported, being found in 99% of bootstrap replicates; however, it does not occur in the Bayesian consensus phylogram (see below).

## 5.3. Bayesian analysis

The overall structure of the Bayesian 50% majority-rule consensus phylogram shown in Fig. 5 is consistent with that of the maximum parsimony phylogram, but there are some significant differences between the two. It should also be noted that because the Bayesian phylogram presented in Fig. 5 is a 50% majority consensus tree not all trees sampled will demonstrate identical gains and losses of particular species at particular nodes, and thus only general patterns of variation in species presence and absence are discussed here.

A deep split between regions 1 and 2, on the one hand, and regions 3–22 on the other is confirmed in 100% of sampled trees. Fifteen losses, including the two crops *Linum usitatissimum* and *Vicia faba* distinguish the founder region from the least derived of the remaining clades, region 4 (central Turkey). Five further losses, including *Cicer arietinum*, and the gain of

Table 2
Character changes on the branches of the optimal parsimony phylogram (Fig. 3)

| Branch number | Gain (+) or loss (−) | Species |
| --- | --- | --- |
| 1 | − | *Aegilops speltoides* |
| | − | *Aizoon hispanicum* |
| | − | *Androsace maxima* |
| | − | *Arnebia linearifolia* |
| | − | *Cephalaria syriaca* |
| | − | *Helianthemum salicifolium* |
| | − | *Hordeum bulbosum* |
| | − | *Hordeum murinum* |
| | − | **Linum usitatissimum** |
| | − | *Lolium temulentum* |
| | − | *Medicago minima* |
| | − | *Medicago radiata* |
| | − | *Poa bulbosa* |
| | − | *Rumex pulcher* |
| | − | **Vicia faba** |
| 2 | − | *Arnebia decumbens* |
| | − | *Bromus sterilis* |
| | − | *Buglossoides tenuiflora* |
| | + | **Lathyrus sativus** |
| 3 | − | **Cicer arietinum** |
| | − | *Verbena officinalis* |
| 4 | − | **Lathyrus sativus** |
| | − | **Pisum sativum** |
| | − | **Vicia ervilia** |
| 5 | − | *Buglossoides arvensis* |
| 6 | + | *Chenopodium album* |
| 7 | + | *Polygonum convolvulus* |
| 8 | + | **Linum usitatissimum** |
| | + | *Polygonum persicaria* |
| 9 | + | *Echinochloa crus-galli* |
| | + | *Galium aparine* |
| | + | *Galium spurium* |
| | + | **Pisum sativum** |
| | + | *Polygonum lapathifolium* |
| 10 | + | *Asperula arvensis* |
| | + | *Chenopodium hybridum* |
| | + | *Polygonum aviculare* |
| | + | *Polygonum minus* |
| | − | *Polygonum persicaria* |
| | + | *Rumex crispus* |
| 11 | + | *Aphanes arvensis* |
| | + | *Atropa bella-donna* |
| | + | *Chenopodium murale* |
| | + | *Plantago lanceolata* |
| | + | *Poa annua* |
| | − | *Polygonum lapathifolium* |
| | + | *Rubus fruticosus* |
| | + | *Rumex acetosella* |
| | + | *Veronica hederifolia* |
| 12 | + | *Chenopodium ficifolium* |
| | + | *Lapsana communis* |
| | + | *Polygonum amphibium* |
| 13 | + | *Bromus secalinus* |
| | + | *Bromus sterilis* |
| | + | *Bromus tectorum* |
| | + | *Chenopodium polyspermum* |
| | + | *Galium palustre* |
| | + | **Papaver somniferum** |
| | + | *Veronica arvensis* |
| | + | *Vicia hirsuta* |
| 14 | + | **Vicia faba** |

Table 2 (*continued*)

| Branch number | Gain (+) or loss (−) | Species |
| --- | --- | --- |
| 15 | − | *Euphorbia helioscopia* |
| | + | **Vicia ervilia** |
| 16 | + | **Avena sativa** |
| | + | *Buglossoides arvensis* |
| | + | *Chenopodium album* |
| 17 | − | **Lens culinaris** |
| 18 | + | *Galium aparine* |

| Autapomorphies | Presence (+) or absence (−) | Species |
| --- | --- | --- |
| Region 1 | + | *Adonis dentata* |
| | + | *Carex divisa* |
| | + | *Coriandrum sativum* |
| | + | *Galium mollugo* |
| | + | *Galium tricornutum* |
| | + | *Rubus sanguineus* |
| | + | *Vaccaria pyramidata* |
| Region 2 | + | *Adonis annua* |
| | + | *Adonis flammea* |
| | + | *Arenaria serpyllifolia* |
| | + | *Bromus arvensis* |
| | + | *Chenopodium album* |
| | + | *Convolvulus arvensis* |
| | + | **Lathyrus sativus** |
| | + | *Papaver argemone* |
| | + | *Plantago lanceolata* |
| | + | *Polygonum convolvulus* |
| | + | *Polygonum corrigioloides* |
| | + | *Polygonum lapathifolium* |
| | + | *Polygonum persicaria* |
| | + | *Portulaca oleracea* |
| | + | *Rumex acetosella* |
| | + | *Urtica urens* |
| Region 3 | + | *Adonis dentata* |
| | + | *Buglossoides tenuiflora* |
| | + | *Hordeum murinum* |
| Region 5 | + | *Rumex sanguineus* |
| Region 7 | + | *Adonis flammea* |
| | + | *Agrimonia eupatoria* |
| | + | *Ajuga chamaepitys* |
| | + | *Anagallis arvensis* |
| | + | *Bromus arvensis* |
| | + | *Buglossoides arvensis* |
| | + | *Chenopodium ficifolium* |
| | + | *Chenopodium polyspermum* |
| | + | **Cicer arietinum** |
| | + | *Convolvulus arvensis* |
| | + | *Fragaria vesca* |
| | + | *Galium mollugo* |
| | + | *Hibiscus trionum* |
| | + | *Hyoscyamus niger* |
| | + | **Lathyrus sativus** |
| | + | *Polygonum persicaria* |
| | + | *Portulaca oleracea* |
| | + | *Verbena officinalis* |
| | + | **Vicia ervilia** |
| | + | *Vicia tetrasperma* |
| Region 14 | + | *Arrhenatherum elatius* |
| | − | *Asperula arvensis* |
| | − | *Chenopodium hybridum* |
| | − | *Echinochloa crus-galli* |
| | − | *Galium spurium* |
| | − | **Lens culinaris** |

Table 2 (*continued*)

| Autapomorphies | Presence (+) or absence (−) | Species |
| --- | --- | --- |
| | − | **Pisum sativum** |
| | + | Plantago major |
| | + | Poa trivialis |
| | + | Polygonum arenastrum |
| | + | Prunella vulgaris |
| | − | Rumex crispus |
| | + | Urtica urens |
| Region 20 | + | Galium tricornutum |
| | + | **Panicum miliaceum** |
| Region 12 | − | Anagallis arvensis |
| | − | Arrhenatherum elatius |
| | + | Chenopodium ficifolium |
| | − | Papaver argemone |
| | − | Phleum pratense |
| | − | Poa trivialis |
| | + | Polygonum amphibium |
| | − | Polygonum arenastrum |
| | − | Prunella vulgaris |
| | − | Rubus fruticosus |
| | − | Veronica hederifolia |
| Region 22 | + | Agrimonia eupatoria |
| | + | Asperula arvensis |
| | + | Bromus arvensis |
| | + | Chenopodium hybridum |
| | + | Euphorbia helioscopia |
| | + | Hyoscyamus niger |
| | + | **Panicum miliaceum** |
| | + | Plantago lanceolata |
| | + | Plantago major |
| | + | Polygonum aviculare |
| | − | Polygonum persicaria |
| | + | Rumex acetosella |
| | + | **Vicia ervilia** |
| | + | **Vicia faba** |
| | + | Vicia tetrasperma |
| Region 21 | − | **Lens culinaris** |
| Region 15 | − | **Lens culinaris** |
| | − | **Triticum monococcum** |
| Region 17 | + | Arenaria serpyllifolia |
| | + | Bromus secalinus |
| | − | **Lens culinaris** |
| | + | Phleum pratense |
| | + | Polygonum aviculare |
| Region 9 | − | **Hordeum vulgare var. nudum** |
| Region 10 | − | **Vicia ervilia** |
| Region 11 | + | Agrostemma githago |
| | + | Fragaria vesca |
| | + | **Linum usitatissimum** |
| | + | **Papaver somniferum** |
| | + | Polygonum convolvulus |
| | + | Polygonum persicaria |
| | + | Portulaca oleracea |
| | + | Rubus fruticosus |
| Region 13 | + | **Papaver somniferum** |
| | + | **Pisum sativum** |
| Region 16 | − | **Hordeum vulgare var. nudum** |
| | − | **Triticum monococcum** |
| Region 18 | − | **Hordeum vulgare** |
| | + | Polygonum persicaria |
| Region 19 | + | **Panicum miliaceum** |

Table 2 (*continued*)

| Autapomorphies | Presence (+) or absence (−) | Species |
| --- | --- | --- |
| Region 6 | + | Agrostemma githago |
| | + | Ajuga chamaepitys |
| | + | Bromus secalinus |
| | + | Coriandrum sativum |
| | + | Galium aparine |
| | + | Galium spurium |
| | + | **Linum usitatissimum** |
| | + | Lolium temulentum |
| | + | **Panicum miliaceum** |
| | + | Portulaca oleracea |
| Region 4 | + | Adonis annua |
| | + | Adonis flammea |
| | + | Carex divisa |
| | + | Chenopodium album |
| | + | Hibiscus trionum |
| | + | Polygonum corrigioloides |
| | + | Vaccaria pyramidata |

Domestic crops are in bold.

*Lathyrus sativus* separate the remaining regions from central Turkey, another strongly supported division, found in 84% of sampled trees. The next least-derived regions are 3 and 6 (Cyprus and Thessaly); the rest of Europe is distinguished from these by the further loss of four species, including *Lathyrus sativus*, *Pisum sativum* and *Vicia ervilia*, a separation supported by a posterior probability of 80%.

In comparison with the preceding changes most of the European regions are relatively underived; that is to say, once crop agriculture arrived in Europe there were relatively few further species losses and gains. However, some clades within the large European group are supported. A group made up of regions 16 and 18 (the Irish and southern Swedish sites) occurs in 60% of the Bayesian samples. The link between the two regions in this analysis is the loss of *Lens culinaris* (lentil) (on branch 9) and the shared gain of a wild species, *Galium aparine* (the same species that was common to both in the maximum-parsimony analysis but which may be a data artefact; see below).

Regions 9, 10 and 11 (Sicily, southern Italy and central and northern Italy, respectively) form a clade by virtue of the gain of *Vicia ervilia*, *Vicia faba* and a single wild species (at branch 7), with a posterior probability of 55%. The two mainland Italian regions form a sub-clade within this grouping, as was the case in the maximum parsimony analysis, being linked by the gain of *Avena sativa* and two wild species (at branch 8), with a very high posterior probability of 91%.

Seven further regions are clearly distinguishable as a single clade in the Bayesian phylogram (with a posterior probability of 56%); they are linked by the gain of six species (including *Linum usitatissimum* and five wild species), and by the loss of *Lens culinaris* (at branch 4). Within this clade, region 15 (northern Britain and Scotland) remains relatively un-derived and is distinguished by only three autapomorphies (all losses, including that of *Triticum monococcum* (einkorn)), while region 14 (southern and central Britain) gains as many as 14 additional wild species (as well as losing one species) to make it one of the more derived regions in the dataset. Distinct from these is

Table 3
Character changes on the branches of the majority-rule consensus phylogram produced by Bayesian analysis of the dataset (Fig. 5)

| Branch No. | Gain (+) or loss (−) | Species |
|---|---|---|
| 1 | − | *Aegilops speltoides* |
| | − | *Aizoon hispanicum* |
| | − | *Androsace maxima* |
| | − | *Arnebia linearifolia* |
| | − | *Cephalaria syriaca* |
| | − | *Helianthemum salicifolium* |
| | − | *Hordeum bulbosum* |
| | − | *Hordeum murinum* |
| | − | ***Linum usitatissimum*** |
| | − | *Lolium temulentum* |
| | − | *Medicago minima* |
| | − | *Medicago radiata* |
| | − | *Poa bulbosa* |
| | − | *Rumex pulcher* |
| | − | ***Vicia faba*** |
| 2 | − | *Arnebia decumbens* |
| | − | *Bromus sterilis* |
| | − | *Buglossoides tenuiflora* |
| | − | ***Cicer arietinum*** |
| | + | ***Lathyrus sativus*** |
| | − | *Verbena officinalis* |
| 3 | − | *Buglossoides arvensis* |
| | − | ***Lathyrus sativum*** |
| | − | ***Pisum sativum*** |
| | − | *Vicia ervilia* |
| 4 | + | *Chenopodium album* |
| | + | *Galium aparine* |
| | − | ***Lens culinaris*** |
| | + | ***Linum usitatissimum*** |
| | + | *Polygonum aviculare* |
| | + | *Polygonum convolvulus* |
| | + | *Polygonum persicaria* |
| 5 | + | *Asperula arvensis* |
| | + | *Chenopodium ficifolium* |
| | + | *Chenopodium hybridum* |
| | + | *Echinochloa crus-galli* |
| | + | *Galium spurium* |
| | + | ***Lens culinaris*** |
| | + | ***Pisum sativum*** |
| | + | *Polygonum lapathifolium* |
| 6 | + | *Bromus secalinus* |
| | + | *Bromus sterilis* |
| | + | *Bromus tectorum* |
| | + | *Chenopodium polyspermum* |
| | + | *Galium palustre* |
| | + | *Lapsana communis* |
| | + | ***Papaver somniferum*** |
| | + | *Veronica arvensis* |
| | + | *Vicia hirsuta* |
| 7 | + | *Euphorbia helioscopia* |
| | + | ***Vicia ervilia*** |
| | + | ***Vicia faba*** |
| 8 | + | ***Avena sativa*** |
| | + | *Buglossoides arvensis* |
| | + | *Chenopodium album* |
| 9 | + | *Galium aparine* |
| | − | ***Lens culinaris*** |

Table 3 (continued)

| Autapomorphies | Presence (+) or absence (−) | Species |
|---|---|---|
| Region 1 | + | *Adonis dentata* |
| | + | *Carex divisa* |
| | + | *Coriandrum sativum* |
| | + | *Galium mollugo* |
| | + | *Galium tricornutum* |
| | + | *Rubus sanguineus* |
| | + | *Vaccaria pyramidata* |
| Region 2 | + | *Adonis annua* |
| | + | *Adonis flammea* |
| | + | *Arenaria serpyllifolia* |
| | + | *Bromus arvensis* |
| | + | *Chenopodium album* |
| | + | *Convolvulus arvensis* |
| | + | ***Lathyrus sativus*** |
| | + | *Papaver argemone* |
| | + | *Plantago lanceolata* |
| | + | *Polygonum convolvulus* |
| | + | *Polygonum lapathifolium* |
| | + | *Polygonum persicaria* |
| | + | *Portulaca oleracea* |
| | + | *Rumex acetosella* |
| | + | *Urtica urens* |
| Region 3 | + | *Adonis dentata* |
| | + | *Bromus sterilis* |
| | + | *Buglossoides tenuiflora* |
| | + | *Hordeum murinum* |
| Region 5 | + | *Buglossoides arvensis* |
| | + | *Rumex sanguineus* |
| Region 7 | + | *Adonis flammea* |
| | + | *Agrimonia eupatoria* |
| | + | *Ajuga chamaepitys* |
| | + | *Anagallis arvensis* |
| | + | *Aphanes arvensis* |
| | + | *Atropa bella-donna* |
| | + | *Bromus arvensis* |
| | + | *Chenopodium murale* |
| | + | *Chenopodium polyspermum* |
| | + | ***Cicer arietinum*** |
| | + | *Convolvulus arvensis* |
| | + | *Fragaria vesca* |
| | + | *Galium mollugo* |
| | + | *Hibiscus trionum* |
| | + | *Hyoscyamus niger* |
| | + | ***Lathyrus sativus*** |
| | + | *Plantago lanceolata* |
| | + | *Poa annua* |
| | − | *Polygonum lapathifolium* |
| | + | *Polygonum minus* |
| | + | *Portulaca oleracea* |
| | + | *Rubus fruticosus* |
| | + | *Rumex crispus* |
| | + | *Verbena officinalis* |
| | + | *Veronica hederifolia* |
| | + | ***Vicia ervilia*** |
| | + | *Vicia tetrasperma* |
| Region 12 | + | *Anagallis arvensis* |
| | + | *Arrhenatherum elatius* |
| | − | *Asperula arvensis* |
| | − | *Chenopodium ficifolium* |
| | − | *Chenopodium hybridum* |
| | + | *Papaver argemone* |
| | + | *Phleum pratense* |
| | + | *Poa trivialis* |
| | + | *Polygonum arenastrum* |

Table 3 (*continued*)

| Autapomorphies | Presence (+) or absence (−) | Species |
|---|---|---|
| | − | *Polygonum aviculare* |
| | + | *Prunella vulgaris* |
| | + | *Rubus fruticosus* |
| | + | *Veronica hederifolia* |
| Region 22 | + | *Agrimonia eupatoria* |
| | + | *Bromus arvensis* |
| | + | *Euphorbia helioscopia* |
| | + | *Hyoscyamus niger* |
| | + | ***Panicum miliaceum*** |
| | + | *Plantago lanceolata* |
| | + | *Plantago major* |
| | + | *Polygonum amphibium* |
| | − | *Polygonum persicaria* |
| | + | *Rumex acetosella* |
| | + | ***Vicia ervilia*** |
| | + | ***Vicia faba*** |
| | + | *Vicia tetrasperma* |
| Region 20 | − | *Chenopodium ficifolium* |
| | + | *Galium tricornutum* |
| | + | ***Panicum miliaceum*** |
| | + | *Polygonum minus* |
| | − | *Polygonum persicaria* |
| | + | *Rumex crispus* |
| | + | *Veronica arvensis* |
| Region 21 | − | *Asperula arvensis* |
| | − | *Chenopodium hybridum* |
| | + | *Lapsana communis* |
| | − | ***Lens culinaris*** |
| | + | *Polygonum amphibium* |
| | − | *Polygonum aviculare* |
| Region 14 | + | *Aphanes arvensis* |
| | + | *Arrhenatherum elatius* |
| | + | *Atropa bella-donna* |
| | + | *Chenopodium murale* |
| | + | *Plantago lanceolata* |
| | + | *Plantago major* |
| | + | *Poa annua* |
| | + | *Poa trivialis* |
| | + | *Polygonum arenastrum* |
| | + | *Polygonum minus* |
| | − | *Polygonum persicaria* |
| | + | *Prunella vulgare* |
| | + | *Rubus fruticosus* |
| | + | *Urtica urens* |
| | + | *Veronica hederifolia* |
| Region 15 | − | *Galium aparine* |
| | − | *Polygonum aviculare* |
| | − | ***Triticum monococcum*** |
| Region 8 | + | *Chenopodium album* |
| | + | *Polygonum convolvulus* |
| Region 9 | − | ***Hordeum vulgare* var. *nudum*** |
| Region 10 | − | ***Vicia ervilia*** |
| Region 11 | + | *Agrostemma githago* |
| | + | *Fragaria vesca* |
| | + | ***Linum usitatissimum*** |
| | + | ***Papaver somniferum*** |
| | + | *Polygonum convolvulus* |
| | + | *Polygonum persicaria* |
| | + | *Portulaca oleracea* |
| | + | *Rubus fruticosus* |
| Region 13 | + | ***Papaver somniferum*** |
| | + | ***Pisum sativum*** |
| | + | ***Vicia faba*** |

Table 3 (*continued*)

| Autapomorphies | Presence (+) or absence (−) | Species |
|---|---|---|
| Region 16 | − | ***Hordeum vulgare* var. *nudum*** |
| | − | ***Triticum monococcum*** |
| Region 18 | − | ***Hordeum vulgare*** |
| | + | *Polygonum persicaria* |
| Region 17 | + | *Arenaria serpyllifolia* |
| | + | *Bromus secalinus* |
| | + | *Chenopodium album* |
| | − | ***Lens culinaris*** |
| | + | *Phleum pratense* |
| | + | *Polygonum aviculare* |
| Region 19 | − | ***Lens culinaris*** |
| | + | ***Panicum miliaceum*** |
| Region 6 | + | *Agrostemma githago* |
| | + | *Ajuga chamaepitys* |
| | + | *Bromus secalinus* |
| | + | ***Cicer arietinum*** |
| | + | *Coriandrum sativum* |
| | + | *Galium aparine* |
| | + | *Galium spurium* |
| | + | ***Linum usitatissimum*** |
| | + | *Lolium temulentum* |
| | + | ***Panicum miliaceum*** |
| | + | *Portulaca oleracea* |
| | + | *Verbena officinalis* |
| Region 4 | + | *Adonis annua* |
| | + | *Adonis flammea* |
| | + | *Carex divisa* |
| | + | *Chenopodium album* |
| | + | *Hibiscus trionum* |
| | + | *Polygonum corrigioloides* |
| | + | *Vaccaria pyramidata* |

Domestic crops are in bold.

a further sub-clade of five regions (western Germany [21], central Europe [20], central Germany [22], northwest Europe mainland [12] and Bulgaria/Macedonia [7]) linked by the gain of eight species, including *Pisum sativum*, *Lens culinaris* and six wild species (on branch 5), with a posterior probability of 71%. Within this sub-clade two of the regions, 12 (the northwestern European mainland) and 22 (central Germany), are linked further (as was the case in the maximum parsimony analysis) by having in common the gain of nine species at branch 6 (including *Papaver somniferum* and eight wild species), with a posterior probability of 54%.

## 6. Discussion

The reticulated structure returned by *NeighborNet* and the relatively low RI value yielded by the maximum parsimony analysis indicate that the dataset contains a large number of conflicting signals. However, the phylogram structures represented in Figs. 3–5 are remarkably consistent, in terms of the east–west alignment of regional groups, with the direction of the spread of Neolithic plant economies from their origins in SW Asia. The Jordanian, Syrian, Israeli, Palestinian and SE Turkish sites where the founder crops evolved (regions 1 and 2) are markedly distinct from the central Turkish, Cypriot and Thessalian Greek sites (regions 3, 4
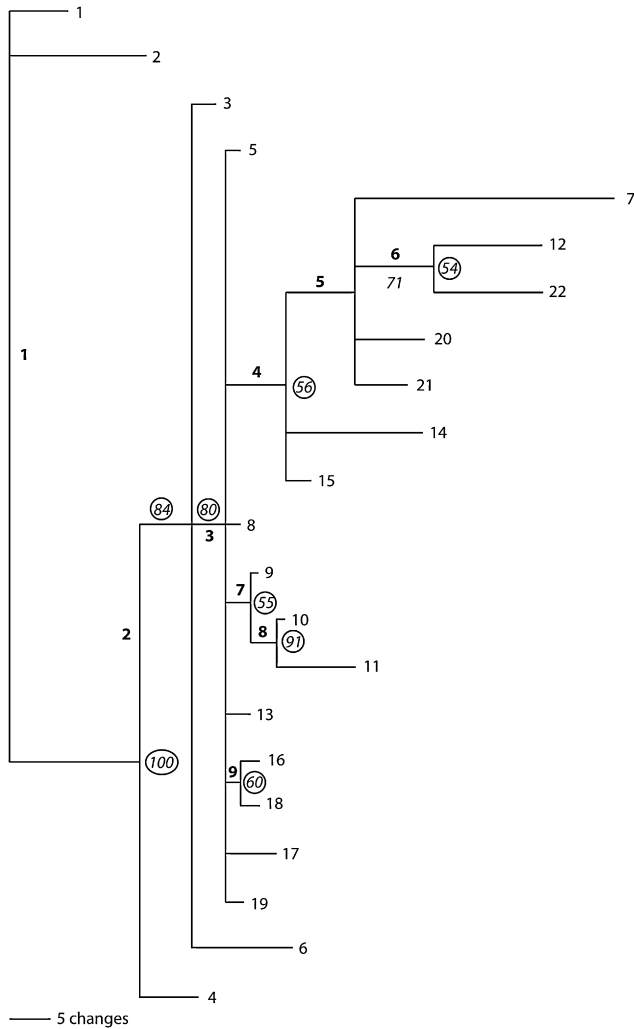
Fig. 5. Results of the Bayesian analysis: 50% majority-rule consensus phylogram produced by Bayesian analysis of the dataset. Numbers in italics are posterior probabilities associated with internal nodes; numbers in bold are branch numbers referring to Table 3.

and 6) to which they spread after the domestication events had taken place. Of note is the marked reduction of diversity at the left hand side of the phylograms. There are 24 losses (to only one gain) on branches 1–3 of the Bayesian phylogram (Fig. 5), including both domestic and wild species; for example, three of the founder crop pulses (*Cicer arietinum, Pisum sativum*, and *Vicia ervilia*) and flax (*Linum usitatissimum*) disappear from the record by branch 3 of the Bayesian phylogram.

Other crop losses at the left of the phylograms include *Vicia faba*, which is present in the Levantine region but is absent from other areas of the eastern Mediterranean. Faba bean is not a founder crop but it evolved at an early date and became incorporated into the farming systems at many early Neolithic sites. Zohary and Hopf (2000, p. 114–115) state that there is uncertainty about the ancestry of *Vicia faba* but early records on Pre-Pottery Neolithic sites suggest that it originated in SW Asia (Kislev, 1985). *Lathyrus sativus* is lost from the record in the maximum parsimony and Bayesian phylograms at

branches 4 and 3, respectively. Like faba bean, grass pea is also absent from the list of founder species but it too has been found frequently on Neolithic sites in the Aegean and west Mediterranean, and in SW Asia (Zohary and Hopf, 2000, p. 119–120). Records for the crop in regions 3, 4 and 6 (Cyprus, central Turkey and Thessaly) are thus consistent with its early cultivation in the Mediterranean basin.

As has been pointed out above, a large number of wild species losses also contribute to the distinction between the Near Eastern outgroup regions and the other regions. They comprise many species that are common in SW Asian/eastern Mediterranean habitats but not in temperate European regions, for example, *Aegilops speltoides, Arnebia decumbens, Arnebia linearifolia, Buglossoides tenuiflora, Cephalaria syriaca, Helianthemum salicifolium* and *Medicago radiata* (*Flora Europaea* online database at http://rbg-web2.rbge.org.uk/FE/fe.html). However, as we have seen, the major distinction in the phylograms is between the founder areas (and to a lesser extent Central Anatolia) and the other east Mediterranean regions, and there is no reason to believe that there would have been significant environmental differences between them to account for the dramatic scale of the losses. While they must partly relate to the natural distribution of continental SW Asian weeds, the losses are likely to result at least in part from the ecological conditions of newly cultivated fields in regions into which agriculture had been recently introduced as well as the anthropogenic effect of transportation of the crop package (Colledge et al., 2004, p. S47).

The divisions of the phylograms beyond the Near East and eastern Mediterranean regions clearly show considerable correspondence to the long-established distinction between Mediterranean and Central European routes for the spread of Neolithic economies within Europe. With the exception of region 7 (Bulgaria/Macedonia), the well-defined grouping of regions 7, 12, 22, 20 and 21 is almost entirely composed of continental LBK and LBK-derived sites, and is characterised by a number of innovations. *Pisum sativum* reappears in the record at branch 5, which distinguishes this clade. Pea is well adapted to both warm Mediterranean and cool temperate conditions (Smartt, 1990, p. 176; Zohary and Hopf, 2000, p. 101); its alliance with the group of regions comprising LBK and early Neolithic sites of western/northwestern Europe is, therefore, entirely consistent with cultivation in more northerly latitudes (see also Zohary and Hopf, 2000, p. 106–108). *Linum usitatissimum* and *Lens culinaris*, which are additions to the same group of regions, are two other founder crops that were also associated with the spread of Neolithic agriculture across the continent and into central Europe (Zohary and Hopf, 2000, p. 98–101, 127–132). The available evidence suggests that the domestication of *Papaver somniferum* (opium poppy) took place in western Europe (Bakels, 1982); there are early finds of seeds on several Linearbandkeramik sites (Knörzer, 1973, 1977, 1997; Bakels, 1979; Kreuz, 1990) so it is likely that the cultivation of opium poppy occurred at an early date in the Neolithic. Its gain at branch 6 of the Bayesian phylogram distinguishes the LBK sites of central Germany (region 22) and the LBK/LBK-derived sites

from the Netherlands, Belgium, northern France and western Germany (region 12) in accordance with this model.

This group is also characterised by very distinctive patterning in gains of wild species, specifically the addition of the typical LBK weed species (Knörzer, 1971; see also Kreuz, 1993) in successive branches of the group of regions comprising the LBK and LBK-derived northwest Europe sites. For example, at least 13 of the species (*Chenopodium album, Polygonum convolvulus, Polygonum persicaria, Echinochloa crus-galli, Galium aparine, Galium spurium, Chenopodium hybridum, Lapsana communis, Bromus secalinus, Bromus sterilis, Bromus tectorum, Veronica arvensis, Vicia hirsuta*) that commonly occur on LBK sites are added progressively at branches 4, 5 and 6 of the Bayesian phylogram. In this instance, therefore, the group is clearly underpinned by the presence of wild species in addition to the predominant crop types.

Virtually all the remaining groups are much less derived. This includes the Mediterranean regions (Crete [5], Sicily [9], southern Italy [10], central/northern Italy [11] and the Iberian peninsula/southern France [13]), but also early Neolithic Hungary and former Yugoslavia (Körös and Starčevo) and, much more surprisingly, Ireland and Scandinavia (regions 16—19). The un-derived nature of the Mediterranean suites of plants all the way to Iberia is in accordance with the well-established model for the spread of farming via the Mediterranean coast and into southern Europe (e.g. Zilhão, 2000, 2001) in environmental conditions not that different from the east Mediterranean. Notable crop gains here include *Vicia ervilia* and *Vicia faba* on branch 7, linking the Italian groups, and *Avena sativa*, which is specific to the two mainland Italian regions.

Indeed, species gains outnumber species losses at the right hand side of the phylograms and questions arise about the reasons for this. As we have seen, some crop gains, like *Papavum somniferum* in the LBK, are almost certainly local innovations. In the case of gains of domestic species that had been part of the original SW Asian crop package but were subsequently lost, like *Vicia faba* and *Vicia ervilia* in Italy or *Pisum sativum, Lens culinaris* and *Linum usitatissimum* in Central Europe, several possibilities exist. One is that they were cultivated in the intervening regions but for sampling reasons do not show up in the data available to us. This seems highly probable in some cases; for example, *Lens culinaris* is lost on branch 4 of the Bayesian phylogram and regained on branch 5. A second is that the crops concerned were indeed lost in the original dispersal events from the Levant but spread subsequently, through contacts that were still being maintained. This could be one of the reasons for the presence of the conflicting signals in the descent pattern seen in the NeighborNet network (Fig. 2). With a few exceptions our sample includes only early Neolithic sites in each region. However, the early Neolithic in the more western and northwestern regions is contemporary with middle and even later Neolithic further east and south and we have not recorded these later sites. It is possible that if we had data from middle to late Neolithic sites in the more southern and eastern parts of our area they would provide evidence for a secondary diffusion of some of these plants.

The underived nature of the plant assemblages from Ireland and Scandinavia (regions 16—19) is probably a result of sampling factors: larger assemblages will automatically include more species. Regions 16—19 contain few sites ($n = 3$, 3, 5, 2, respectively). Moreover, at least in the case of Scandinavia, the majority of identifications are from impressions, not carbonised material. Evidence of cereal grains and chaff is commonly preserved as impressions in pottery and daub; however, wild taxa are rarely found in this form so it is possible that this accounts for the less diverse range of species represented there. In this regard it is relevant to note that the mean numbers of taxa (prior to the exclusions made for this analysis) identified on sites in the three regions (Denmark, 10; southern Sweden, 7; northern Sweden, 8) are much lower than the overall mean for the entire data set (21, $n = 241$). This difference in the mode of preservation of the plant taxa could therefore account for the un-derived appearance of the Scandinavian sites.

The close link in both phylograms between southern Sweden and Ireland appears to be largely accounted for by the addition of *Galium aparine* (at branch 18 in the maximum parsimony phylogram and at branch 9 in the Bayesian phylogram). However, this may be an example of a taxon that many authors are disinclined to identify beyond the genus level (although keys have been devised to aid the identification of *Galium* species, for example, see Lange, 1979, p. 203—206). As our analysis uses only plants identified to species, and not just to genus, the link between these two regions may be solely due to shared identification practices.

Yet these explanations do not fully account for other apparent anomalies in the phylogram structures. Region 7, comprising Karanovo and Starčevo sites mainly in Bulgaria, relatively close to the east Mediterranean and Anatolia and not very far north, is the most derived region in the whole analysis. This suggests that a very distinctive local farming pattern had evolved here (see Marinova, 2001; Kreuz et al., 2005). Our results indicate that it bears a close similarity to the LBK, with which it is neither culturally nor geographically closely linked, in contrast to the conclusions of Kreuz et al. (2005) (see the supplementary material for further discussion).

The dissimilarity between region 7 and the western Balkan region (8) is also difficult to explain since they are geographically adjacent and culturally connected, as the sites in region 8 belong to the Körös and Starčevo cultures. Region 8 is much less derived than region 7 (the latter is distinguished by 30 autapomorphies, the former by only two) and corresponds much more closely to what one would expect of this region in the early Neolithic on the basis of the pattern in the rest of the phylogram. Allowing for sampling variation there is no reason why this should not be the ancestor of the more derived LBK pattern.

However, the LBK plant spectrum itself, and presumably the agricultural system of which it is a product, does not seem to be ancestral to any of the early Neolithic plant assemblages of NW Europe. Even if the un-derived state of the Scandinavian suites of plants is largely accounted for by the nature of the samples, this does not hold for northern (region

15) or southern Britain (region 14), especially the latter, which is represented by the archaeobotanical assemblages from 19 different site-phases. The reasons for this may be similar to those discussed above in relation to the loss and gain of *Vicia faba* and other crops. Bakels (in press) has shown that the crops and weeds associated with the Rössen and Michelsberg cultures in the Netherlands are markedly different from those of the preceding LBK. Moreover, it is likely that the SE Netherlands was abandoned at the end of the LBK and then reoccupied in the Rössen phase (Bakels, in press). This suggests that the Rössen phase farming system, rather than the LBK one, may have been ancestral to the sort of early farming system that characterised early Neolithic Britain, with its greater frequency of free threshing wheat and/or naked barley. This would also fit in better with the dates since Rössen belongs to the 5th millennium BC and the British early Neolithic begins c. 4000 BC. In other words, the British system may be part of a secondary rather than primary spread of crops and associated weeds, bringing species that had been present in the primary spread in regions further to the south and east, including the Mediterranean, but not in the LBK.

## 7. Conclusions

We can now return to the questions raised at the beginning of the paper. The analysis does support the claim that there should be a phylogenetic signal in the data given that the spread of crop agriculture represents a dispersal process from a broadly single origin. However, the signal is complicated by the fact that once the crop agriculture complex arrived in Europe in many areas it underwent only limited further changes and there are thus many possible ways for one region to be connected to another in a reticulated fashion. In addition, secondary spreads of some crops and associated weeds probably occurred, a phenomenon not unlike the dialect chains that can occur in the context of language dispersal and likewise leading to complex links between regions. There are also some issues about the representativeness of the data, particularly for southern Scandinavia, because of the nature of the samples.

By and large the branching pattern does indeed correspond to a combination of geographical proximity and ecological/climatic similarity, so that archaeobotanical assemblages from areas closer/similar to the source are less derived - have undergone less evolutionary change - than those from further away. However, setting aside the Scandinavian regions, whose lack of fit to this expectation probably arises from the nature of the samples, there are anomalies. The position of southern Britain (region 14) less derived than the earlier LBK and further away geographically, is probably to be accounted for by secondary crop diffusion processes of the kind already noted. More interesting are some of the other divergences.

Cyprus is remarkably derived for a region which is so close to the founder areas. This must say something about the nature of the processes acting during the dispersal (see below). The evolutionary similarity of Cyprus to the Greek regions and Europe in general also supports the arguments for an island

and coastal route from Cyprus, rather than an Anatolian route for the spread of crop agriculture and other aspects of the Neolithic to the Aegean, including Thessaly (Perlès, 2001; Colledge et al., 2004). This fits in with the recent demonstration by Perlès (2005) that a whole suite of specifically Anatolian cultural elements do not occur in Greece.

By contrast, the plant assemblage from early Neolithic Bulgaria (region 7), the first stop on the route of a continental spread of cereal agriculture via central and northwest Anatolia, is the most derived region in Europe, when it should be one of the least derived and might be expected to be similar to central Anatolia (region 4). This is a major puzzle. Perlès (2005) has postulated that there was a separate spread of Neolithic cultural features through Anatolia to the Balkans, on the basis of a series of elements common to both but not found in Greece. If this is the case then the derived nature of the Bulgarian plant assemblage raises interesting questions about subsistence practices in northwest Anatolia, about which still very little is known. It seems that even if Perlès is correct about the origin of some aspects of the Balkan Early Neolithic assemblage, and even if future research suggests that the Bulgarian suite of plants had its origin in NW Anatolia, the highly derived Bulgarian plant spectrum cannot be considered ancestral to the Körös and Starčevo assemblages of region 8, which look much more like descendants of the Greek/east Mediterranean line as well as possessing the underived features which make them the plausible ancestors to the LBK complex that would be expected on other grounds.

The general patterns seen in the geographical distribution of the different plant spectra are also largely valid for the relations between the evolutionary history of the plant economies and the archaeological cultural patterns. The distinction between the Mediterranean and Central European Neolithic dispersal routes and associated cultures, for example, is clearly apparent in the composition of the suites of plants,, but the anomalies remain. Cyprus, for example, is culturally as well as geographically close to the Levant (Peltenburg et al., 2001), but by contrast the archaeobotanical assemblages have much less in common. Early Neolithic Bulgaria is culturally very similar to the Körös and Starčevo cultures of Hungary and former Yugoslavia (region 8) but the composition of the plant suites represented in the two regions could hardly be more different. Southern Britain (region 14) looks less anomalous from a cultural as opposed to a geographical perspective for the reasons already stated.

The pattern of species losses and gains that emerges from the analysis has already been extensively discussed. What is clear from the phylograms yielded by the maximum parsimony and Bayesian analyses is the limited distinctiveness of many of the European regions. The evolutionary changes that occur between the Levant and the Aegean/southeast Europe, involving a very large number of losses mainly of wild species, are greater than all those that are found within Europe on the maximum parsimony tree. This pattern is almost as strong in the Bayesian consensus tree, which has higher posterior probabilities for the east Mediterranean branches than for all intra-European branches with the exception of one. As noted above, while this must partly relate to the natural distribution of continental SW Asian weeds, it is likely that the

main effects arise from the transmission process, and specifically from the ecological conditions in the fields in newly colonised regions and the anthropogenic effect of transportation of the crop package.

However, this observation itself raises interesting issues. Presumably the same points apply throughout the spread of farming beyond continental SW Asia, so why are the losses so limited after the initial expansion? The massive losses here could be regarded as further evidence for the initial maritime spread discussed above, since this would be far more likely to be associated with significant transport effects than overland movements. But that cannot be the whole story because we know that farming spread by sea along the north coast of the Mediterranean and our results emphasise its uniformity. However, we also know that contact along the Mediterranean coast was maintained. The spread of farming here was not a one-off event, after which contact was lost. Perhaps the initial maritime spread of farming out of continental SW Asia was precisely such a one-off event.

North of the Mediterranean area the question arises as to whether the derived features of the LBK and a few other regional plant assemblages can be accounted for in terms of adaptation of the agricultural system to new conditions. Given its location it does not seem possible to account for the derived nature of the Bulgarian pattern in this way and other work we have undertaken (Colledge et al., 2005; cf. also Bakels, in press) does not strongly support an explanation for the derived LBK pattern in terms of adaptation to new ecological and climatic conditions. This does not mean that the spread of agriculture was not an adaptive process. It was, in the sense that it represented a highly portable subsistence system that would have supported higher densities of people per unit area than available alternatives in most of the areas to which it spread, and those who adopted it would, on average, have been more reproductively successful than those who did not (Shennan, 2002, in press). However, agricultural practices constituted a complex in which the elements were adapted first and foremost to each other and an associated way of life; a portable constructed niche in other words. It thus seems to us likely that most of the modifications that appeared during the descent and dispersal of crop agriculture from the Near East to Europe arose as a result of factors affecting the transmission of the complex, but probably not all the transmission patterns of which we have evidence in our data are associated with the primary dispersal.

Finally, we can briefly address what has been gained through the use of a specifically phylogenetic analysis. At one level it has simply confirmed what has long been assumed in terms of the pattern of spread of crop agriculture. However more importantly it has enabled us to specify rigorously a hypothetical history of evolutionary changes in the complex of crops and associated weeds and to relate it to such contextual variables as geographical location and cultural affinity. This has led to consideration of the processes producing that history, including some reasons why the dispersal process did not produce a perfect tree phylogeny, as well as to the identification of some specific anomalies, such as the nature

of the Bulgarian assemblages, which raise further questions for the future.

## Appendix A. Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jas.2007.02.022.

## References

Bakels, C., 1979. Linearbandkeramik Früchte und Samen aus den Niederlanden, In: U. Körber-Grohne (zusammengetragen) Festschrift Maria Hopf: zum 65. Geburtstag am 14. September 1979. Archaeo-Physika 8:1−10. Rheinland-Verlag, Köln,.

Bakels, C.C., 1982. Der Mohn, die Linearbandkeramik und das westliche Mittelmeergebiet. Archäologisches Korrespondenzblatt 12, 11−13.

Bakels, C., in press. Nature or culture? Cereal crops raised by Neolithic farmers on Dutch loess soils, In: Colledge, S., Connolly, J. (Eds), The Origins, Spread and Use of Domestic Plants in Neolithic Southwest Asia and Europe. Left Coast Press, California.

Bogaard, A., 2002. Questioning the relevance of shifting cultivation to Neolithic farming in the loess belt of Europe: evidence from the Hambach forest experiment. Vegetation History and Archaeobotany 11, 155−168.

Bogaard, A., 2004. Neolithic Farming in Central Europe: An Archaeobotanical Study of Crop Husbandry Practices. Routledge, London.

Bryant, D., Moulton, V., 2004. NeighborNet: An Agglomerative Method for the Construction of Planar Phylogenetic Networks. Mol. Biol. Evol. 21 (2), 255−265.

Bryant, D., Filimon, F., Gray, R.D., 2005. Untangling our past: languages, trees, splits and networks. In: Mace, R., Holden, C., Shennan, S. (Eds.), The Evolution of Cultural Diversity: A Phylogenetic Approach. UCL Press, London, pp. 67−84.

Collard, M., Shennan, S.J., Tehrani, J.J., 2006. Branching, blending, and the evolution of cultural similarities and differences among human populations. Evolution and Human Behavior 27, 169−184.

Colledge, S., Conolly, J. (Eds), in press. The Origins, Spread and Use of Domestic Plants in Neolithic Southwest Asia and Europe. Left Coast Press, California.

Colledge, S., Conolly, J., Shennan, S., 2004. Archaeobotanical evidence for the spread of farming in the eastern Mediterranean. Current Anthropology 45 (suppl.), S35−S58.

Colledge, S., Conolly, J., Shennan, S., 2005. The Evolution of early neolithic farming from SW Asian origins to NW European limits. European Journal of Archaeology 8 (2), 137−156.

Greenhill, S.J., Gray, R.D., 2005. Testing population dispersal hypotheses: Pacific settlement, phylogenetic trees and Austronesian languages. In: Mace, R., Holden, C., Shennan, S. (Eds.), The Evolution of Cultural Diversity: A Phylogenetic Approach. UCL Press, London, pp. 33−54.

Harris, D., 2006. The multidisciplinary study of agricultural origins: "One World Archaeology" in practice? In: Layton, R., Shennan, S., Stone, P. (Eds.), A Future for Archaeology. UCL Press, London, pp. 229–244.

Jones, G., 1991. Numerical analysis in archaeobotany. In: van Zeist, W., Wasylikowa, K., Behre, K.E. (Eds.), Progress in Old World Palaeoethnobotany. A.A. Balkema, Rotterdam, pp. 63–80.

Kislev, M.E., 1985. Early Neolithic horsebean from Yiftah'el. Israel. Science 228, 319–320.

Knörzer, K., 1971. Urgeschichtliche Unkräuter im Rheinland ein Beitrag zur Entstehungsgeschichte der Segetalgesellschaften. Vegetatio 23, 89–110.

Knörzer, K.-H., 1973. Botanisches Material. Der Bandkeramische Siedlungsplatz Langweiler 2. Rheinische Ausgrabungen 13, 139–152.

Knörzer, K.-H., 1977. Plfanzliche Grossreste des Bandkeramischen Siedlungsplatzes Langweiler 9. Der Bandkeramische Siedlungsplatz Langweiler 9. Rheinische Ausgrabungen 18, 279–303.

Knörzer, K.-H., 1997. Botanische Untersuchung von 16 neolithischen Siedlungsplätzen im Bereich der Aldenhovener Platte, Kr. Düren und Aachen. Studien zur neolithischen Besiedlung der Aldenhovener Platte und ihrer Umgebung. Rheinische Ausgrabungen 43, 650–684.

Kreuz, A., 1990. Die ersten Bauern Mitteleuropas—eine archäobotanische Untersuchung zu Umwelt und Landwirtschaft der Ältesten Bandkeramik. Analecta Praehistorica Leidensia 23, 1–256.

Kreuz, A., 1993. Einheimische oder fremde Pflanzen? Überlegungen zur Herkunft potentieller Unkräuter und ihrer Verbreitung zur Zeit der Bandkeramik, in: Kalis, A.J, Meurers-Balke, J. (Hrsg.) 7000 Jahre bäuerliche Landwirtschaft. Entstehung, Erforschung, Erhaltung., 20 Aufsätze zu Ehren von Karl-Heinz Knörzer. Archaeo-Physika 13, Köln und Bonn, pp. 23–33.

Kreuz, A., Marinova, E., Schäfer, E., Wiethold, J., 2005. A comparison of early Neolithic crop and weed assemblages from the Linearbandkeramik and the Bulgarian Neolithic cultures: differences and similarities. Vegetation History and Archaeobotany 14 (4), 237–258.

Lange, E., 1979. Verkohlte Pflanzenreste aus den slawischen Siedlungsplätzen Brandenburg und Zirzow (Kr. Neubrandenburg), In: Körber-Grohne, U. (zusammengetragen), Festschrift Maria Hopf: zum 65. Geburtstag am 14. September 1979. Archaeo-Physika 8:191–207. Rheinland-Verlag, Köln.

Marinova, E.M., 2001 Vergleichende paläoethnobotanische Untersuchung zur Vegetationsgeschichte und zur Entwicklung der prähistorischen Landnutzung in Bulgarien. PhD. Institut für Paläontologie. Universität Bonn.

Netting, R.McC., 1993. Smallholders, Householders: Farm Families and the Ecology of Intensive, Sustainable Agriculture. Stanford University Press, Stanford.

Peltenburg, E., Colledge, S., Croft, P., Jackson, A., McCartney, C., Murray, M., 2001. Neolithic dispersals from the Levantine Corridor: a Mediterranean perspective. Levant 33, 35–64.

Perlès, C., 2001. The Early Neolithic in Greece. The First Farming Communities in Europe. Cambridge University Press, Cambridge.

Perlès, C., 2005. From the Near East to Greece: let's reverse the focus: cultural elements that didn't transfer. In: Lichter, C. (Ed.), How Did Farming Reach Europe? (BYZAS 2). German Archaeological Institute, Istanbul, pp. 275–290.

Ronquist, F., Huelsenbeck, J.P., 2005. MrBayes v3.1 (Bayesian Analysis of Phylogeny). <http://mrbayes.net>.

Shennan, S.J., 2002. Genes, Memes and Human History: Darwinian Archaeology and Cultural Evolution. Thames and Hudson, London.

Shennan, S.J. in press. The spread of farming into Central Europe and its consequences: evolutionary models, In: Kohler, T., van der Leeuw, S.E. (Eds) Model-based Archaeology. SAR Press, Santa Fe, NM.

Smith, B.D., 2001. Low-level food production. Journal of Archaeological Research 9, 1–43.

Smartt, J., 1990. Grain Legumes: Evolution and Genetic Resources. Cambridge University Press, Cambridge.

Swofford, D.L., Begle, D.P., 1993. PAUP: phylogenetic analysis using parsimony version 3.1. Laboratory of Molecular Systematics, Smithsonian Institution, Washington DC.

Swofford, D.L., 1998. PAUP* 4. Phylogenetic analysis using parsimony (*and Other Methods). Sinauer, Sunderland, MA. Version 4.

Weiss, E., Kislev, M.E., Hartmann, A., 2006. Autonomous cultivation before domestication. Science 312, 1608–1610.

Willcox, G., 2002. Geographical variation in major cereal components and evidence for independent domestication events in the Western Asia. In: Cappers, R.T.J., Bottema, S. (Eds.), The Dawn of Farming in the Near East. Studies in Early Near Eastern Production, Subsistence and Environment 6, 133–140 (ex oriente, Berlin).

Willcox, G., 2005. The distribution, natural habitats and availability of wild cereals in relation to their domestication in the Near East: multiple events, multiple centres. Vegetation History and Archaeobotany 14 (4), 534–541.

Zilhão, J., 2000. From the Mesolithic to the Neolithic in the Iberian Peninsula. In: Price, T.D. (Ed.), Europe's First Farmers. Cambridge University Press, Cambridge, pp. 144–182.

Zilhão, J., 2001. Radiocarbon evidence for maritime pioneer colonization at the origins of farming in west Mediterranean Europe. Proceedings of the National Academy of Sciences USA 98, 14180–14185.

Zohary, D., Hopf, M., 2000. Domestication of Plants in the Old World, third ed. Oxford University Press, Oxford.