

scientific data



OPEN

DATA DESCRIPTOR

A pan-European dataset revealing variability in lithic technology, toolkits, and artefact shapes ~15-11 kya

Shumon T. Hussain¹✉, Felix Riede¹, David N. Matzig¹, Miguel Biard², Philippe Crombé³, Javier Fernández-Lopéz de Pablo⁴, Federica Fontana⁵, Daniel Groß⁶, Thomas Hess¹, Mathieu Langlais⁷, Ludovic Mevel⁸, William Mills⁹, Martin Moník¹⁰, Nicolas Naudinot¹¹, Caroline Posch¹², Tomas Rimkus¹³, Damian Stefański¹⁴ & Hans Vandendriessche³

Comparative macro-archaeological investigations of the human deep past rely on the availability of unified, quality-checked datasets integrating different layers of observation. Information on the durable and ubiquitous record of Paleolithic stone artefacts and technological choices are especially pertinent to this endeavour. We here present a large expert-sourced collaborative dataset for the study of stone tool technology and artefact shape evolution across Europe between ~15.000 and 11.000 years before present. The dataset contains a compendium of key sites from the study period, and data on lithic technology and toolkit composition at the level of the cultural taxa represented by those sites. The dataset further encompasses 2D shapes of selected lithic artefact groups (armatures, endscrapers, and borers/perforators) shared between cultural taxa. These data offer novel possibilities to explore between-regional patterns of material culture change to reveal scale-dependent processes of long-term technological evolution in mobile hunter-gatherer societies at the end of the Pleistocene. Our dataset facilitates state-of-the-art quantitative analyses and showcases the benefits of collaborative data collation and synthesis.

Background & summary

The Late Glacial (c. 14.7-11.6 ka BP) is a pivotal period in European prehistory. It links the later part of the Weichselian Pleniglacial (c. 20-14.7 ka BP) – the supposed apogee of Upper Paleolithic techno-cultural efflorescence – with the subsequent Holocene (c. 11.8 ka BP-today) and its diverse woodland forager, pastoral and early agricultural societies. Situated at the Pleistocene-Holocene juncture and characterized by a series of volatile but high resolution environmental and climatic upheavals, the Late Glacial has been a central arena of scholarly debate on the impact of climate change on human societies¹⁻⁵, on palaeodemography⁶⁻¹⁰, the evolution of so-called ‘complex’ or ‘transegalitarian’ hunter-gatherers^{11,12}, and the diversification of human lifeways and adaptations that preceded post-Pleistocene trajectories of cultural evolution^{13,14}. By the same token, this period

¹Department of Archaeology and Heritage Studies, Aarhus University, Moesgård Allé 20, 8270, Højbjerg, Denmark.

²INRAP, Centre Île-de-France, Institut National de Recherches Archéologiques Préventives, 18 rue Chapelle, 89510 PASSY/UMR 8068 TEMPS, Technologie et Ethnologie des Mondes Préhistoriques, Paris, Nanterre, France.

³Department of Archaeology, Ghent University, Sint-Pietersnieuwstraat 35, 9000, Ghent, Belgium. ⁴I.U. de Investigación en Arqueología y Patrimonio Histórico, University of Alicante, Alicante, Spain. ⁵Dipartimento di Studi Umanistici – Sezione di Scienze Preistoriche e Antropologiche, University of Ferrara, Ferrara, Italy. ⁶Museum Lolland-Falster, Frisegade 40, 4800, Nykøbing F, Denmark. ⁷CNRS UMR 5199 PACEA, University of Bordeaux, Bordeaux, France. ⁸CNRS UMR UMR 8068 TEMPS, Paris, Nanterre, France. ⁹ZSBA, Schloß Gottdorf, Schloßinsel 1, 24837, Schleswig, Germany. ¹⁰Department of Geology, Faculty of Science, Palacky University Olomouc, Olomouc, Czech Republic. ¹¹Université Côte d’Azur, UMR 7264 CEPAM, CNRS, Nice, France. ¹²Natural History Museum Vienna, Burgring 7, 1010, Vienna, Austria. ¹³Institute of Baltic Region History and Archaeology, Klaipėda University, Klaipėda, Lithuania. ¹⁴Archaeological Museum in Kraków, Kraków, Poland. ✉e-mail: s.t.hussain@cas.au.dk

has also been stereotyped as documenting cultural loss and technological simplification¹⁵. Increasingly, ancient genomic data is similarly informing interpretations of movement and contact in this period^{9,16}, yet many of these interpretations remain difficult to substantiate given the lack of relevant large-scale comparative archaeological data which can be readily mobilized to test competing hypotheses or to explore scale-dependent spatiotemporal dynamics of within-period change above the level of sites or regions. Yet, due to its comparatively well-resolved human and environmental archives and limited recovery bias compared to older periods, European Late Glacial archaeology harbors the potential to benchmark hunter-gatherer research in other periods of the human deep past. To do so, information on stone artefact variability, production technologies and design choices (henceforth: lithic data) is vital for our understanding of early human evolution as stone artefacts are abundant, durable and yield rich insights into fundamental dimensions of human behavior^{17,18}.

We here describe an open-access dataset¹⁹ compiled by a concerted effort of lithic experts across Europe to help overcoming this problem and to begin asking pan-European questions on synchronic and diachronic techno-cultural variability, and to facilitate critical interrogations of the scope and quality of current knowledge and the significance of higher-level taxonomic denominations to organize the record (e.g. Final Magdalenian, Hamburgian, Azilian, Federmesser, Ahrensburgian, Swiderian, Belloisian, etc.)^{20,21}. The dataset¹⁹ was constructed with the aim of 1) rendering long-standing hypotheses and narrative framings of the Late Glacial amenable to systematic empirical scrutiny, and 2) breaking away from terminological and semantic discussions and instead to open new possibilities for data-driven investigations of millennial-scale patterns and trends in human technological behavior beyond labels and categorizations largely contingent on research history. Two complementary but interrelated concerns motivate our effort: a) the growing recognition of scale in archaeology and cultural evolutionary studies with the resulting need to take serious hitherto largely neglected macro-archaeological processes and phenomena^{22,23}, and b) the impeding effects of regional research traditions and culture-historical frames of reference for regional and continental-scale comparison^{20,24}, to address both cross-sectional and longitudinal questions in Late Glacial archaeology and beyond. Our efforts resonate with calls for synthesis and higher-level archaeological analysis^{21,25} using ‘big data’ or at least unprecedentedly large datasets, while also duly accounting for the fragmentary and complex nature of the available material registers^{26–28}. The latter draws attention to a nascent transformation in knowledge ecologies based on an ethos of sharing and collaborative inquiry^{25,29–31} necessary to develop data infrastructures promoting broad-scale comparative investigations at cross-regional scales. The here-presented dataset¹⁹ is a first attempt to work towards such a shared data infrastructure for Late Glacial Europe, to build a scalable higher-order data framework for lithic analysts and to probe into the untested utility and potential of such integrated collaborative work. In contrast to previous database projects in Pleistocene archaeology^{32,33}, our point of departure is not individual sites and the lithic assemblages they host, but amalgamated higher-order units which reflect the current state of research in different European regions, allowing us to transparently assemble data on different lithic domains operating on varying scales of observation, in turn offering unique possibilities of archaeological analysis and comparison.

Methods

Taxonomic units, macro-regions and time slices. The deep prehistory of European forager populations is commonly framed in cultural-historical terms^{1,2,5,34–36} – that is, as a sequence of variously overlapping taxonomic units, although the precise definition of these units as ‘techno-complexes’, ‘techno-cultures’ or ‘industries’ may differ between research contexts^{24,37}. Taxonomic units or ‘named archaeological cultures’ (NACs) are the most widely recognized higher-order archaeological units used to interrogate the long-term evolution of Pleistocene societies. NACs thus offer a suitable higher-order reference system for organizing and synthesizing currently available knowledge on lithic behavioral variability and evolution across different parts of Late Glacial Europe. NACs are situated in space and time: we employ macro-regions reflecting geographic and ecological differences but also scholarly expertise and predefined time slices (binned in equidistant 1k-year intervals) to construct the data framework. Geographic macro-regions ($n = 16$) and employed time slices ($n = 4$) are shown in Fig. 1 (cf. Supplementary Table 1). Time slices I to IV span the period between 15 and 11 ka BP, including the terminal part of the Pleniglacial (Late Upper Paleolithic), the Late Glacial (Final Paleolithic) and the incipient Holocene (earliest Mesolithic). The individual NACs ($n = 86$) were assembled from the literature based on the current state-of-the-art of lithic research in a given macro-region and then assigned to a single or multiple time slices depending on the latest relative, stratigraphic and/or chronometric dating evidence. For each NAC, a broad range of lithic variables pertaining to different levels of technological organization were then extracted from the currently available literature by expert editors acquainted with a given macro-region.

Collaborative data compilation and recording process. The workflow of collaborative data compilation, preparation and synthesis is outlined in Fig. 2. Each macro-region was handled by an expert editor intimately familiar with the regional lithic evidence and the latest archaeological debates and interpretations. In a first step, regional lithic experts carefully reviewed the literature and identified the latest culture-historical framework accepted for a given macro-region, thereby assembling a regional inventory of NACs. Expert editors were then confronted with a first draft recording scheme integrating a broad range of lithic variables and asked to synthesize and test-record the published lithic data for each inventoried NAC within their macro-region accordingly. Initial insights and encountered problems were subsequently discussed by all research participants on a two-day online workshop to adjust and improve the data infrastructure and to align expectations and basic definitions as well as to address open questions²¹. Based on the critical feedback received, the lithic data structure was updated, and all initial entries thoroughly checked by the CLIOARCH team. Datasets were then sent back to the expert editors, who had to review and revise their provided data and address any identified issues and inconsistencies. Expert editors were supported by CLIOARCH core members throughout the entire process to ensure overall data coherence and to provide feedback and/or a second opinion if needed. The revised macro-regional datasets were then

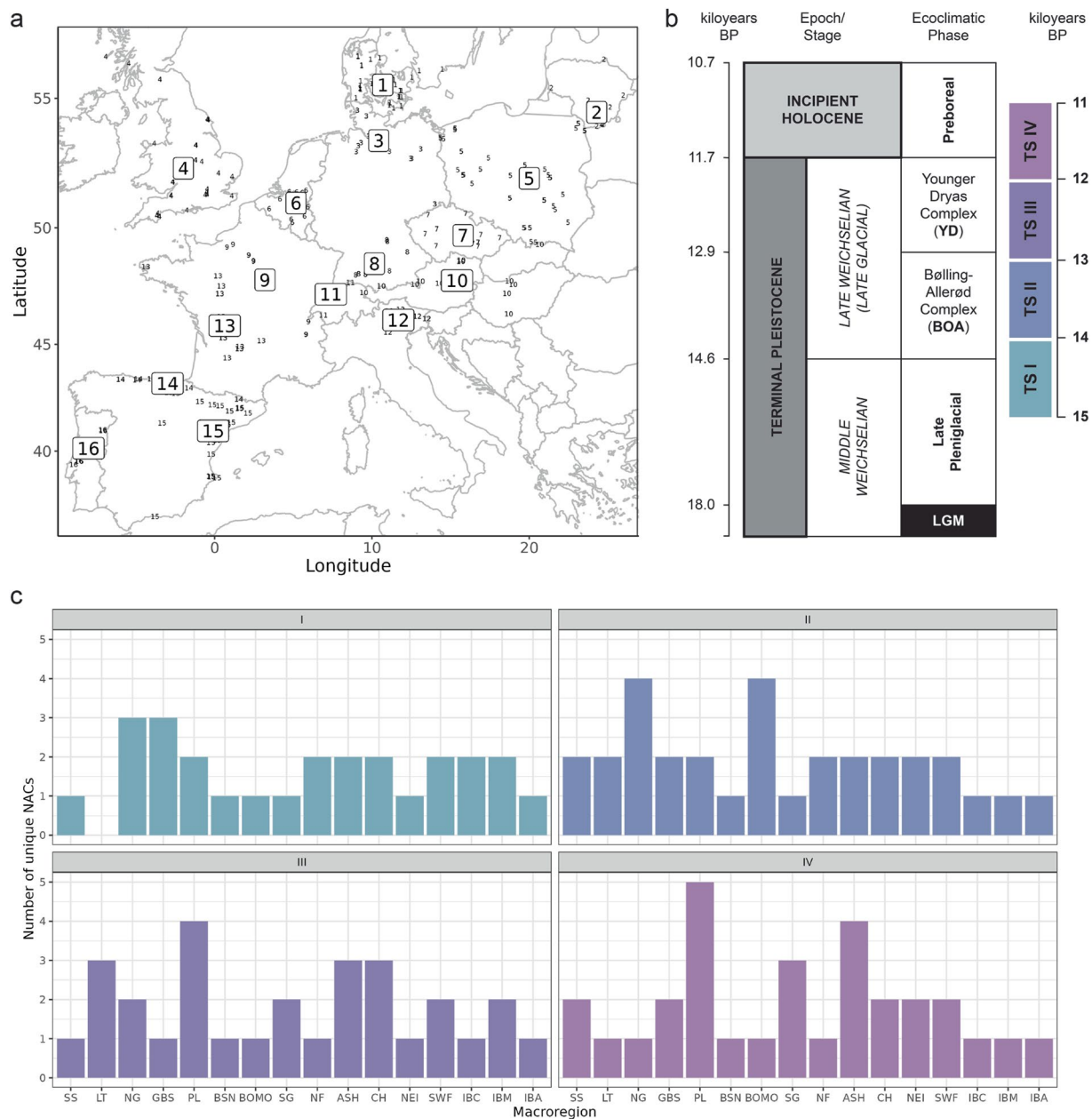


Fig. 1 Spatiotemporal characteristics of the dataset¹⁹. **(a)** Geographic overview of macro-regions (boxed numbers) and their corresponding key sites (small numbers); 1 = Southern Scandinavia (SS), 2 = Lithuania (LT), 3 = Northern Germany (NG), 4 = Britain (GBS), 5 = Poland (PL), 6 = Belgium and Southern Netherlands (BSN), 7 = Bohemia and Moravia (BOMO), 8 = Southern Germany (SG), 9 = Northern France (NF), 10 = Austria, Slovakia and Hungary (ASH), 11 = Switzerland (CH), 12 = Northeastern Italy (NEI), 13 = Western France (SWF), 14 = Cantabrian Spain (IBC), 15 = Mediterranean Iberia (IBM), 16 = Atlantic Iberia (IBA). **(b)** Time slices I to IV in relation to established chronological schemes in Late Glacial/incipient Holocene archaeology. **(c)** Number of expert-submitted named archaeological cultures (NACs) per time slice and macro-region.

submitted for final data integration and preparation. The recorded individual lithic variables are shown in Fig. 3 and described in detail in Tables 2–9 (see below) and in Data Records.

Dataset design and data collection principles. The 1511NAC dataset¹⁹ is designed to facilitate macro-archaeological analyses of lithic diversity, complexity and cultural evolution in Europe between 15 and 11 ka BP²¹. Part of this endeavor is the examination of potential offsets between different lithic data domains and/or dimensions of lithic production technologies and toolkits. The overall structure of the dataset¹⁹ reflects this analytical aim (Table 1): the first data domain (I) gathers NAC-level information on the structural composition of lithic toolkits based on a broad distinction of lithic morphotypes divided into ‘armatures’ and ‘domestic’ tools respectively (Tables 2, 3); the second data domain (II) draws together NAC-level lithic information on

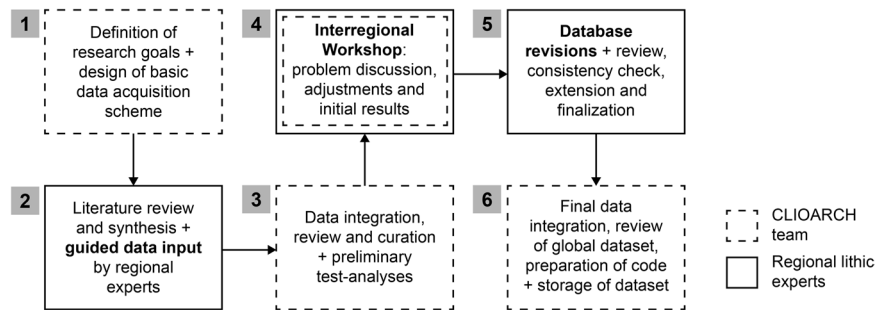


Fig. 2 Workflow of collaborative data compilation and synthesis. Numbers indicate successive stages of data processing, from definition and preparation to validation.

the organization of production technologies in the form of discrete traits, focusing especially on the laminar component (Tables 4–9); the third data domain (III), finally, consists of a library of 2D outline shapes of key tool groups (armatures, endscrapers and borers/perforators) extracted from the literature and digitally prepared for analysis with site-level resolution. The first and second data domains are fed with presence/absence data (categorical) coded as binary characters (1/0), or as n/a in cases where the respective information was not available or ambiguous. Cases for which a decision between 1 or 0 could not reasonably be made based on the present state of publication were originally recorded as ‘?’, but have been amended to ‘n/a’ as the latter is more conservative and inter-observer disparities are otherwise likely to introduce yet another artificial layer of variability. Presence data coded as 1 represents NAC-wide recurrence of occurrence, and thus records technological traits and tool morphotypes that are consistently but not necessarily always observed within archaeological sites grouped under a given regional NAC. The third data domain consists of digitized shape data (continuous) from individual key sites attributed to regional NACs, provided as prepared image files and outlines. The contents of all data domains are described in more detail in the next section.

Data Records

Folder structure. The complete 1511NAC dataset is openly available at Zenodo¹⁹ and structured as shown in Fig. 4. The data repository comprises four principal folders: *1511NAC_Dataset*, *1_data*, *2_scripts* and *3_output*. Generally speaking, the folder *1511NAC_Dataset* contains all data entries relevant for researchers who are interested in using the dataset¹⁹. The folders *1_data*, *2_scripts* and *3_output* contain the raw data and R scripts that were created and used to generate the dataset¹⁹ and the figures in this paper. Each principal folder is described in more detail below.

1511NAC_Dataset

- The *artefact_images_by_region* subfolder holds the prepared images of all single lithic artefacts in .jpg format, organized per region and ordered according to the three broader artefact groupings (data domain III/2D artefact outlines, see below). Each .TPS file (thin-plate-spline) contains the pixel-to-centimeter ratio information for each set of lithic artefacts per region per artefact grouping, enabling further metric analysis.
- The *outlines* subfolder contains all .RDS 2D outlines generated from the prepared artefact images. For each of the three larger artefact groupings (armature, endscrapers, borers/perforators, see below), a single .RDS file containing the respective artefact outlines with their metadata and metric measurements is provided, as well as a separate .csv file holding their respective metadata and metric measurements alone, and thus facilitate simple metric analyses. To display the .RDS files properly in R, the *Momocs* package ($\geq 1.4.0$ ³⁸) must be installed and loaded.
- The *references* subfolder gathers all primary literature referenced in the dataset¹⁹ prepared in different file types to be directly imported into currently available bibliographic management tools.
- The three master tables are provided as separate .csv files. They contain the metadata of all archaeological key sites represented in this dataset¹⁹ (*MASTERTABLE_v7_all_revised_sites_with_outline_siteID.csv*) as well as the recorded trait data of data domains I and II (Toolkits: *MASTERTABLE_v7_all_revised_tools_with_Tsrange_uniqueID.csv*; and Technology: *MASTERTABLE_v7_all_revised_technology_with_TSrange_uniqueID.csv*).

1_data

This folder contains both the raw and the final data as contained in the *1511NAC_Dataset*, as well as the data required to reproduce the figures and tables included in this paper. The files and the dataset¹⁹ included here are primarily provided for full transparency. They allow researchers to trace the entire process of data processing leading to the final dataset¹⁹ supplied.

- The *artefact_images_by_region* subfolder is identical as the above-described.
- The *raw* subfolder contains the original, unprocessed .csv files holding the site-specific metadata as well as the trait-data of data domains I and II as submitted by the macro-regional expert editors.

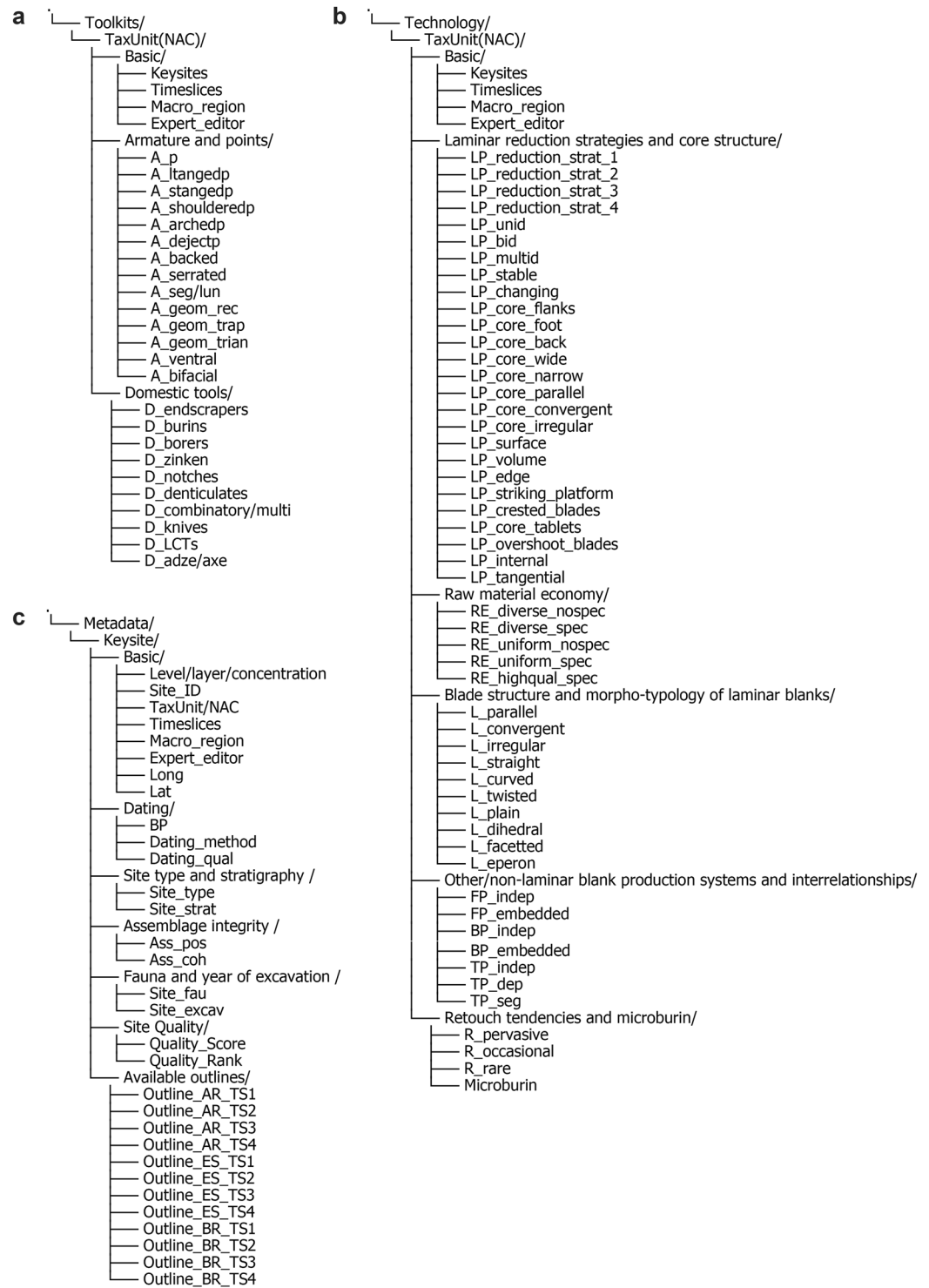


Fig. 3 Trait structure of main data domains and site-specific metadata. (a) Toolkits data domain (I). (b) Technology data domain (II). (c) Metadata.

- The *prepared* subfolder contains all data derived from the raw master tables and the prepared lithic artefact images, etc., as well as all transformed data. The folder therefore contains all data necessary to reproduce the figures and tables in this paper and includes the same final outline data and master tables as contained in the main *1511NAC_Dataset* folder.
- The *references* subfolder is identical as the above-described.

Data domains	Sub-domains	Number of recorded traits	Data type	Total observations	Data resolution
<i>Toolkit</i>	Armature/points; Domestic tools	24	Discrete	86	NAC-level
<i>Technology</i>	Laminar reduction strategies and core structure; Raw material economy; Blade structure and morphology of laminar blanks; Other/non-laminar blank production systems and their interrelationships; Retouch tendencies (+ microburin)	52	Discrete	86	NAC-level
<i>2D Artefact outlines</i>	Armature (AR); Endscraper (ES); Borer (BR)	34 (AR)	Continuous	5986 (AR = 3512, ES = 1930, BR = 544)	Object-level (and site-level)
Metadata	Site context and data quality; Calculated site-quality scores; Register of processed and recorded tool outlines	25	Various	—	Site-level

Table 1. Structure of the dataset¹⁹ and descriptive parameters of the main data domains.

Code	Description
A_p	Simple points
A_langedp	Large tanged points (width >15 mm)
A_stangedp	Small tanged points (width <15 mm)
A_shoulderep	Shouldered points
A_archedp	Arched/Arch-backed points
A_dejectp	Dejected/Angle-backed points
A_backed	(Other) Backed pieces
A_serrated	Serrated/denticulated implements
A_seg/lun	Segments/lunates
A_geom_rec	Rectangular (geometric) microliths
A_geom_trap	Trapezoid (geometric) microliths
A_geom_trian	Triangular (geometric) microliths
A_ventral	Partial ventral retouch
A_bifacial	Complete bifacial retouch (shaping)

Table 2. Categorical NAC-level lithic armature data recorded as systematic presences/absences in data domain I (Toolkits).

Code	Description
D_endscrapers	Endscrapers
D_burins	Burins
D_borers	Borers/perforators (including <i>becs</i> and other atypical forms)
D_zinken	<i>Zinken</i>
D_notches	Notched pieces
D_denticulates	Denticulated pieces
D_combinatory/multi	Combinatory tools/multitools
D_knives	Large blade knives/heavily retouched larger blades
D_LCTs	(Other) Large Cutting Tools
D_adze/axe	Flake adzes/axes

Table 3. Categorical NAC-level domestic lithic tool data recorded as systematic presences/absences in data domain I (Toolkits).

2_scripts

This folder exists purely for the purpose of reproducibility and contains the R scripts used to prepare and extract the data which are then deposited in *1511NAC_Dataset*, but also the scripts used to generate the figures and tables of this paper.

- *01_datasheet_modifications.R* is used to prepare the spreadsheets containing information on archaeological key sites and those holding data domains I and II.
- *02_datasheet_summaries_plots.R* is used to data-wrangle and to create the figures and tables in this paper related to site-specific metadata and data domains I and II.
- *03_get_outlines.R* is used to extract outlines from the supplied .jpg images, to further prepare them, and to calculate their associated metric measurements.

Code	Description
LP_reduction_strat_1	One single focal reduction strategy
LP_reduction_strat_2	Two focal reduction strategies
LP_reduction_strat_3	Three focal reduction strategies
LP_reduction_strat_4	Four focal reduction strategies
LP_unid	Undirectional laminar production
LP_bid	Bidirectional laminar production
LP_multid	Multidirectional laminar production
LP_stable	Reduction patterns are stable across core life-histories
LP_changing	Reduction patterns change across core life-histories
LP_core_flanks	Prepared/well-defined core-flanks
LP_core_foot	Prepared/defined core foot
LP_core_back	Prepared/defined core back
LP_core_wide	Wide reduction surfaces
LP_core_narrow	Narrow reduction surfaces
LP_core_parallel	Geometry of reduction surfaces highlights parallel configurations
LP_core_convergent	Geometry of reduction surfaces highlights convergent configurations
LP_core_irregular	Irregular reduction surface geometries

Table 4. Categorical NAC-level data on laminar reduction strategies, core structure and exploitation surface geometry recorded as systematic presences/absences in data domain II (Technology).

LP_surface	Surface exploitation of cores (surface-near volumes are targeted for exploitation)
LP_volume	Volume(tric) exploitation of cores (core reduction concepts in principle allow for the exploitation of whole core volumes)
LP_edge	(Narrow) Edge exploitation (production of blanks from natural or artificial narrow edges without exploitation of whole core volumes, including burin-like productions)
LP_striking_platform	Prepared striking platform of cores
LP_crested_blades	Crested/neocrested blades
LP_core_tablets	Core tablets
LP_overshoot_blades	Preparatory/corrective overshoot blades
LP_internal	Internal knapping (knapping gesture results in point of impact on the interior of the striking platform, often associated with thick blank platforms/butts and pronounced striking features such as bulbs)
LP_tangential	Near-edge/peripheral knapping (knapping gesture results in point of impact very close or at the core edge, often resulting in marginal blank platforms and diffuse striking features)

Table 5. Categorical NAC-level data on core volume management and preparation recorded as systematic presences/absences in data domain II (Technology).

Code	Description
RE_diverse_nospec	Focus on a diverse raw material base but no obvious specialization (i.e., no specific link between a particular raw material and a particular reduction strategy)
RE_diverse_spec	Focus on a diverse raw material base with evidence on specialization on (a) selected raw material(s)
RE_uniform_nospec	Focus on uniform raw material base yet no obvious specialization, i.e., adaptation of the knapping strategy to the given raw material (properties, forms)
RE_uniform_spec	Focus on a uniform raw material base with evidence for specialization, i.e., adaptation of knapping strategies to raw material conditions/properties
RE_highqual_spec	Global focus on high quality raw material with evident adaptation of knapping strategies to the properties of the raw material

Table 6. Categorical NAC-level raw material economy data recorded as systematic presences/absences in data domain II (Technology).

- *04_summary_artefacts.R* is used to data-wrangle and to create the 2D outline-related figures and tables shown in this paper.

3_output

This folder contains the original figures and tables created for this article.

Further information and explanation on file naming conventions in the dataset¹⁹ can be found in the Supplementary Information.

Code	Description
L_parallel	Parallel outlines
L_convergent	Convergent outlines
L_irregular	Irregular outlines
L_straight	Straight laminar profiles
L_curved	Curved laminar profiles
L_twisted	Twisted laminar profiles
L_plain	Plain platforms
L_dihedral	Dihedral platforms
L_facetted	Facetted platforms
L_eperon	<i>En éperon</i> platform preparation

Table 7. Categorical NAC-level data on blade structure and the morpho-typology of laminar blanks recorded as systematic presences/absences in data domain II (Technology).

Code	Description
FP_indep	Independent flake production
FP_embedded	Embedded flake production (dependent on laminar production(s))
BP_indep	Independent bladelet production
BP_embedded	Embedded bladelet production (dependent on blade or flake production(s))
TP_indep	Tool production/manufacture independent of blank production (i.e., no systematic link between specific blank and tool forms/types)
TP_dep	Tool production/manufacture dependent on or follows/ anticipates blank production (i.e., evidence for a systematic link between specific blank and tool forms/types)
TP_seg	Segmentation of domestic tools and armature (blank-tool relations differ between the two, different blanks are selected and transformed in both cases, either within the same reduction strategy or the two tool categories are produced from separate reduction strategies)

Table 8. Categorical NAC-level data on other (non-laminar) reduction systems and system-interrelationships recorded as systematic presences/absences in data domain II (Technology).

Code	Description
R_pervasive	Assemblage-level retouch is pervasive (>20%)
R_occasional	Assemblage-level retouch is occasional (10–20%)
R_rare	Assemblage-level retouch is rare (<10%)
Microburin	Microburin products and by-products

Table 9. Categorical NAC-level data on retouch-tendency and microburin technology (all lithic artefacts ≥ 2 cm) recorded as systematic presences/absences in data domain II (Technology).

Analytical units and data domains in detail. The general trait-structure of the dataset¹⁹ is shown in Fig. 3. The primary analytical units are the macro-regional NACs and in the case of outline data, both NACs and individual archaeological key sites. We provide each data domain as a separate .csv file for easy and flexible implementation. Linked metadata provides further contextual information on the key sites respectively attributed to macro-regional NACs, such as lithic assemblage provenance, site quality, radiometric dating, and the underlying bibliographic sources.

Cultural taxonomic units. The dataset¹⁹ comprises 86 separate NAC entries. Supplementary Table 3 shows how these are spread over time slices I to IV and the 16 captured macro-regions. These NACs cover a large portion of Northern and Western Europe and circumscribe the following geographic macro-regions: Atlantic Iberia, Cantabrian Iberia, Mediterranean Iberia, Western France, Britain, Northern France, the Low Countries, Switzerland, Southern Germany, Northern Germany, Southern Scandinavia, Northeastern Italy, Austria/Slovakia/Hungary, Moravia/Bohemia, Poland, and Lithuania. NAC naming conventions follow common literature practice, but priority was always given to the state-of-the-art understanding of regional experts. NACs were either recorded as bounded units (e.g., ‘Magdalenian’, ‘Federmesser’, ‘Ahrensburgian’ or ‘Belloisian’) or as chronological sub-units (e.g., ‘Final Magdalenian’, ‘Havelte’, ‘Early Azilian’ or ‘Late Laborian’), depending on the importance of the latter in the archaeological discourse and the prevailing opinion/consensus in the respective body of expert literature. Individual NACs identified and recorded in this way are linked to a portfolio of key archaeological sites anchoring them in space and time and exemplifying the lithic information assembled in the toolkit and technology data domains (n = 350). These sites tend to be the best-studied site exemplars of a NAC in

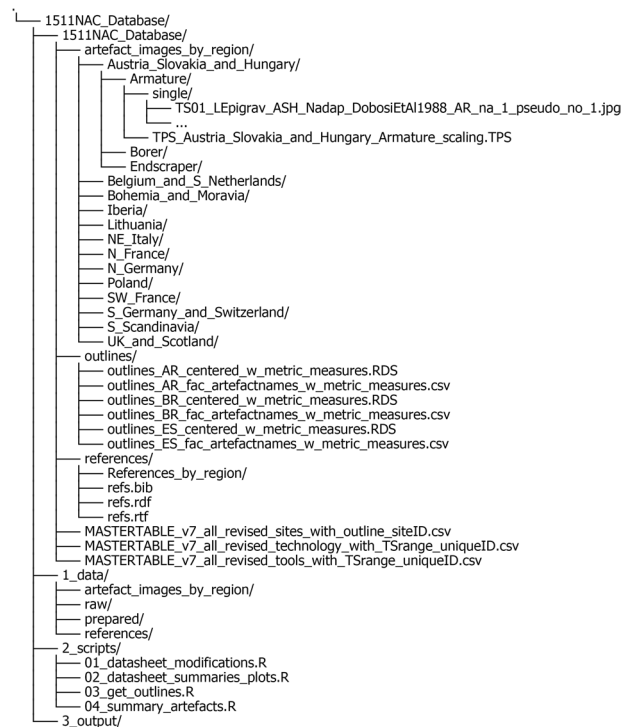


Fig. 4 Folder tree structure of the data repository¹⁹.

each macro-region, hold key information on the chronological position of said NAC, and/or frame a rich source of drawings or photographs of NAC-specific lithic tools.

Toolkit data. Discrete binary toolkit data (presence/absence) for different NACs is provided for 24 tool morphotypes subdivided into two broader groups of ‘armatures’ and ‘domestic’ tools, constituting data domain I. The armatures group mainly gathers various lithic point types, backed implements and microlithic forms, while the domestic tools group covers the near-ubiquitous Upper Paleolithic tool types of burins, endscrapers and borers/perforators. The following morphotypes are included in the armatures group ($n = 14$): simple points, large tanged points (width > 15 mm), small tanged points (width < 15 mm), shouldered points, arched/arch-backed points, dejected/angle-backed points, (other) backed pieces, serrated/denticulated implements, segments/lunates, rectangular (geometric) microliths, trapezoid (geometric) microliths and triangular (geometric) microliths, as well as separate entries for the recurrence of NAC-level partial ventral and bifacial retouch (cf. Table 2). The domestic tools group sports the following morphotypes ($n = 10$): endscrapers, burins, borers (including perforators, *becs* and other atypical forms), *Zinken* (borers with a typical, strongly bent nose), notched pieces, denticulated pieces, combinatory tools/multitools, large blade knives/heavily retouched larger blades, (other) LCTs (Large Cutting Tools) and flake adzes/axes (cf. Table 3). Presence data (1) indicates NAC-level recurrence and thus inter-assemblage importance of these morphotypes.

Technology data. Discrete binary data on technology (presence/absence) for different NACs is recorded in 54 individual techno-typological traits, organized in five sub-domains which together make up data domain II. Again, these data operate on the aggregate level and the recorded information refers to NAC-wide trends and patterns not necessarily reducible to observations pertaining to individual archaeological sites. We assume that NAC-level records represent the observations made on multiple archaeological key sites and so provide better information overall. Building on a long tradition of lithic technological analysis, the five technology sub-domains are: (A) laminar reduction strategies and core structure, (B) raw material economy, (C) blade structure and morpho-typology of laminar blanks, (D) other blank production systems and system interrelationships, (E) retouch tendencies and microburins. Sub-domain A encompasses 26 traits pertaining to the number of co-existing laminar core reduction strategies (distinct modes of working selected lithic raw material nodules), the extent and dynamics of core preparation (e.g., investment and stability of approach), the geometry of reduction surfaces (shape and edge configuration), and the knapping technique (internal vs. tangential) (cf. Tables 4, 5). Sub-domain B draws together 5 traits related to the relationship between raw material choice and reduction strategies (cf. Table 6). Sub-domain C features 10 traits linked to the morphology, profile and platform configuration of laminar blanks (cf. Table 7). Sub-domain D gathers 7 traits targeting the interdependency of different blank production systems (flake, blade, bladelet) and the relationship between blank production and tool manufacturing (cf. Table 8). Sub-domain E, finally, compiles 4 traits pertaining to the scope/intensity of recurrent assemblage-wide retouch (frequency of tools/modified pieces) and as the microburin technique (cf.

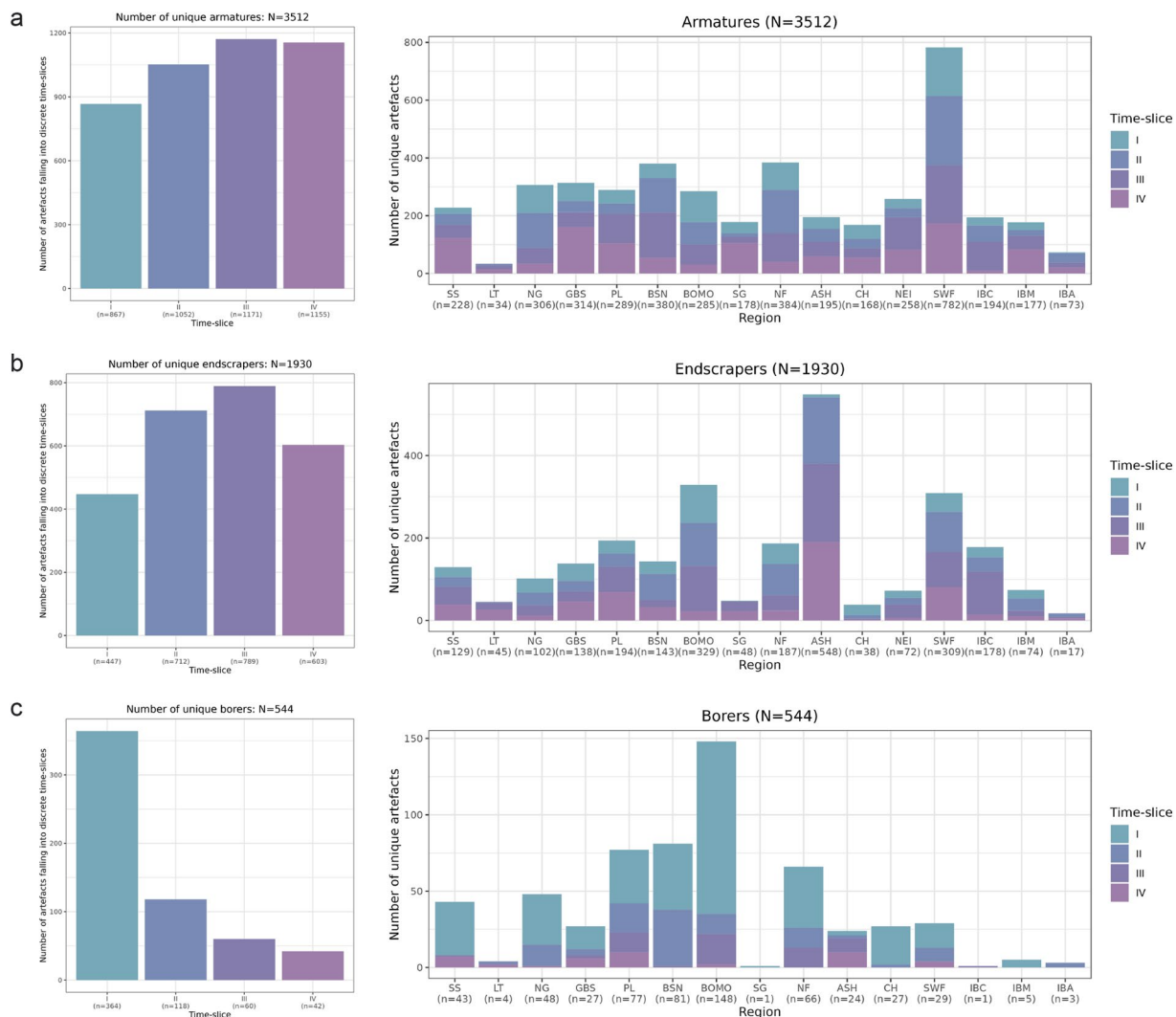


Fig. 5 Overview of lithic artefact outlines in the dataset¹⁹. **(a)** Number of digitized armature outlines (AR) for each time slice (*left*) and their distribution across individual macro-regions (*right*). **(b)** Number of digitized endscrepper outlines (ES) for each time slice (*left*) and their distribution across individual macro-regions (*right*). **(c)** Number of digitized borer/perforator outlines (BR) for each time slice (*left*) and their distribution across individual macro-regions (*right*). Note that the borer dataset includes incomplete artefacts.

Table 9). These discrete trait data enable the examination of relationships between sub-domains and broader spatiotemporal trait dynamics.

2D artefact outlines. The 2D outline library (data domain III) hosts 5986 individual lithic tool outlines, compiled and grouped in three broader categories: armatures (AR), endscreppers (ES), and borers (BR) with reference to the categories listed in the toolkit data domain. Burins were not collected because of the huge variability of artefact shapes associated with this tool group and because the so captured variability is shaped by extreme levels of blank-derived noise. The AR group comprises 3512 artefact outlines linked to tool morphotypes classified as ‘armatures’. The ES group contains 1930 outlines identified as ‘endscreppers’. The BR group contains 544 outlines identified as ‘borers’ and perforators. AR and ES outlines were collected from complete lithic artefacts only, or artefacts for which the missing part (for example, the apex of a point) could unambiguously be reconstructed. BR outlines do not respect this criterion and must therefore be handled with caution; here, focus rests on the working end of the artefacts rather than the total outline. These data are provided both as fully processed outlines saved as .RDS in R to preserve as much information as possible, and as individual image files pre-processed into binary artefact images which can directly be fed into outline extraction and shape-description workflows³⁹. These data can thus be deployed in outline analyses, using geometric morphometric methods of shape space analysis⁴⁰. They provide the hitherto largest repository of key tool forms from the European Late Glacial. Each artefact image is labelled according to the conventions defined in the Supplementary Information (Part 4, cf. Supplementary Table 8), specifying its NAC, time slice, macro-region, site of origin, and bibliographic reference(s). In addition, AR outlines come with supplementary information on point axiality, distinguishing

Code	Description
Long	Longitude coordinates (given in decimal degree)
Lat	Latitude coordinates (given in decimal degree)
BP	Broad dating range in cal. BP, e.g., “15-13 calBP”
Dating_method	Dating method, e.g., “14 C”, “TL”, “typology”, “technology”, “biostratigraphy” or “geology”
Dating_qual	Dating quality (drop-down): “Reliable” or “Problematic”
Site_type	Type of site (drop-down): “Openair”, “Cave” or “Rockshelter”
Site_strat	Stratification (or not) of site (drop-down): “Stratified” or “Surface”
Ass_pos	Assemblage position (drop-down): “Primary/insitu” or “Secondary/relocated”
Ass_coh	Assemblage coherence (drop-down): “Homogeneous” or “Mixed”
Site_fau	Presence/preservation of faunal material (drop-down): “Preserved” or “No”
Site_excav	Period of excavation (drop-down): “Before WW2”, “1950–1980”, “1980–2000” or “After 2000”

Table 10. Basic site-specific information and data quality indicators.

between distal (‘d’) and proximal (‘p’) lithic points if possible. A breakdown of the spread of outlines across time slices and macro-regions is provided in Fig. 5.

The outline datasets¹⁹ also harbors metric artefact-level information, even though this data is only available at the sub-sample level and was not systematically collected. Where scale bars were available on the submitted images (AR: n = 2966; ES: n = 1628; and BR: n = 498), a pixel-to-centimeter ratio/scaling factor was created for each image separately using *tpsUtil* and *tpsDig2* in R⁴¹. This scaling factor was then implemented in the outline extraction process³⁹ using the *outlineR* package⁴² to automatically associate a scaling factor – where applicable – with individual outlines, allowing for the quick and ready extraction of a multitude of user-defined artefact-specific metric measurements. The dataset¹⁹ supplies the following artefact-level measurements extracted via the *Momocs* package³⁸: length (cm), width (cm), area (cm²), perimeter (cm), calliper (cm), circularity and eccentricity (see³⁸ for a definition and description of these specific measurements).

Metadata. Contextual and supplementary information on the selected archaeological key sites (Supplementary Tables 2, 5) anchoring regional NACs in the dataset¹⁹ is organized according to three sub-domains. The first harbors information on site location (geographic coordinates), dating range, dating method, dating quality, the type of site, the nature, context, and coherence of the lithic assemblage(s) in question, faunal preservation as well as the original year of excavation (Table 10). Sub-domain two addresses the quality and reliability of the key sites listed by calculating quality scores (1 to 6) and ranks (1 to 4) based on information provided in the first sub-domain (Supplementary Tables 4, 6). Lower quality scores broadly indicate less reliable archaeological contexts expressed in lower ranks. The third and last sub-domain is a register of processed tool outlines stored in the outline library so that analysts can quickly identify the distribution of available outlines across sites (Supplementary Table 7).

Dataset coding and abbreviations used. Metadata (NACs, macro-regions, expert editors, etc.), tool groups, technological traits, 2D artefact outlines and archaeological site contexts are coded as abbreviations. All abbreviations used in the dataset¹⁹ are listed in the foregoing tables or are explained in the Supplementary Information.

Technical Validation

Data quality and reliability. Quality scores were calculated for each archaeological site context based on the associated metadata provided (Supplementary Table 4). The score was derived by summing all positive entries (1s) for the following data categories: dating quality = reliable, site stratigraphy = stratified, assemblage position = primary/*in-situ*, assemblage coherence = homogeneous, and excavation year = 1980–2000 (Supplementary Table 6). Recorded years of excavation = after 2000 (1s) were double weighted, resulting in a possible score of 6 for the most dependable archaeological contexts. Figure 6a shows the geographic distribution of quality scores across Europe, pointing to north-south and east-west quality gradients with Western and Southern European sites scoring generally higher. This pattern is mainly a function of the availability and inclusion of open-air vs. sheltered sites (caves, rockshelters) in the dataset¹⁹. In addition, it may highlight differential depositional and preservation conditions of stone artefact assemblages on the North European Plain and in high-latitude landscapes, which have been strongly affected by glacial activity, for example across the northeastern extent of the study area. Some of the included lithic data from Eastern and parts of Central Europe may thus be less reliable than their Southern and Western European counterparts and this is in part also a consequence of the varying scope and intensity of Late Glacial research in these regions in recent decades. This situation is also well-reflected in the distribution of average quality scores across macro-regions plotted in Fig. 6b. Lithuania and the macro-region comprised of Austria, Slovakia and Hungary host a notable number of comparatively problematic data contexts and should accordingly be treated with some caution. Interestingly, the average quality scores of all included macro-regions, taken together, slightly deteriorate over time (from TS I to IV), and not the other way around as may be expected *a priori*. This similarly suggests that research history and/or unequal levels of sedimentation may play a role here, although past human behavior may also be implicated, for example due to more mobile forager groups and hence possibly more elusive material signatures in some parts of Northern and Eastern Europe. Note that this quality

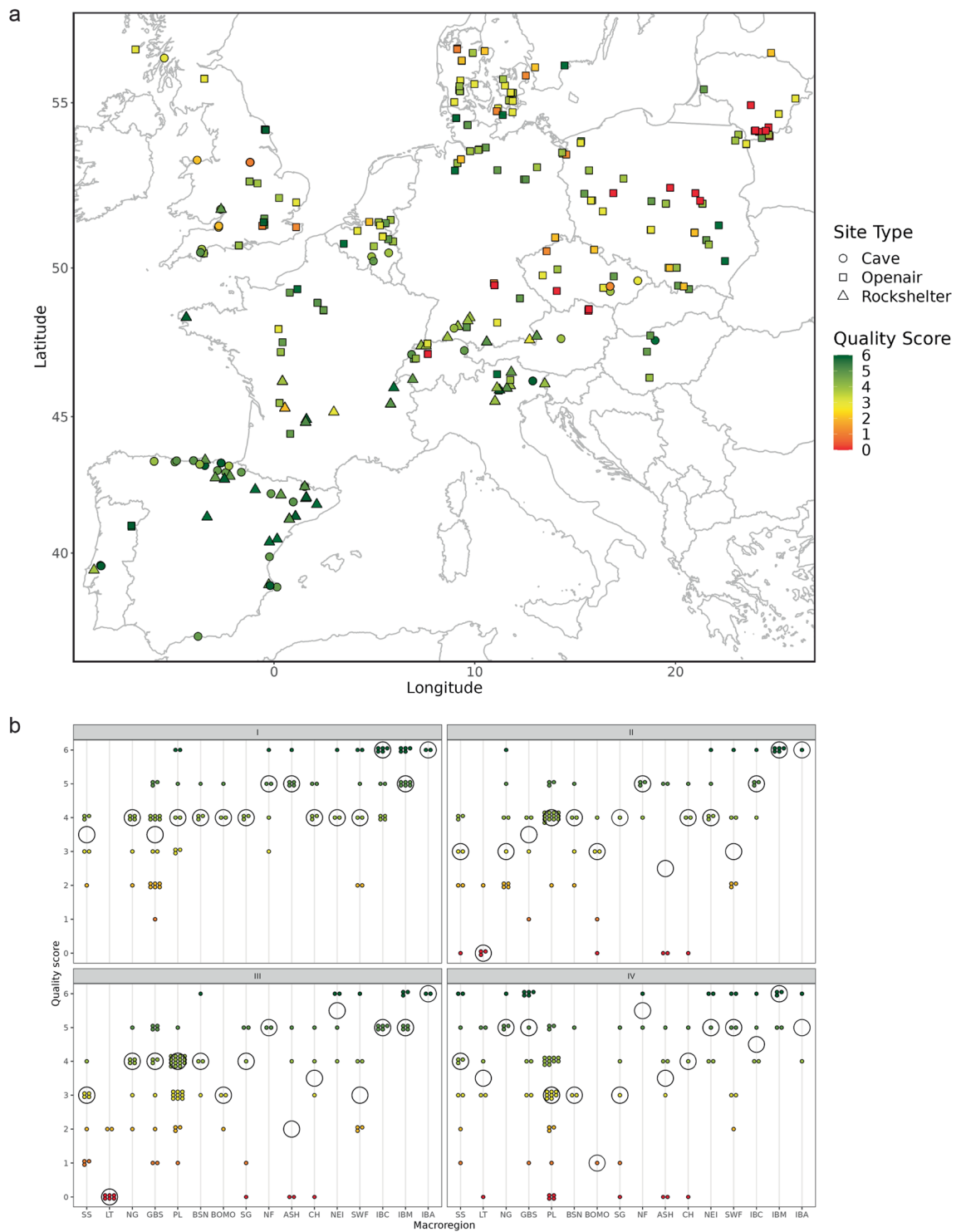


Fig. 6 Data quality. **(a)** Quality scores computed for each archaeological key site (see text for explanation) from 0 = poor to 6 = excellent. **(b)** Detailed break-down of site-specific quality scores in relation to time slices and individual macro-regions (Supplementary Table 4). Black circles indicate median values computed for each macro-region within a given time slice.

assessment does not take into consideration how the lithic data was analyzed and formed, for example whether or not the latest methods and analytical standards were employed.

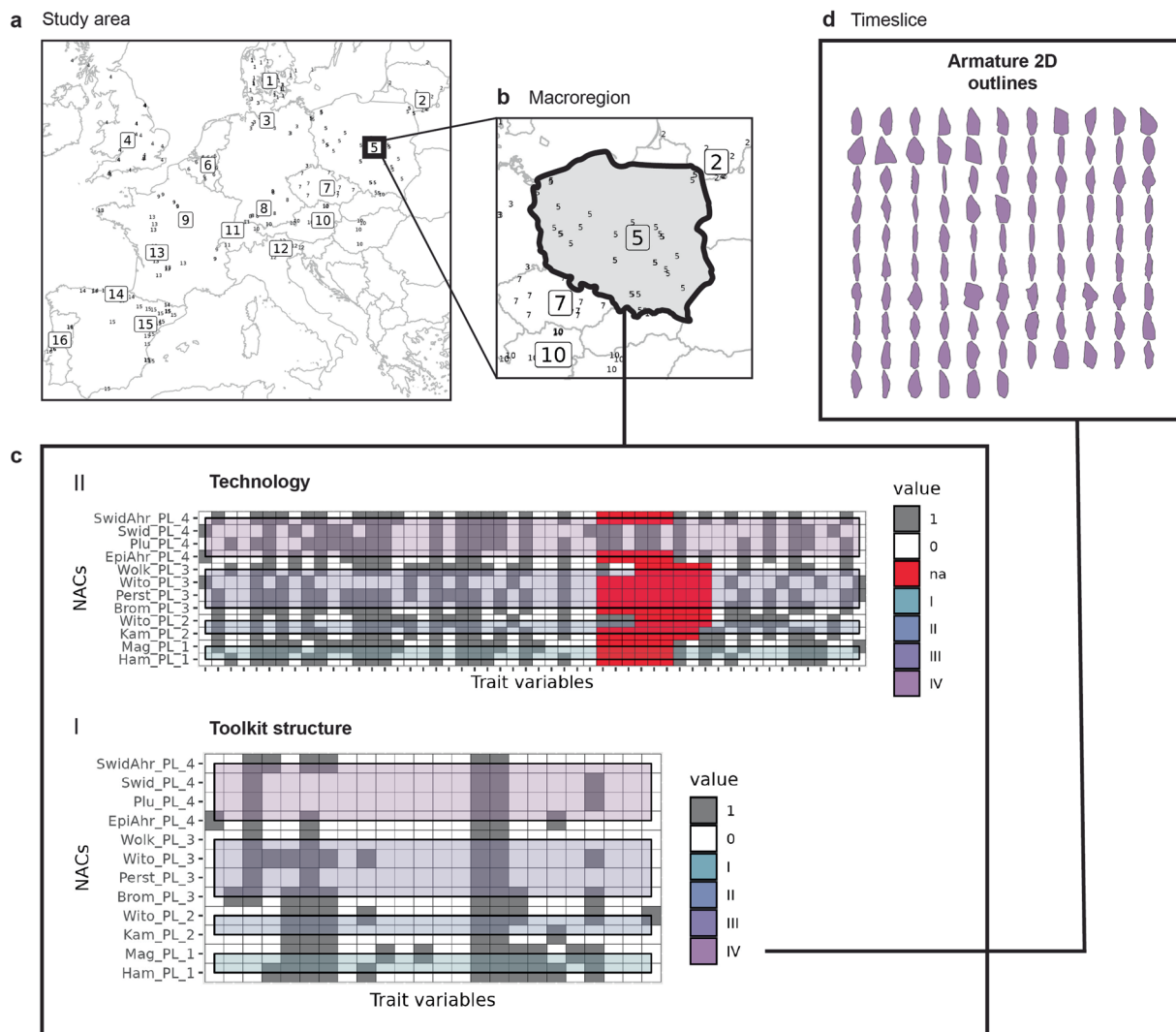


Fig. 7 Overview and exemplification of lithic data structure. **(a)** Pan-European scale of observation with all 16 macro-regions (boxed numbers) and their associated archaeological key sites (small numbers). **(b)** Macro-regional scale of observation (example: Poland); **(c)** Macro-regional data matrix of observations on toolkit composition (*I*) and technology (*II*) coded as presences/absences (na = information not available in the literature) and sorted according to time slices. **(d)** Digitized 2D outlines of lithic armature within a selected time slice (example: TS IV) of the same macro-region. This data structure facilitates the diachronic and synchronic analysis of within and cross-domain lithic variation.

Usage Notes

Data structure and spatiotemporal scale. The dataset¹⁹ records multi-scale observations on lithic technology for each macro-region (Fig. 7). Data on reduction technologies and toolkit composition are provided as NAC-level presence/absence matrices (macro-regional scale) but the resolution of the two data domains differs, as toolkit trait variables refer to highly generalized lithic tool morphotypes while technological traits represent a mix of well-established descriptive variables and more synthetic observations. Variability in the technology domain is thus not directly comparable to variability in the toolkit domain, but we generally expect that technology-toolkit relationships reflect relevant information on technical choices and cultural evolution. In other words, patterns in macro-regional technological data should be related to the structural organization of the corresponding toolkits, but the relationship must not always be strictly co-evolutionary as technical trade-offs may for example promote informal tool-use not documented in the dataset⁴³. Conversely, tools shapes may be more readily copied and tool concepts may thus be more mobile between communities (reflected in *Tools*-based similarity and co-occurrence of rarer forms such as *Zinken*, specific geometric microliths or tanged points), whereas specific ways of working lithic raw material volumes (combinations of *Technology* variables) are often expected to diffuse less easily^{44,45}, but this hypothesis should be empirically explored and different techno-cultural contexts may yield varying dynamics^{46,47}. The offered data structure so enables the comparative investigation of patterned co-variation between technology and toolkit data, and to test whether such relationships differ between regions and time slices. NAC-level data resolved on the macro-regional scale (Fig. 7a,b) can thus

generally be examined diachronically (analytical units: different time slices) or synchronically (analytical units: same time slices across regions).

The outline data adds depth to this structure. On the temporal axis, the artefact outlines are also resolved according to the four time slices, while spatially they are linked to individual archaeological key sites and thus enable the investigation of geographically continuous dynamics of shape change in three different artefact registers (armatures, endscrapers and borers). Again, the relationship between patterns recorded in artefact outline assemblages and NAC-level toolkits should not be arbitrary, but the nature of the interactions is open to characterization. Depending on the spatiotemporal, and possibly technological, context, diversity in armature morphotypes (*Tools*) may be directly linked to global diversity patterns in armature outlines (*Outlines*), or not. This is because *Outline* data should capture more subtle levels of variation but also since the variability within morphotypes (as defined in *Tools*) may vary between technological contexts. Diversity in tool outlines may similarly be expected to be elevated in cases where lithic technology is rooted in a conceptual and operational distinction between lithic blank and tool production, so that tool confection (modification) becomes the key means to install and differentiate tool functions, or, alternatively, reflect the grounding diversity of forms obtained during blank production, or a combination of both. These and cognate hypotheses can empirically be tested using the present dataset¹⁹, and changing structuration principles at the intersection of technology, toolkit structure and artefact shape may in this way be discovered and pinpointed in time and space.

Inter-domain dynamics. The multi-domain and multi-scale structure of the dataset¹⁹ supports the systemic analysis of long-term dynamics of lithic change. By quantifying the amplitude and rate of change in varying data domains, researchers can examine patterns of synchronicity or lack thereof in the evolution of stone artefact technologies on a pan-European scale, and in this way contribute to debates on the temporality of material culture change^{48–52} as well as broader discussions on the evolution of cultural systems^{53,54}. Confronting the dynamics of change across data domains thereby contributes to the identification, quantification, and hence qualification of different tempi and modes of change, possibly revealing that technological choices, toolkit compositions, and artefact shapes are subjected to differential dynamics of long-term technological evolution. Under which conditions such asynchronies in lithic evolution emerged, or whether we should instead regard such patterns as the default mode of material culture change, is an important emerging question and the dataset¹⁹ makes it possible to begin systematically addressing it. Inter-domain analysis and comparison finally provide a new means to assess the general character and the complexity of cultural evolution between ~15,000 and 11,000 years ago. Complexity can for example be understood as a property of domain interaction, and mapping out domain-specific dynamics of change can thus provide new perspectives on questions of linearity and nonlinearity in the millennial-scale evolution of Late Pleistocene/earliest Holocene human forager technologies.

Further remarks and perspectives of data usage. Archaeological data are voluminous, distributed across many language and publication formats. They are rarely available in standardized digital formats. Mindful of the analytical decisions that had to be taken in transforming qualitative observations and legacy images into the dataset¹⁹ presented here^{55,56}, we recommend due diligence in its usage. That said, computational tools are offering new avenues for research within lithic analysis specifically^{57,58}, and for understanding large-scale patterns of cultural evolution⁵⁹. Data availability is a key element in this endeavor^{60,61} and the dataset¹⁹ presented here is the largest of its kind at this moment, assembled through the integration of domain-specific (Late Glacial archaeology and lithic technology) and computational expertise.

By linking original sources to coded data, we offer users the possibility to trace the material of interest back to source for validation. As our technical validation indicates, data availability and quality varies by region, creating an imbalance between Western Europe and the remainder of the study region. Especially in parts of Central and Eastern Europe, data availability is compromised by research historical factors, while taphonomic factors may have led to generally poorer preservation conditions in Northern Europe. In assembling the present dataset¹⁹, our aim has been to facilitate inter-regional analysis. Users seeking to ask questions at smaller geographic scales should be mindful of the regional integration of information within the toolkit and technology domains. This integration precludes comparisons of lithic technological variability between sites and at within-region scales. Furthermore, the 2D outline data derives from legacy sources and different production styles may impact accuracy⁶². Drawings and photos have, however, been shown to provide reliable information for geometric morphometric analysis^{40,63}, lending confidence that these data can be used for comparative inter-regional analyses. Notably, the acquisition of additional outline data for lithic artefacts is straightforward, and we hope that users will complement the dataset¹⁹ at hand with further specimens. Our dataset¹⁹ provides the possibility of quantifying cultural similarity in an information-rich manner. In tandem with the increasing availability of palaeogenomic information, the dataset¹⁹ presented here thereby offers the opportunity to explicitly compare patterns of biological and cultural relatedness during this pivotal epoch in European prehistory, and therefore to contribute to general, discipline-transcending questions on early human biocultural evolution.

Code availability

The full dataset¹⁹ and the code used in this paper are freely available at Zenodo: <https://doi.org/10.5281/zenodo.7940337>. All data, analyses, and visualizations presented in this paper were prepared in R 4.2.2 under Ubuntu 18.04.5 LTS (64-bit) using the following R packages: *data.table* ($\geq 1.14.8$), *dplyr* ($\geq 1.1.2$), *forcats* ($\geq 1.0.0$), *ggforce* ($\geq 0.4.1$), *ggplot2* ($\geq 3.4.2$), *ggpointgrid* ($\geq 1.2.0$), *ggridges* ($\geq 0.5.4$), *magrittr* ($\geq 2.0.3$), *Momocs* ($\geq 1.4.0$), *outlineR* ($\geq 0.1.0$), *raster* ($\geq 3.6-20$), *readr* ($\geq 2.1.4$), *remotes* ($\geq 2.4.2$), *rgeos* ($\geq 0.6-2$), *rworldmap* ($\geq 1.3-6$), *sp* ($\geq 1.6-0$).

Received: 2 June 2023; Accepted: 22 August 2023;

Published online: 07 September 2023

References

- Valentin, B. Jalons pour une paléohistoire des derniers chasseurs: XI^e-VI^e millénaire avant J.-C. (Publications de la Sorbonne, 2008).
- Marchand, G. *Préhistoire atlantique: fonctionnement et évolution des sociétés, du Paléolithique au Néolithique*. (Éditions Errance, 2014).
- Riede, F. *Splendid isolation: the eruption of the Laacher See Volcano and Southern Scandinavian Late Glacial hunter-gatherers*. (Aarhus University Press, 2017).
- Crombé, P. Mesolithic projectile variability along the southern North Sea basin (NW Europe): Hunter-gatherer responses to repeated climate change at the beginning of the Holocene. *PLOS ONE* **14**, e0219094 (2019).
- Grimm, S. B. *Resilience and reorganisation of social systems during the Weichselian lateglacial in North-West Europe: an evaluation of the archaeological, climatic, and environmental record*. (Verlag des Römisch-Germanischen Zentralmuseums; Schnell & Steiner, 2019).
- Gamble, C., Davies, W., Pettitt, P. & Richards, M. Climate change and evolving human diversity in Europe during the last glacial. *Phil. Trans. R. Soc. Lond. B* **359**, 243–254 (2004).
- Fernández-López de Pablo, J. *et al.* Palaeodemographic modelling supports a population bottleneck during the Pleistocene-Holocene transition in Iberia. *Nat Commun* **10**, 1872 (2019).
- Posth, C. *et al.* Pleistocene Mitochondrial Genomes Suggest a Single Major Dispersal of Non-Africans and a Late Glacial Population Turnover in Europe. *Current Biology* **26**, 827–833 (2016).
- Charlton, S. *et al.* Dual ancestries and ecologies of the Late Glacial Palaeolithic in Britain. *Nat Ecol Evol* **6**, 1658–1668 (2022).
- Schmidt, I., Gehlen, B. & Zimmermann, A. Population estimates for the Final Palaeolithic (14,000 to 11,600 years cal. BP) of Europe – challenging evidence and methodological limitations. in *On the Move - Mobility of people, objects and ideas during the European Upper Palaeolithic* 221–237 (Société préhistorique française, 2021).
- Hayden B. Transegularian Hunter/Gatherers. in *The Power of Feasts: From Prehistory to the Present* (ed. Hayden, B.) 47–108 (Cambridge University Press, 2014).
- Arnold, J. E. *et al.* Entrenched Disbelief: Complex Hunter-Gatherers and the Case for Inclusive Cultural Evolutionary Thinking. *J Archaeol Method Theory* **23**, 448–499 (2016).
- Riede, F. The Loss and Re-Introduction of Bow-and-Arrow Technology: A Case Study from the Northern European Late Paleolithic. *Lithic Technology* **34**, 27–45 (2009).
- Ivanovaite, L., Serwatka, K., Hoggard, C. S., Sauer, F. & Riede, F. All these Fantastic Cultures? Research History and Regionalization in the Late Palaeolithic Tanged Point Cultures of Eastern Europe. *Eur. j. archaeol* **23**, 162–185 (2020).
- Straus, L. G. The Archaeology of the Pleistocene—Holocene Transition in Southwest Europe. in *Humans at the End of the Ice Age: The Archaeology of the Pleistocene—Holocene Transition* (eds. Straus, L. G., Eriksen, B. V., Erlandson, J. M. & Yesner, D. R.) 83–99 (Springer US, 1996).
- Manninen, M. A. *et al.* First encounters in the north: cultural diversity and gene flow in Early Mesolithic Scandinavia. *Antiquity* **95**, 310–328 (2021).
- Kuhn, S. L. *The Evolution of Paleolithic Technologies*. (Routledge, 2020).
- Liu, C. & Stout, D. Inferring cultural reproduction from lithic data: A critical review. *Evolutionary Anthropology: Issues, News, and Reviews* **32**, 83–99 (2022).
- Hussain, S. T. *et al.* Research compendium for 'A pan-European dataset revealing variability in lithic technology, toolkits, and artefact shapes ~15–11 kya'. *Zenodo* <https://doi.org/10.5281/ZENODO.7940337> (2023).
- Riede, F. *et al.* Cultural taxonomies in the Paleolithic—Old questions, novel perspectives. *Evol Anthropol* **29**, 49–52 (2020).
- Hussain, S. T. *et al.* Towards a collaborative meta-analysis of the Final Palaeolithic/earliest Mesolithic in Europe. Report on the 2nd CLIOARCH Workshop, 26th–27th November 2020. *Archäologische Informationen* 415–420, <https://doi.org/10.11588/AI.2020.1.81428> (2021).
- Chatters, J. C., Kuijt, I. & Prentiss, A. M. *Macroevolution in human prehistory: evolutionary theory and processual archaeology*. (Springer, 2009).
- Perreault, C. *The Quality of the Archaeological Record*. (The University of Chicago Press, 2019).
- Reynolds, N. & Riede, F. House of cards: cultural taxonomy and the study of the European Upper Palaeolithic. *Antiquity* **93**, 1350–1358 (2019).
- Altschul, J. H. *et al.* Fostering synthesis in archaeology to advance science and benefit society. *Proc. Natl. Acad. Sci. USA* **114**, 10999–11002 (2017).
- Hofmann, K. P. *et al.* Ding-Editionen. Vom archäologischen (Be-)Fund übers Corpus ins Netz. *iDAI.publications* **2019**, 1–12 (2019).
- Diachenko, A. Archaeology and the nation state. The case of eastern Europe. *Arch. Dial.* **23**, 3–10 (2016).
- Hakonen, A. Communities Beyond Society: Divergence of Local Prehistories on the Bothnian Arc, Northern Europe. *Open Archaeology* **7**, 211–230 (2021).
- McAnany, P. A. & Rowe, S. M. Re-visiting the field: Collaborative archaeology as paradigm shift. *Journal of Field Archaeology* **40**, 499–507 (2015).
- Wylie, A. A plurality of pluralisms: Collaborative practice in archaeology. in *Objectivity in Science* 189–210 (Springer, 2015).
- Galeazzi, F. & Richards-Rissetto, H. Editorial Introduction: Web-based Archaeology and Collaborative Research. *Journal of Field Archaeology* **43**, 1–8 (2018).
- Van Andel, T. & Davies, W. *Neanderthals and modern humans in the European landscape during the last glaciation: archaeological results of the Stage 3 Project*. (McDonald Institute for Archaeological Research; Distributed by Oxbow Books, 2003).
- Kandel, A. W. *et al.* The ROCEEH Out of Africa Database (ROAD): A large-scale research database serves as an indispensable tool for human evolutionary studies. *PLOS ONE* **18**, e0289513 (2023).
- Djindjian, F., Kozłowski, J. K., Otte, M., Delporte, H. & Kozłowski, J. *Le paléolithique supérieur en Europe*. (Armand Colin, 1999).
- Kozłowski, J. K. Paléolithique supérieur et Mésolithique en Méditerranée: cadre culturel. *L'Anthropologie* **109**, 520–540 (2005).
- Kozłowski, S. K. Mapping the Central/East European terminal palaeolithic/earliest mesolithic. *Archaeologia Baltica* **7**, 29–35 (2006).
- Reynolds, N. Threading the Weft, Testing the Warp: Population Concepts and the European Upper Paleolithic Chronocultural Framework. in *Culture History and Convergent Evolution: Can We Detect Populations in Prehistory?* (ed. Groucutt, H. S.) 187–212 (Springer International Publishing, 2020).
- Bonhomme, V., Picq, S., Gaucherel, C. & Claude, J. Momocs: Outline Analysis Using R. *J. Stat. Soft.* **56** (2014).
- Matzig, D. N. *outlineR: Artefact Processing and Extraction Protocol v1*. <https://www.protocols.io/view/outliner-arteifact-processing-and-extraction-protoc-bygaptse>, <https://doi.org/10.17504/protocols.io.bygaptse> (2021).
- Matzig, D. N., Hussain, S. T. & Riede, F. Design Space Constraints and the Cultural Taxonomy of European Final Palaeolithic Large Tanged Points: A Comparison of Typological, Landmark-Based and Whole-Outline Geometric Morphometric Approaches. *J Paleol Arch* **4**, 27 (2021).
- Rohlf, F. The tps series of software. *Hystrix, the Italian Journal of Mammalogy* **26** (2015).
- Matzig, D. N. *outlineR: An R package to derive outline shapes from (multiple) artefacts on JPEG images*. *Zenodo* <https://doi.org/10.5281/ZENODO.4527470> (2021).
- Guéret, C. & Jacquier, J. La transition Pléistocène-Holocène dans le Nord-Ouest de l'Europe vue par le prisme de la tracéologie lithique: essai de synthèse. in *Préhistoire de l'Europe du Nord-Ouest: mobilité, climats et identités culturelles, XXVIII^e Congrès Préhistorique d'Amiens, 30 Mai-4 Juin 2016. Volume 2: Paléolithique supérieur ancien, Paléolithique final - Mésolithique* 363–379 (Société Préhistorique Française, 2019).

44. Tostevin, G. B. *Seeing lithics: a middle-range theory for testing for cultural transmission in the pleistocene*. (Oakville, CT: Oxbow Books, 2013).
45. Roussel, M., Soressi, M. & Hublin, J.-J. The Châtelperronian conundrum: Blade and bladelet lithic technologies from Quinçay, France. *Journal of Human Evolution* **95**, 13–32 (2016).
46. Laurent, K. *The prehistoric apprentice: investigating apprenticeship, know-how and expertise in prehistoric technologies = L'apprenti préhistorique: appréhender l'apprentissage, les savoir-faire et l'expertise à travers les productions techniques des sociétés préhistoriques*. (The Czech Academy of Sciences, Institute of Archaeology, Brno, 2018).
47. Valentin, B., Weber, M.-J. & Bodu, P. Initialisation and progression of the core reduction process at Donnemarie-Dontilly (Seine-et-Marne, France), site of the Belloisian tradition: New interpretative key for comparisons with contemporaneous industries and 'Federmesser-Gruppen' assemblages. *Bulletin de la Société préhistorique française* **111**, 659–678 (2014).
48. Naudinot, N. *et al.* Divergence in the evolution of Paleolithic symbolic and technological systems: The shining bull and engraved tablets of Rocher de l'Impératrice. *PLoS ONE* **12**, e0204464 (2017).
49. Perlès, C. Temps of Change: When Soloists don't play Together. Arrhythmia in 'Continuous' Change. *J Archaeol Method Theory* **20**, 281–299 (2013).
50. Kuhn, S. L. Questions of Complexity and Scale in Explanations for Cultural Transitions in the Pleistocene: A Case Study from the Early Upper Paleolithic. *J Archaeol Method Theory* **20**, 194–211 (2013).
51. Maier, A. *et al.* Cultural evolution and environmental change in Central Europe between 40 and 15 ka. *Quaternary International* **581–582**, 225–240 (2021).
52. Perlès, C. Cultural Implications of Uniformity in Ornament Assemblages: Paleolithic and Mesolithic Ornaments From Franchthi Cave, Greece. *PaleoAnthropology* 196–207, <https://doi.org/10.4207/PA.2019.ART131> (2019).
53. Buskell, A., Enquist, M. & Jansson, F. A systems approach to cultural evolution. *Palgrave Commun* **5**, 1–15 (2019).
54. Jansson, F., Buskell, A. & Enquist, M. Cultural Systems. *OSF* <https://doi.org/10.31235/osf.io/6szvj> (2022).
55. Hofmann, K. P. Dingidentitäten und Objekttransformationen. Einige Überlegungen zur Edition von archäologischen Funden. in *Objektpistemologien. Zum Verhältnis von Dingen Und Wissen 179–215* (Edition Topoi, 2018).
56. Katsianis, M., Kalayci, T. & Sarris, A. Bridging Digital Approaches and Legacy in. *Archaeology. Digital* **2**, 538–545 (2022).
57. Shott, M. Digitizing archaeology: a subtle revolution in analysis. *World Archaeology* **46**, 1–9 (2014).
58. Grosman, L. Reaching the Point of No Return: The Computational Revolution in Archaeology. *Annual Review of Anthropology* **45**, 129–145 (2016).
59. Riede, F. *et al.* A meta-analysis of Final Palaeolithic/earliest Mesolithic cultural taxonomy and evolution in Europe. Preprint at Zenodo, <https://doi.org/10.5281/ZENODO.8195588> (2023).
60. Marwick, B. & Birch, S. E. P. A Standard for the Scholarly Citation of Archaeological Data as an Incentive to Data Sharing. *Adv. archaeol. pract.* **6**, 125–143 (2018).
61. Marwick, B. *et al.* Open science in archaeology. *SAA Archaeological Record* **17**, 8–14 (2017).
62. Lopes, D. M. *Drawing in a Social Science. vol. 1* (Oxford University Press, 2018).
63. Timbrell, L. *et al.* Testing inter-observer error under a collaborative research framework for studying lithic shape variability. *Archaeol Anthropol Sci* **14**, 209 (2022).

Acknowledgements

We would like to thank all student assistants who supported this project at various stages of progression. STH, FR and DM acknowledge funding through the ERC Consolidator Grant project CLIOARCH from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 817564).

Author contributions

Conceptualization: S.T.H., F.R.; Methodology: S.T.H., D.M., F.R.; Software: D.M., S.T.H.; Validation: S.T.H., D.M., F.R., M.B., P.C., F.F., D.G., T.H., M.L., J.F.L.d.P., L.M., W.M., M.M., N.N., C.P., T.R., D.S., H.V.; Formal analysis: D.M., S.T.H.; Investigation: S.T.H., D.M., F.R., M.B., P.C., F.F., D.G., T.H., M.L., J.F.L.d.P., L.M., W.M., M.M., N.N., C.P., T.R., D.S., H.V.; Resources: S.T.H., D.M., F.R., M.B., P.C., F.F., D.G., T.H., M.L., J.F.L.d.P., L.M., W.M., M.M., N.N., C.P., T.R., D.S., H.V.; Data curation: S.T.H., D.M.; Writing – Original Draft: S.T.H., D.M., F.R., M.B., P.C., F.F., D.G., T.H., M.L., J.F.L.d.P., L.M., W.M., M.M., N.N., C.P., T.R., D.S., H.V.; Writing – Review & Editing: S.T.H., D.M., F.R., M.B., P.C., F.F., D.G., T.H., M.L., J.F.L.d.P., L.M., W.M., M.M., N.N., C.P., T.R., D.S., H.V.; Visualization: D.M., S.T.H.; Supervision: S.T.H., F.R.; Project administration: S.T.H.; Funding acquisition: F.R.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-023-02500-9>.

Correspondence and requests for materials should be addressed to S.T.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.