

Supplementary Information of Hussain *et al.* A pan-European dataset revealing variability in lithic technology, toolkits, and artefact shapes ~15-11 kya. *Sci. Data* (2023).

SI Part 1

Regional and site-level units of analysis

Supplementary Table 1. Overview of macro-regional units of analysis and their corresponding IDs and regional expert IDs (n=16).

Macroregional Unit	ID/Macro_region_code	Expert ID
Southern Scandinavia	SS	FR
Lithuania	LT	TR
Northern Germany	NG	DG
Britain	GBS	WM
Poland	PL	DS
Belgium and Southern Netherlands	BSN	HV, PC
Bohemia and Moravia	BOMO	MM
Southern Germany	SG	TH
Northern France	NF	LM, MB
Austria, Slovakia and Hungary	ASH	CP
Switzerland	CH (or SW)	TH
Northern/North-eastern Italy	NEI	FF
Western France	SWF	ML, NN
Cantabrian Spain	IBC	JFLdP
Mediterranean Iberia	IBM	JFLdP
Atlantic Iberia	IBA	JFLdP

Supplementary Table 2. Overview of total number of recorded key sites per macro-region (a full list of sites can be extracted from the original dataset published with this paper). Note that these numbers are not directly comparable since layer or concentration distinctions have sometimes been recorded as separate occurrence even when the larger site-context is identical, and sometimes not depending on the nature of the context: macro-regions marked with * involve site entries resolved on the level of excavation context or concentration; the actual number of archaeological sites is thus lower.

Macroregional Unit	Number of key sites
Southern Scandinavia	33
Lithuania	17
Northern Germany	19
Britain	40
Poland	83*

Belgium and Southern Netherlands	16
Bohemia and Moravia	12
Southern Germany	10
Northern France	9
Austria, Slovakia and Hungary	13
Switzerland	11
Northern/North-eastern Italy	13
(South)Western France	18
Cantabrian Spain	19
Mediterranean Iberia	30
Atlantic Iberia	7
<i>total</i>	350

Supplementary Table 3. Tabulation of recorded named archaeological cultures (NACs) ordered according to macro-region and arranged in equidistant time slices (TS I to IV).

<i>Time- slice →</i>	<i>TS I</i>	<i>TS II</i>	<i>TS III</i>	<i>TS IV</i>
<i>Macro- region ↓</i>				
<i>Southern Scandina via (SS)</i>	Havelte	Federmesser, Brommean	Brommean	Ahrensburgian, Maglemose
<i>Lithuania (LT)</i>		Hamburgian, Federmesser	Federmesser, Brommean, Ahrensburgian	Ahrensburgian
<i>Northern Germany (NG)</i>	Hamburgian, Federmesser	Havelte, Federmesser, Brommean	Federmesser, Ahrensburgian	Maglemose
<i>Britain (GBS)</i>	Creswellian/Late Magdalenian, Backed Blade or Hengistbury- type CBP, Havelte	Backed Blade or Hengistbury-type CBP, Havelte	Allerød CBP	Long Blade Industries/Epi- Ahrensburgian (FBBT), Early Mesolithic

<i>Poland (PL)</i>	Magdalenian, Hamburgian	Kamienna variant, Witowian	Witowian, Brommean, Perstunian, Wolkuszan	Wolkuszan, Epi- Ahrensburgian, Swiderian, Swiderian- Ahrensburgian, Pludian
<i>Belgium and Southern Netherlands (BSN)</i>	Magdalenian	Federmesser	(Epi-)Ahrensburgian	Epi-Ahrensburgian
<i>Bohemia and Moravia (BOMO)</i>	Magdalenian	Magdalenian, Epimagdalenian, ABP Federmesser, ABP Tishnovian	ABP Tishnovian	Mesolithic
<i>Southern Germany (SG)</i>	Late/Transitional Magdalenian	Late Palaeolithic	Late Palaeolithic, Atzenhofer group	Late Palaeolithic, Atzenhofer group, Beuronian A
<i>Northern France (NF)</i>	Upper Magdalenian, Early Azilian	Early Azilian, Late Azilian	Late Azilian	Belloisian
<i>Austria, Slovakia and Hungary (ASH)</i>	Late Magdalenian, Late Epigravettian	Aziloid Alps, Aziloid Tradition Eastern Central Europe	Aziloid Alps, Aziloid Tradition Eastern Central Europe, Late Palaeolithic	Aziloid Alps, Aziloid Tradition Eastern Central Europe, Late Palaeolithic, Recent Epigravettian
<i>Switzer- land (CH)</i>	Late Magdalenian/Techno- mplex E, Older Azilian	Older Azilian, Fazies Fürsteiner	Older Azilian, Fazies Fürsteiner, Younger Azilian	Fazies Fürsteiner, Early Mesolithic I
<i>North- eastern Italy (NEI)</i>	Late Epigravettian (Phase 2)	Late Epigravettian (Phase 2), Late Epigravettian (Phase 3)	Late Epigravettian (Phase 3)	Late Epigravettian- Early Sauveterrian, Early Sauveterrian
<i>Western France (SWF)</i>	Upper Magdalenian, Early Azilian	Early Azilian, Late Azilian	Late Azilian, Early Laborian	Early Laborian, Late Laborian
<i>Cantabria n Spain (IBC)</i>	Final Magdalenian	Azilian	Azilian	Azilian

<i>Mediterranean Iberia (IBM)</i>	Upper Magdalenian, Final Magdalenian, Azilian	Azilian, Epipalaeolithic/Epimagdalenian	Azilian, Epipalaeolithic/Epimagdalenian, Sauveterrian/Sauveterroid Epipalaeolithic	Sauveterrian/Sauveterroid Epipalaeolithic
<i>Atlantic Iberia (IBA)</i>	Upper Magdalenian	Final Magdalenian	Azilian	Azilian

SI Part 2

Site quality

Data quality is assessed on the level of individual archaeological sites and presented as a ‘Quality_Score’ (see main text and **Supplementary Table 6** below).

Supplementary Table 4. Quality score per region per time-slice.

<i>Macroregion-code and number of observations</i>	Discrete time- slice	Quality score (median)
<i>ASH (n=5)</i>	I	5
<i>BSN (n=4)</i>	I	4
<i>BOMO (n=3)</i>	I	4
<i>NEI (n=3)</i>	I	4
<i>NG (n=6)</i>	I	4
<i>SG (n=3)</i>	I	4
<i>CH (n=5)</i>	I	4
<i>SWF (n=7)</i>	I	4
<i>PL (n=8)</i>	I	4
<i>IBC (n=7)</i>	I	6
<i>IBM (n=10)</i>	I	5
<i>IBA (n=2)</i>	I	6
<i>SS (n=6)</i>	I	3.5
<i>NF (n=5)</i>	I	5
<i>GBS (n=18)</i>	I	3.5
<i>ASH (n=4)</i>	II	2.5
<i>BSN (n=5)</i>	II	4

<i>BOMO</i> (<i>n</i> =8)	II	3
<i>LT</i> (<i>n</i> =4)	II	0
<i>NEI</i> (<i>n</i> =5)	II	4
<i>NG</i> (<i>n</i> =9)	II	3
<i>CH</i> (<i>n</i> =4)	II	4
<i>SG</i> (<i>n</i> =1)	II	4
<i>SWF</i> (<i>n</i> =6)	II	3
<i>PL</i> (<i>n</i> =17)	II	4
<i>IBC</i> (<i>n</i> =6)	II	5
<i>IBM</i> (<i>n</i> =4)	II	6
<i>IBA</i> (<i>n</i> =1)	II	6
<i>SS</i> (<i>n</i> =8)	II	3
<i>NF</i> (<i>n</i> =4)	II	5
<i>GBS</i> (<i>n</i> =4)	II	3.5
<i>SS</i> (<i>n</i> =9)	III	3
<i>ASH</i> (<i>n</i> =4)	III	2
<i>BSN</i> (<i>n</i> =4)	III	4
<i>BOMO</i> (<i>n</i> =3)	III	3
<i>LT</i> (<i>n</i> =8)	III	0
<i>NEI</i> (<i>n</i> =4)	III	5.5
<i>NG</i> (<i>n</i> =7)	III	4
<i>CH</i> (<i>n</i> =4)	III	3.5
<i>SG</i> (<i>n</i> =5)	III	4

<i>SWF</i> (<i>n</i> =6)	III	3
<i>PL</i> (<i>n</i> =29)	III	4
<i>IBC</i> (<i>n</i> =6)	III	5
<i>IBM</i> (<i>n</i> =6)	III	5
<i>IBA</i> (<i>n</i> =2)	III	6
<i>NF</i> (<i>n</i> =2)	III	5
<i>GBS</i> (<i>n</i> =11)	III	4
<i>ASH</i> (<i>n</i> =6)	IV	3.5
<i>BSN</i> (<i>n</i> =3)	IV	3
<i>BOMO</i> (<i>n</i> =1)	IV	1
<i>LT</i> (<i>n</i> =6)	IV	3.5
<i>NEI</i> (<i>n</i> =5)	IV	5
<i>NG</i> (<i>n</i> =4)	IV	5
<i>CH</i> (<i>n</i> =3)	IV	4
<i>SG</i> (<i>n</i> =5)	IV	3
<i>SWF</i> (<i>n</i> =7)	IV	5
<i>PL</i> (<i>n</i> =29)	IV	3
<i>IBC</i> (<i>n</i> =4)	IV	4.5
<i>IBM</i> (<i>n</i> =5)	IV	6
<i>IBA</i> (<i>n</i> =2)	IV	5
<i>SS</i> (<i>n</i> =10)	IV	4
<i>NF</i> (<i>n</i> =2)	IV	5.5
<i>GBS</i> (<i>n</i> =11)	IV	5

The 1511NAC Dataset

Metadata

Supplementary Table 5. General abbreviations (provided as numbers and as text entries).

Code	Description
KeySite	Name of selected key site
Level/layer/ concentration	Context of selected lithic assemblage from key site
Site_ID	Unique site identifier
TaxUnit	Name of associated archaeological taxonomic unit
TaxUnit_unique	Unique TaxUnit identifier
Timeslice	Time-slice attribution (“1”, “2”, “3” or “4”); if a given TaxUnit spans multiple Timeslices, these are listed as comma separated numbers (e.g. “1,2”)
Expert_editor	Code for data provider
Macro_region	Macro-region for which lithic data was collected (frame of reference for TaxUnit determination)
Macro_region_code	Unique identifier of macro-regions

Site/context data module

In contrast to the previous two data modules, the site/context module compiles categorical and non-categorical information on the level of individual key sites chosen to represent larger taxonomic archaeological units. Contextual information is either given as numbers, text, a combination of the two, or as predefined categorical selections (drop-down).

Supplementary Table 6. Calculated site-quality scores

Code	Description
Quality_Score	Total sum of following scores: ‘1’ for Dating_qual = ‘Reliable’, ‘1’ for Site_strat = ‘Stratified’, ‘1’ for Ass_pos = ‘Primary/insitu’, ‘1’ for Ass_coh = ‘Homogeneous’ + ‘1’ for Site_excav = ‘1980-2000’ or ‘2’ for Site_excav = ‘After 2000’ (total possible score = ‘6’)
Quality_Rank	‘1’ if Quality_Score = 5-6, ‘2’ if Quality_Score = 3-4, ‘3’ if Quality_Score = 1-2 or ‘4’ if Quality_Score = 0

Supplementary Table 7. Register of processed and recorded tool outlines (provided as ‘1’ for available and ‘0’ for not available).

Code	Description
Outline_AR_TS1	Armature outlines for Timeslice 1
Outline_AR_TS2	Armature outlines for Timeslice 2
Outline_AR_TS3	Armature outlines for Timeslice 3
Outline_AR_TS4	Armature outlines for Timeslice 4
Outline_ES_TS1	Endscraper outlines for Timeslice 1
Outline_ES_TS2	Endscraper outlines for Timeslice 2
Outline_ES_TS3	Endscraper outlines for Timeslice 3
Outline_ES_TS4	Endscraper outlines for Timeslice 4
Outline_BR_TS1	Borer/perforator outlines for Timeslice 1
Outline_BR_TS2	Borer/perforator outlines for Timeslice 2
Outline_BR_TS3	Borer/perforator outlines for Timeslice 3
Outline_BR_TS4	Borer/perforator outlines for Timeslice 4

Outline library

Data structure and recording scheme

The outline dataset consists of individual image files, which have been pre-processed into binary artefact images, and which can directly be fed into an outline extraction and shape-description workflow (see detailed description of extraction and data preparation methods in the main paper). Original artefact imagery does not necessarily contain only a single lithic artefact but sometimes gather multiple objects. This being said, the automated extraction protocol used generates individual files for each extracted object-outline (Matzig 2021a, b). It is these individual object-outline image files (images as data), which are stored as JPEG files in the dataset. Fully pre-processed artefact outlines are also made available and are stored separately in the dataset.

Individual outlines are grouped and coded according the three main tool categories structuring the outline sample (Armature = ‘AR’; Endscraper = ‘ES’; Borer/perforator = ‘BR’). Individual files/images are linked by naming conventions (see below) reflecting their associated time slice, taxonomic unit, site, literature source, etc.

Note that for the purpose of our research goals/interests, these broader categories (AR, ES, BR) were not further disintegrated into individual lithic types, morphotypes or techno-types that are commonly distinguished in the literature.

Furthermore, only complete or reconstructed pieces have been included in the outline dataset for lithic armatures (AR) and endscrapers (ES), but not for borers/perforators (BR) because in the latter case the analytical aim of the data collection was not to analyse whole-outline variability but the shape configuration of borer working-ends (open outlines).

This is the general structure of the employed naming convention:

- **Timeslice_TaxonomicUnit_Macroregion_KeySite_LiteratureSource_ToolCat_Axiality_SubID1_SubID2**

Supplementary Table 8. Break-down of the different parts of this label structure and how they are expressed in the naming of individual image files.

Timeslice	‘TS1’, ‘TS2’, ‘TS3’ or ‘TS4’ for Timeslice 1, 2, 3 or 4; image data that relates to multiple Timeslices is coded in the following way: Timeslices 1-2 = ‘TS12’, Timeslices 3-4 = ‘TS34’ or Timeslices 2-4 = ‘TS234’
TaxonomicUnit	Identifier taken from ‘TaxUnit_unique’ (see Metadata)
Macroregion	Identifier taken from ‘Macro_region_code’ (see Metadata)
KeySite	Identifier taken from ‘Site_ID’ (see Metadata)
LiteratureSource	Short literature reference without spaces, e.g., ‘Mayer2000’; in case of multiple authors, e.g., ‘WebsterHut1998’ or ‘Greenwoodetal2003’
ToolCat	Code for overarching/main tool categories mentioned above: ‘AR’ for Armature, ‘ES’ for Endscrapers, and ‘BR’ for Borers/perforators (including <i>Zinken</i>)

Axiality	Axiality of points/orientation of tool axis in relation to blank axis: ‘d’ = distal point (point axis aligns with blank axis), ‘p’ = proximal point (point axis inversed in relation to blank axis), or ‘na’ = not applicable/not determinable
SubID1	Sub-id assigned to image files in numerical order that share all previous information
SubID2	Sub-id assigned to individualized object-outlines in numerical after automated outline extraction

References

Matzig, D.N. 2021a. outlineR: An R package to derive outline shapes from (multiple) artefacts on JPEG images. *Zenodo*. <https://doi.org/10.5281/ZENODO.4527469>.

Matzig, D.N. 2021b. *outlineR: Artefact Processing and Extraction Protocol v1*.

DOI:dx.doi.org/10.17504/protocols.io.bygaptse (external link: https://www.protocols.io/view/outliner-artefact-processing-and-extraction-protoc-rm7vz3pp8gx1/v1_09/22/2021).