ABSTRACT

This paper presents a compilation of radiocarbon dates associated with the process of arrival, development and consolidation of the first farming communities that settled between the north-western Mediterranean Arch and the High Rhine area approximately between 5900 and 2000 cal BC covering a large geographical area previously out of the main focus of $^{14}$C data compilations. The database includes dates from scientific publications and it incorporates several fields for evaluating the quality and the reliability of the available samples. It overall provides 3617 radiocarbon dates that allow the scientific community to build chronological models that can be combined with other proxies such as spatial location, type of sample or chronocultural phase.
(1) OVERVIEW

CONTEXT

The development and consolidation of farming practices has been one of the most crucial progresses for human communities. The origin of the farming populations that spread across Europe was in southwest Asia. This process started in the Aegean around 6500 BC and lasted about 2500 years [1, 2]. In the area between the northwestern Mediterranean Arch and the High Rhine, this economic shift occurred around 5900–4600 BC with the arrival of communities mainly from the central and eastern Mediterranean regions [3–6] and, to a lesser extent, from central Europe [7–9]. During the following 2000 years, farming practices evolved within a general framework of intensive and small-scale agricultural farming, until the widespread use of the plough that roughly coincides with the beginning of the Bronze Age (around 2100 BC), when more extensive farming practices developed and the productivity per household also increased [10, 11].

Thanks to the rise of performed radiocarbon measurements in the field of archaeological research, publications and funded projects where 14C dates are a key proxy have significantly increased during the last decades, generating large databases with chronological information of different periods and regions [e. g. 12, 13–17]. A quick search in Dimensions [18] for the words radiocarbon and Archaeology provides a total result of 40415 publications, 50 datasets, 249 grants, 48 patents and 69 policy documents between 2010 and 2020.

The database here presented has been compiled as part of the 4-year project AgriChange [19] funded by the Swiss National Science Foundation (SNF). The goal of this project is to collect data on crops, local climate and agricultural risk-reduction strategies to trace the factors influencing agricultural and land use changes during the Neolithic in the aforementioned territories.

This dataset aims to be a user-friendly tool to build different chronological models, thanks to the high variability of information fields included, which allow multiple approaches of analysis.

SPATIAL COVERAGE

Description: The dataset includes information from the current administrative units of six different European countries. From Italy, the provinces of: Veneto, Emilia-Romagna, Friuli-Venezia Giulia, Lombardia, Piemonte, Trentino-Alto Adige, Valle d’Aosta and Liguria. From France, the regions of: Rhône-Alpes, Provence-Alpes-Côte d’Azur, Franche-Comté, Alsace, Languedoc-Roussillon, Occitaine, Midi-Pyrénées and Grand Est. From Spain, the region of Catalunya, as well as from Andorra and Liechtenstein. As these boundaries are based on current administrative divisions and they are of limited value for our analyses, up to 14 ecoregions according to environmental and topographic criteria were differentiated [9: SM1] (Figure 1)

![Figure 1](study_area.png)

*Figure 1* Study area. White dots represent sites with radiocarbon information.
The following geographic information is provided referenced with GCS WGS 1984 using decimal degrees as angular unit:

Northern boundary: 48.310890 dd
Southern boundary: 40.522964 dd
Eastern boundary: 13.907915 dd
Western boundary: 0.159223 dd

TEMPORAL COVERAGE
Ca. 5900 – 2000 cal BC. Nonetheless, all the published dates associated with a Neolithic episode were recorded, even those ones that signify stratigraphic movements or that come from old measurements. Thus, if all dates are considered, the chronological range of the database covers ca. 34000 cal BC to 1800 cal AD.

(2) METHODS

The dates that make up the corpus of the database come from two main sources. On the one hand, the available published data from new AMS measurements of seed/fruit samples carried out in the framework of the AgriChange project [19]. And on the other hand, data collected from published sources such as scientific papers, published radiocarbon databases [e.g. 20] and grey literature [21]. A total of 3617 radiocarbon dates have been collected. Information related to the dates was likewise gathered from the same type of sources.

STEPS
Six main fields of information related to each radiocarbon date have been considered:

1. Site information. Where the ID of each date, name and type of the site appears. Only when given in publications do site coordinates appear in GCS WGS 1984.

2. Date information. Here the main information of a date (laboratory code, BP and SD) and the calibration of the raw data at 95.4%, using OxCal v 4.4.2 and the atmospheric curve IntCal20 [22, 23], of confidence -in calendar years- and its mean are provided. As in the previous case, fields related with the dating technique (AMS or LSC/GSC) and the physicochemical values of the sample dated ($\delta^{13}C$, %C, %N and C/N) will appear if they were provided in publications, since they allow to test the degree of incidence of the reservoir effect in the case of human and faunal bone samples, so as to evaluate if they must be calibrated through Marine calibration curve (e.g. 24).

3. Accuracy. This field has the purpose of evaluating the radiocarbon measurement result and the sample information. One date will be reliable (=yes) or unreliable (=no) according to several filtering criteria [9]. Unreliable include an explanation of their rejection (Why field). Additionally, there is also a weighted coefficient of the relationship between the BP and its SD. More details both on the rejection explanations and the weighted coefficient of the dates are explained in Quality Control section.

4. Sample and stratigraphic information. The information of the sample is provided in four degrees. From a broad description (e. g. seed, bone or wood) to the species (e.g Triticum monoccocum, Ovis aries or Corylus avellana), as well as intermediate categories (if no species are given) such as the provenance (if the samples belong to a, e. g., cereal, fauna or shrub) and the family/subfamily (e. g. Cerealia, Betulaceae, Caprinae). Furthermore, the stratigraphic relationship of the dated sample is given in any of the following options: level, structure and/or stratigraphic unit (SU). Finally, the chronological phase (Early, Middle or Late Neolithic) and the related technocomplex (e. g. Cardial, Chasséen or Horgen) are also provided.

5. Territorial information. This field records the location of the site according to the current administrative borders (municipality, province, region and country). Each municipality belongs to an ecoregion defined by ecological and topographic similarities [9]. The geographical coordinates, represented in GCS WGS 1984, refers to the municipality where the site is located.

6. References. The reference of where the radiocarbon information comes from and if it was obtained from a published database.

7. Observations. Lastly, we added a field for eventual important comments from both the authors of the consulted literature and the database administrators.

The data is stored in a purposely designed FileMaker file [25]. This database has been created as part of the broader AgriChange database, which contains information related to the agricultural practices (archaeobotanical data, crop stable isotope analysis, storage features, and crop pests) of sites with Neolithic records in the study area of the project.
SAMPLING STRATEGY
The available published radiocarbon dates of the north-western Mediterranean Arch and High Rhine Neolithic contexts were systematically recorded until June 2020.

QUALITY CONTROL
The reliability of the recorded radiocarbon dates was tested through various qualitative controls which led to their acceptance or rejection. Reasons for the rejection of the samples can be grouped into four. The main reason for rejecting a radiocarbon measurement is related to an archaeological bias: lack of description of the sample, lack of stratigraphic/contextual information or both are the main rejecting reasons. Another important reason to reject a sample measurement are taphonomic biases not detected during excavations, with the consequence that the dated samples do not belong to the layers where they were found. The judgements of the site archaeologists (i.e. exclusion of a date due to an inconsistency with the associated record) have hence been taken into account. As a final point, the radiocarbon measurements with higher SD than 100 have been discarded due to the high inaccuracy that they imply.

In relation to the latter point, SD values above 100 are informing (almost always) of old measurements made using outdated low-resolution techniques, such as LSC or GSC.

The SD is in relation to the error calculation of the sample measurement. To quantify the quality of the subsequent calibration result, the so-called “QuaDate” in the database was estimated, which provides a percentage value (of the SD in relation to the BP date) that allows defining the BP-SD relationship of each dating such as high quality (HQ), normal (N) or low quality (LQ). Thus, this coefficient provides a quantitative information on the precision of the time range given by the calibration process. In no case, this “QuaDate” value was used as a reason to accept or reject a radiocarbon measurement, but rather it is understood as a complement to the dates after applying the archaeological filtering criteria.

Finally, a couple of aspects should be nuanced:

1) Among the unreliable dates there is a group of dates that can still be used (“unreliable but socio-eco.” box in the database). These dates do not correspond with the desired event to be dated (i.e. a particular deposit) due to postdepositional processes, but they are direct evidence of economic or social practices, such as identified crop and livestock or human remains.

2) On the other hand, there is a large amount of long-lived (L-l) charcoal samples. This corresponds to unidentified charcoal fragments, coming from closed deposits, such as hearths. Some of these samples could be affected by the so-called old-wood effect [26]; thus their exclusion from broad analyses should depend on the research questions. For our purposes, their utility only affects the probability density but not the temporal ranges (Figure 2).

CONSTRAINTS
As mentioned above, some fields of the database have not been filled systematically due to the uncertainty of the sample or the dated context.

(3) DATASET DESCRIPTION

Table 1 provides a summary of the database according to ecoregion with information related to the reliability of the available dated samples (see the field “Why” in the “Accuracy” block of the database for further details on unreliable dates). The dataset sheds light on the differences in the quality and quantity of the radiocarbon data, mainly according to the main historical questions and record availability.

After the analysis of the radiocarbon dates, Table 1 shows that 1672 radiocarbon dates have

Figure 2 Summed Calibrated Dates Probability Distribution of all radiocarbon dates from the database (brown) and those ones classified as reliable (green).
been catalogued as reliable, while 2004 as unreliable (Figure 3a). If we analyse the reason for date rejections, in most of the cases, samples were rejected due to a poor sample description, an unreliable sample selection, a high standard deviation or postdepositional incidences.

Regarding the reliable dates, the SD is usually lower than 50 and never beyond 70, while for unreliable dates the SD is often around or greater than 100 (Figure 3b).

Finally, the quality of the available radiocarbon dates available per period can be evaluated comparing the percentage of the SD in relation to the BP date (Figure 3c). This calculation allows to characterise the Early Neolithic as the best dated period (with higher-quality dates), largely due to the recent projects aimed to temporally characterize this moment and we can therefore highlight the need for higher-quality dates for the Late Neolithic period in the study area.

Focusing on the results for the Early Neolithic in Table 1, human and wild faunal bones, as well as domestic fauna and crop remains, are the best represented samples. Domestic taxa have been particularly targeted in both Mediterranean shores’ ecoregions. Without a doubt, this fact is due to the will to temporalize the first farming communities by dating direct evidence. Indeed, the neolithization of the region had a marked maritime character, as already highlighted by other authors [i.e. 26, 27]. As can be seen in Figure 4a, the highest concentrations of dates are on the shores of Liguria, Languedoc and Catalonia, as well as the areas of the mid-lower Rhone course and Trieste karst. However, almost the opposite evidence is observed in the rest of the ecoregions.

Table 1 Table-summary of the data from the database. Bone* = faunal bones + dentitions, Charcoal* = charcoals + woods, Others* = carbonates, organic sediment, food crust, etc.
Samples of human bone are also an important group that may over-represent some areas, especially in the northern shores of the Mediterranean. In some cases, such as Les Bréguières (Mougins, Alpes-Maritimes), Arma dell’Aquila or grotta Pollera (both in Finale Ligure, Liguria) [28–31], there are up to 10-20 dates for the same event. This, together with systematic dating programmes at some sites, such as Arene Candide (Finale Ligure, Liguria) and Abri de Pendimoun (Castellar, Alpes-Maritimes), results in the concentrations of dates that are observed in the kernel density maps (Figure 4a).

The Middle Neolithic dataset is also highly influenced by dated human remains, which represent the highest proportion of all dated samples, particularly in three areas: the Alps (mainly Sion and surroundings), the Swiss Plateau (Chamblandes technocomplex) and the southern shores of the Mediterranean (Sepulcres de Fossa technocomplex). In relation to the Swiss Plateau (Figure 4b), a concentration of data points around lakes can be observed, obviously connected to the start of pile-dwelling settlements in the area [32]. By this time, dated domestic taxa (both carpological and faunal remains) significantly decrease, providing only the ecoregions of the Rhone valley and the southern shores of the Mediterranean some relevant numbers of dates on crops. Due to the effect of the concentration of the dominant assemblages mentioned above, the foci with the highest density of datings are concentrated at the mouth of the Llobregat river and its surroundings, as well as the Valais and the lakes north of the Alps. Although in a lesser degree, the occupations from the course of the Rhone river and both slopes of the Pyrenees also stand out (Figure 4b).

For the Late Neolithic period, the focus on human bone dates reaches its maximum. Without taking into account the long-lived charcoal samples, human remains represent more than 50% of all dates and, as can be seen in Figure 4c, most of them are concentrated in three areas: the Swiss Plateau, and the Mediterranean shores’ ecoregions. However, this dynamic shows different burial phenomena. While in the Swiss Plateau most of these records are related to dolmens, in the Mediterranean areas they come from caves. On the other

---

**Figure 3** Stats on the reliability of the samples. A: Accepted (light grey) and rejected (dark grey) samples for reliability: DA = Discarded by authors, HSD = high standard deviation, PP = postdepositional processes, SD = sample description, SS = sample selection. B: Barplots showing standard deviation of reliable and unreliable dates. C %BP Date-SD per phase: EN = Early Neolithic, MN = Middle Neolithic, LN = Late Neolithic.

**Figure 4** Kernel density maps with the reliable dates by period. a = Early Neolithic, b = Middle Neolithic, c = Late Neolithic.
hand, chronological information on farming practices in this period is minimal, as for other regions in Europe [33], but it does not seem to have anything to do with the importance of farming itself, as proposed by other authors [34]. Instead, it is more likely an artefact of taphonomic issues and lack of systematic radiocarbon dating of high-quality samples.

**OBJECT NAME**
AgriChange_14Cdatabase_references.txt – references cited in the database.

**DATA TYPE**
Primary, secondary, processed and interpreted data.

**FORMAT NAMES AND VERSIONS**
.csv, .xlsx, .fmp, .txt

**CREATION DATES**
The records were created between 2014–2020 as part of Berta Morell’s PhD research project [35] and the SNF AgriChange funded project “Small seeds for large purposes: an integrated approach to agricultural change and climate during the Neolithic in Western Europe” (grant number: PP00P1_170515).

**DATASET CREATORS**
Berta Morell-Rovira (University of Vigo and Eberhard Karls University of Tübingen), Héctor Martínez-Grau (University of Basel) and Ferran Antolín (University of Basel) have been in charge of collecting the published radiocarbon data. Hector Martínez-Grau homogenized and supervised the final version.

**LANGUAGE**
English. However, some stratigraphic information maintains the original language of the excavation record. Thus, for example, for the English words *level/layer* we can find *nivell/capa* (Catalan), *nivel/capa* (Spanish), *Level/Schicht* (German), *livello/strato* (Italian) or *niveau/couche* (French).

**LICENSE**
This dataset was deposited and has been released under a Creative Commons Attribution 4.0 International license.

**REPOSITORY LOCATION**
https://doi.org/10.5281/zenodo.4541470

**PUBLICATION DATE**
11/12/2020

(4) **REUSE POTENTIAL**
This database provides the most updated available radiocarbon datasets of northwestern Mediterranean and north-south western alpine Neolithic contexts.

The data can both be used as provided, since an extensive previous work has been done to filter the dates and their contexts; or, alternatively, and thanks to the exhaustive collating of the data associated with each date, any user of the database can apply their own criteria to decide which dates have greater or lesser degree of reliability for their own research questions. These dates can be used for palaeodemographic models based on summed radiocarbon probability distributions, different Bayesian modelling to determine the chronology or temporal dynamics of different regions or Neolithic techno-complex, as well as for planning guidelines for future dating practices/strategies in a regional scale, considering the quality and quantity of available dates.

Moreover, the database also provides geographical references to geospatial analysis and the dated plant and animal remains can be used as proxies for the spread of farming practices across the region.

**NOTE**
1 Done on 18th September 2020.

**ACKNOWLEDGEMENTS**
The collection of radiocarbon data and their stratigraphic association and cultural contexts would not have been possible without the help of several of the archaeologists responsible for the sites or the various study areas. It is here impossible to give a comprehensive list of names but we want to express our gratitude to all of them. We would also like to acknowledge the comments of the anonymous reviewers, that have undoubtedly improved the quality of the dataset and manuscript.

**FUNDING INFORMATION**
This research took place under the project funded by the Swiss National Science Foundation entitled “Small seeds for large purposes: an integrated approach to agricultural change and climate during the Neolithic in Western Europe” (AgriChange Project, SNSF Professorship grant number: PP00P1_170515, PI: Ferran Antolín).

**COMPETING INTERESTS**
The authors have no competing interests to declare.
AUTHOR AFFILIATIONS

Héctor Martínez-Grau  [orcid.org/0000-0002-5839-3084]
Integrative Prehistory and Archaeological Science (IPAS/IPNA),
University of Basel, CH

Berta Morell-Rovira  [orcid.org/0000-0002-3998-1921]
Departamento de Historia, Arte e Xeografía, University of Vigo, Spain

Ferran Antolín  [orcid.org/0000-0002-0533-5788]
Integrative Prehistory and Archaeological Science (IPAS/IPNA),
University of Basel, CH;
Department of Natural Sciences, German Archaeological Institute, Germany

REFERENCES


35. Morell B. La cronología como medio de interpretación social: los contextos funerarios del NE de la Península Ibérica entre finales del V e inicios del IV milenio cal BC. 2019. Universitat Autònoma de Barcelona.