

# SCIENTIFIC REPORTS



OPEN

## Reconciling material cultures in archaeology with genetic data: The nomenclature of clusters emerging from archaeogenomic analysis

Stefanie Eisenmann<sup>1,2</sup>, Eszter Bánffy<sup>3</sup>, Peter van Dommelen<sup>4</sup>, Kerstin P. Hofmann<sup>3</sup>, Joseph Maran<sup>5</sup>, Iosif Lazaridis<sup>2,6,7</sup>, Alissa Mittnik<sup>1,2,6</sup>, Michael McCormick<sup>2,8,9</sup>, Johannes Krause<sup>1,2,10</sup>, David Reich<sup>2,6,7,9,11</sup> & Philipp W. Stockhammer<sup>1,2,12</sup>

Genome-wide ancient DNA analysis of skeletons retrieved from archaeological excavations has provided a powerful new tool for the investigation of past populations and migrations. An important objective for the coming years is to properly integrate ancient genomics into archaeological research. This article aims to contribute to developing a better understanding and cooperation between the two disciplines and beyond. It focuses on the question of how best to name clusters encountered when analysing the genetic makeup of past human populations. Recent studies have frequently borrowed archaeological cultural designations to name these genetic groups, while neglecting the historically problematic nature of the concept of cultures in archaeology. After reviewing current practices in naming genetic clusters, we introduce three possible nomenclature systems ('numeric system', 'mixed system (a)', 'geographic-temporal system') along with their advantages and challenges.

Recent methodological advances including the advent of second generation short read sequencing technologies, the application of targeted hybridisation capture, and the recognition of petrous bones as rich sources for preservation of DNA, have transformed ancient DNA analysis into a revolutionary new tool for investigating the past<sup>1–4</sup>. The exponential increase in the publication of ancient genomes, however, has not been matched by the development of a theoretical framework for the discussion of ancient DNA results and their contextualisation within the fields of history and archaeology<sup>5</sup>. A particularly striking instance of this is given by two ancient DNA papers published in 2015 by Haak *et al.* and Allentoft *et al.* that detected fundamental changes in the central European gene pool during the 3<sup>rd</sup> millennium BCE due to a massive gene influx ultimately deriving from the Pontic steppe region<sup>6,7</sup>. They revived the discussion of large-scale migrations in prehistory, an idea that had been substantially dismissed in archaeology since the 1960s<sup>8,9</sup>. The genetic evidence for large-scale movements of people became undeniable in light of the DNA data, and so the question was no longer about whether ancient DNA analysis can be trusted, but how the results should be interpreted and presented. For example Furrholt 2016, Vander Linden 2016, and Heyd 2017 all accepted the genetic findings, but expressed concern that the studies did not sufficiently deal with the complexities of migrations in that they summarised their findings with simplified migration models involving groups of people (populations) moving

<sup>1</sup>Max Planck Institute for the Science of Human History, Department of Archaeogenetics, Jena, 07745, Germany.

<sup>2</sup>Max Planck—Harvard Research Center for the Archaeoscience of the Ancient Mediterranean (MHAAM), Jena, Germany & Cambridge, MA, USA. <sup>3</sup>Römisch-Germanische Kommission, Deutsches Archäologisches Institut, Frankfurt, a. M., 60325, Germany. <sup>4</sup>Joukowsky Institute for Archaeology and the Ancient World, Brown University, Providence, RI, 02912, USA. <sup>5</sup>Institut für Ur- und Frühgeschichte und Vorderasiatische Archäologie, Ruprecht-Karls-Universität Heidelberg, Heidelberg, 69117, Germany. <sup>6</sup>Department of Genetics, Harvard Medical School, Boston, MA, 02115, USA. <sup>7</sup>Broad Institute of Harvard and MIT, Cambridge, MA, 02142, USA. <sup>8</sup>Department of History, Harvard University, Cambridge, MA, 02138, USA. <sup>9</sup>Initiative for the Science of the Human Past at Harvard, Harvard University, Cambridge, MA, USA. <sup>10</sup>Institute for Archaeological Sciences, University of Tübingen, Tübingen, 72074, Germany. <sup>11</sup>Howard Hughes Medical Institute, Harvard Medical School, Boston, MD, 20815, USA. <sup>12</sup>Institut für Vor- und Frühgeschichtliche Archäologie und Provinzialrömische Archäologie, Ludwig-Maximilians-Universität München, Munich, 80799, Germany. Correspondence and requests for materials should be addressed to S.E. (email: [eisenmann@shh.mpg.de](mailto:eisenmann@shh.mpg.de)) or P.W.S. (email: [philipp.stockhammer@lmu.de](mailto:philipp.stockhammer@lmu.de))

from point A to point B and the subsequent intermixing with another group at point B<sup>8,10,11</sup>. Due to the limited space available to the authors in the journals where the papers were published, the two ancient DNA studies did not include substantial sections reflecting on the meanings of their results that could have headed off possible misinterpretations. Some archaeologists interpreted the papers as simplistically equating people with culture<sup>11</sup>.

The debate over the two 2015 papers made one thing very clear: in order to be able to engage in fruitful discussions geneticists and archaeologists need to agree on shared language, terms and concepts. Effective communication among researchers requires a common vocabulary. The goal of this paper is to address these issues of terminology and to make some suggestions about naming conventions researchers can choose to use in order to facilitate discussions across fields.

In part (I) we review the naming conventions for genetic clusters used in the past five years. Part (II) deals with the meaning of archaeological cultural designations that are frequently borrowed to name genetically distinguished groups of individuals and aims at providing geneticists with a better understanding of the problematic nature of these terms in the history of archaeological research. In part (III) three possible genetic nomenclatures are being discussed.

## Current Practice in the Naming of Genetic Clusters

Studies on ancient DNA extracted from human remains suffered from a major drawback during the 1990's and early years of the 21<sup>st</sup> century. The study of ancient human DNA was regarded as likely not feasible, due to the fact that modern human contamination was shown to be abundant in ancient remains and could not be distinguished from the DNA of our ancestors<sup>12–14</sup>. The advent of next generation sequencing technologies and the establishment of reliable criteria to authenticate ancient DNA, such as DNA damage patterns, however, have made it possible in the last five years to generate and analyse authentic genome-wide ancient DNA data from a large number of individuals<sup>15,16</sup>. This data can be analysed to identify statistical groupings of individuals who share more genetic variants with each other than with individuals outside these groups. Such groups are often called 'genetic clusters'.

The majority of genome-wide ancient DNA data comes from Western Eurasia. Studies on these data have revealed that the region's genetic landscape has been shaped by at least two major migrations after the initial settlement by hunter-gatherers in the Upper Palaeolithic – one from the Near East by the first farmers of Anatolian origin starting around 7000 BCE<sup>17–19</sup> and a second one during the 3<sup>rd</sup> millennium BCE, when descendants of steppe pastoralists spread to Northern and Central Europe<sup>6,7</sup>.

The first genome-wide ancient DNA studies reporting these findings grouped samples mainly using geographical terms in combination with relative time periods and/or subsistence practices, like *Mesolithic Europeans* or *Scandinavian hunter-gatherers* (see Table 1, 1–3)<sup>1,19,20</sup>.

The amount of genome-wide data has continued to increase massively, leading to the identification of more and more genetic groups and sub-groups. Since 2015, nearly all DNA papers on Western Eurasia have used what we would like to call "mixed systems" for the naming of genetic clusters. All of these systems are based on the same four elements for inventing names that they variably combine and order: 1. geographical terms (like "Scandinavian"), 2. relative time periods (like "Early Neolithic"), 3. subsistence practices (like "hunter-gatherer"), and 4. names of archaeological cultures (like "Yamnaya") (For more detail on naming conventions and names, see Table 1).

In 2015 Haak *et al.*<sup>7</sup> introduced a mixed system that has been adopted by many subsequent studies<sup>17,21–27</sup>. In Table 1 we call this nomenclature "mixed system (a)". The "mixed system (a)" usually combines two of the four elements mentioned. For example individuals from present-day Spain dating to the Early Neolithic are called *Spain\_EN*, individuals from the site Esperstedt in Germany, dating to the Middle Neolithic are designated *Esperstedt\_MN*. The term *Western European hunter-gatherer*, originally introduced by Lazaridis *et al.*<sup>19</sup>, was retained in Haak *et al.*<sup>7</sup>. In other cases the authors decided to use names of archaeological cultures that are usually also, but not always, combined with the relative dating. This led to names like *Bell\_Beaker\_LN* (Bell Beaker Late Neolithic), *Starcevo\_EN* (Starčevo Early Neolithic), and *Yamnaya*<sup>7</sup>.

This means that using the 'mixed system (a)' to name a newly observed cluster usually offers three possible options: <1: geographic term (country or region or site)\_relative time span> or <2: geographic term\_subsistence practice> or <3: archaeological culture (\_relative time span)>. Nevertheless, the selection of terms is neither unreflective nor arbitrary, since the 'mixed system (a)' incorporates some features which are of considerable importance for the applicability of nomenclatures in general. These can be summarised in five words: brevity, coherence, accessibility, flexibility, and stability.

**Brevity:** Each name used in the papers should be as short as possible while containing sufficient detail to distinguish the cluster from other genetic clusters in the same study and beyond.

**Coherence:** Individuals from genetically distinguishable groups should not be given the same name; individuals from genetically indistinguishable ones should.

**Accessibility:** The names given to genetic clusters should be recognisable and easy to remember. In the case of the 'mixed system (a)' this is often implemented by using archaeological cultural designations, like *Bell\_Beaker* and *Corded\_Ware*, which have well-recognised meanings that make them accessible to a wider audience. However, a potential pitfall in borrowing already existing names from archaeology is that an archaeological cultural designation may or may not have a one-to-one correspondence with a genetic cluster. For example individuals associated with the Bell Beaker Complex are not genetically homogeneous across Europe, and thus it is in genetic terms appropriate to use classifications that distinguish subgroups, e.g. *Beaker-associated Iberia*<sup>28</sup>.

**Flexibility:** The nomenclature needs to be flexible enough to adjust when there are new genetic findings. An appropriately flexible nomenclature should offer the possibility of both subdividing previously named groups into smaller ones and merging clusters which were at first found to be distinguishable. Depending on the scope and perspective of a study, the terms to describe genetic clusters may change to some extent from paper to paper. Sometimes groups need to be lumped together for trans-continental analysis and sometimes they need to be split for fine-grained study.

Publication	Nomenclature used	main Genetic Clusters *newly encountered	newly reported Individuals
<b>Lazaridis, I. et al. 2014</b> Ancient human genomes suggest three ancestral populations for present-day Europeans. <i>Nature</i> <b>513</b> , 409–413.	<b>geography or geography + subsistence practice</b>	<ul style="list-style-type: none"> <li>• Ancient North Eurasians (ANE)</li> <li>• *Scandinavian hunter-gatherers (SHG)</li> <li>• *Western European hunter-gatherers (WHG)</li> <li>• Early European Farmers</li> </ul>	9
<b>Gamba, C. et al. 2014</b> Genome flux and stasis in a five millennium transect of European prehistory. <i>Nat. Commun.</i> <b>5</b> , 5257.	<b>subsistence practice or relative dating</b>  limited use of cultural designations (Körös) in naming	<ul style="list-style-type: none"> <li>• individual labeling of 13 samples: <ul style="list-style-type: none"> <li>• KO1, KO2 (KO = Körös-Neolithic)</li> <li>• 11 other samples: relative dating + number (NE1, NE2, CO1, BR1, IR1, etc.)</li> </ul> </li> <li>• cluster names: <ul style="list-style-type: none"> <li>• Hunter-gatherers</li> <li>• *Neolithic group (meaning Neolithic samples from Hungary)</li> </ul> </li> </ul>	13
<b>Skoglund, P., et al. 2014</b> Genomic Diversity and Admixture Differs for Stone-Age Scandinavian Foragers and Farmers. <i>Science</i> <b>344</b> , 747–750.	<b>relative date + geography (+subsistence practice)</b>	<ul style="list-style-type: none"> <li>• Paleolithic Siberian</li> <li>• Mesolithic Europeans</li> <li>• Late Scandinavian hunter-gatherers</li> <li>• *Neolithic Scandinavian farmers</li> <li>• Chalcolithic farmer</li> </ul>	11
<b>Haak, W. et al. 2015</b> Massive migration from the steppe was a source for Indo-European languages in Europe. <i>Nature</i> <b>522</b> , 207–211.	<b>mixed system (a)</b> <ul style="list-style-type: none"> <li>• geography+ subsistence</li> <li>• geography + relative date</li> <li>• archaeological culture (+relative date)</li> </ul> keeping names given in Lazaridis et al. 2014, adding many new	<ul style="list-style-type: none"> <li>• Eastern European hunter-gatherers (EHG)</li> <li>• Early Neolithic (EN); comprising Starcevo_EN, LBK_EN, Spain_EN, and others</li> <li>• Middle Neolithic (MN); comprising Esperstedt_MN and others</li> <li>• *Yamnaya</li> <li>• *Corded Ware</li> <li>• Late Neolithic/Bronze Age (LN/BA); comprising *Bell_Beaker_LN, Unetice_EBA, and others</li> </ul>	69
<b>Allentoft, M. E. et al. 2015</b> Population genomics of Bronze Age Eurasia. <i>Nature</i> <b>522</b> , 167–172.	<b>mixed system</b> 3 elements, order varying <ul style="list-style-type: none"> <li>• relative date + subsistence + geography</li> <li>• geography + relative date + archaeological culture</li> </ul> e.g. “WHG + SHG” slightly reframed to “Mesolithic hunter-gatherers (West, Scandinavia)”	<ul style="list-style-type: none"> <li>• Mesolithic hunter-gatherers</li> <li>• Neolithic farmers + *Copper Age Remedello</li> <li>• European Late Neolithic and Bronze Age (*Corded Ware, *Bell Beakers, Scandinavia, and others)</li> <li>• Bronze Age steppe/Caucasus <ul style="list-style-type: none"> <li>• *Yamnaya,</li> <li>• *Sintashta,</li> <li>• *Armenia</li> <li>• *Afanasiovo</li> <li>• *Andronovo</li> </ul> </li> <li>• and others in Asia from Bronze Age to Iron Age</li> </ul>	101
<b>Mathieson, I. et al. 2015</b> Genome-wide patterns of selection in 230 ancient Eurasians. <i>Nature</i> <b>528</b> , 499–503.	<b>mixed system (a)</b>	<ul style="list-style-type: none"> <li>• WHG</li> <li>• EHG</li> <li>• SHG</li> <li>• Bronze Age (steppe)</li> <li>• *Eneolithic Samara</li> <li>• *Srubnaya</li> <li>• Sintashta/Andronovo</li> <li>• *Anatolia Neolithic</li> <li>• Early Neolithic (Europe); comprising LBK EN, Iberia EN</li> <li>• Middle Neolithic (Europe)</li> <li>• Chalcolithic Iberia</li> </ul>	163
<b>Günther, T. et al. 2015</b> Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. <i>Proc. Nat. Acad. Sci. USA</i> <b>112</b> , 11917–11922.	<b>subsistence practice sub-level: Relative date + Geography</b>	<ul style="list-style-type: none"> <li>• Hunter-gatherer</li> <li>• Farmer <ul style="list-style-type: none"> <li>• *Chalcolithic Iberian</li> <li>• Neolithic Hungarian</li> </ul> </li> </ul>	8
<b>Jones, E. R. et al. 2015</b> Upper Palaeolithic genomes reveal deep roots of modern Eurasians. <i>Nat. Commun.</i> <b>6</b> , 8912.	<b>mixed system (a)</b>	<ul style="list-style-type: none"> <li>• *Caucasus hunter-gatherer (CHG)</li> <li>• Western and Scandinavian hunter-gatherers (WHG, SHG)</li> <li>• Eastern Hunter-gatherers (EHG)</li> <li>• (Neolithic) farmers</li> <li>• (Bronze Age) Yamnaya</li> </ul>	3
<b>Cassidy, L. M. et al. 2016</b> Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. <i>Proc. Nat. Acad. Sci. USA</i> <b>113</b> , 368–373.	<b>mixed system</b> formatting and abbreviations slightly varying from (a)	<ul style="list-style-type: none"> <li>• Western HGs, Eastern HGs, Scandinavian HGs</li> <li>• Paleolithic HGs</li> <li>• Early Neolithic</li> <li>• Middle Neolithic</li> <li>• Western and Central European Late Neolithic to Bronze Age (Spanish Chalcolithic, Hungarian Bronze, *Irish Bronze Age, Yamnaya, and others)</li> </ul>	4
<b>Fu, Q. et al. 2016</b> The genetic history of Ice Age Europe. <i>Nature</i> <b>534</b> , 200–205.	<b>genetic nomenclature based on type-sites</b> “CHG” renamed to “Satsurblia”	<ul style="list-style-type: none"> <li>• Satsurblia (= CHG)</li> <li>• *Vestonice</li> <li>• Mal'ta</li> <li>• *El Mirón</li> <li>• *Villabruna</li> </ul>	51
Continued			

Publication	Nomenclature used	main Genetic Clusters *newly encountered	newly reported Individuals
<b>Broushaki, F. et al. 2016</b> Early Neolithic genomes from the eastern Fertile Crescent. <i>Science</i> 353, 499–503.	<b>geography + relative time</b> or <b>geography + subsistence practice</b> macro-scaling: WHG, CHG, SHG and EHG lumped together and grouped under name “European foragers”	<ul style="list-style-type: none"> <li>• European foragers</li> <li>• Aegean and European Neolithic farmers</li> <li>• *Neolithic Iranians</li> </ul>	4
<b>Lazaridis, I. et al. 2016</b> Genomic insights into the origin of farming in the ancient Near East. <i>Nature</i> 536, 419–424.	<b>mixed system (a)</b> “Natufian” only archaeological cultural designation used to name a genetic cluster term “Yamnaya” replaced by “steppe”	<ul style="list-style-type: none"> <li>• EHG</li> <li>• WHG</li> <li>• CHG</li> <li>• SHG</li> <li>• *Natufians</li> <li>• *Neolithic Levant (Levant_N)</li> <li>• *Neolithic Iran (Iran_N)</li> <li>• *Chalcolithic Iran (Iran_ChL)</li> <li>• Armenia (Armenia_ChL, Armenia_EBA, Armenia_MLBA)</li> <li>• Anatolian/European farmers (Anatolia_N, Europe_EN, Europe_MNChL)</li> <li>• Early/Middle Bronze Age steppe</li> <li>• Europe/steppe continuum (Steppe_MLBA, Europe_LNBA)</li> <li>• *Bronze Age Levant (Levant_BA)</li> <li>• *Chalcolithic Anatolia (Anatolia_ChL)</li> </ul>	44
<b>Kılınc, G. M. et al. 2016</b> The Demographic Development of the First Farmers in Anatolia. <i>Curr. Biol.</i> 26, 2659–2666.	<b>geography + relative time + site for newly sampled individuals</b> and <b>maintaining before established names (mixed system a)</b>	<ul style="list-style-type: none"> <li>• CHG</li> <li>• WHG</li> <li>• EHG</li> <li>• Swedish hunter-gatherers</li> <li>• *Central Anatolian Neolithic, Tepecik-Ciftlik</li> <li>• *Central Anatolian Neolithic, Boncuklu</li> <li>• *Northwest Anatolian Neolithic/Chalcolithic, Barcin</li> </ul>	9
<b>Lazaridis, I. et al. 2017</b> Genetic origins of the Minoans and Mycenaeans. <i>Nature</i> 548, 214–218.	<b>mixed system (a)</b>	<ul style="list-style-type: none"> <li>• European hunter-gatherers</li> <li>• EHG</li> <li>• WHG</li> <li>• SHG</li> <li>• Late Neolithic/Bronze Age Europe/Steppe (Steppe_EMBA, and others)</li> <li>• Anatolian/European farmers (Anatolia_N, *Greece_N, Europe_EN, Europe_MNChL)</li> <li>• Levant (Natufian, and others)</li> <li>• Caucasian/Iran (...)</li> <li>• *Aegean/Anatolian Bronze Age</li> <li>• *Anatolia_BA</li> <li>• *Minoan_Lasithi</li> <li>• *Minoan_Odigitria</li> <li>• *Crete_Armenoi</li> <li>• *Mycenaean</li> <li>• *Modern Greeks</li> </ul>	19
<b>Jones, E. R. et al. 2017</b> The Neolithic Transition in the Baltic Was Not Driven by Admixture with Early European Farmers. <i>Curr. Biol.</i> 27, 576–582.	<b>geography+ subsistence practice</b> or <b>geography + relative time</b>	e.g. Latvia_HG, Anatolian Chalcolithic, Scandinavian HG	8
<b>Martiniano, R. et al. 2017</b> The population genomics of archaeological transition in west Iberia. Investigation of ancient substructure using imputation and haplotype-based methods. <i>PLoS genet.</i> 13, e1006852.	<b>mixed system</b>	<ul style="list-style-type: none"> <li>• Western_HG1</li> <li>• Western_HG2</li> <li>• Scandinavian_HG</li> <li>• Caucasus_HG</li> <li>• Atlantic_Neolithic, Neolithic Portugal</li> <li>• Portugal_BA</li> <li>• Yamnaya_Afanasiovo</li> <li>• Sintashta_Andronovo</li> <li>• Russia_LBA_IA</li> <li>• AegeanEN_HungarianLKB</li> <li>• HungarianMLN_SpainCardialEN</li> </ul>	14
<b>Saag, L. et al. 2017</b> Extensive Farming in Estonia Started through a Sex-Biased Migration from the Steppe. <i>Curr. Biol.</i> 27, 2185–2193.e6.	<b>mixed system (a)</b>	<ul style="list-style-type: none"> <li>• Steppe EMBA</li> <li>• Steppe MLBA</li> <li>• European LNBA</li> <li>• Anatolian/Levantine EF</li> <li>• CCC</li> <li>• CWC</li> <li>• Caucasus HG</li> <li>• and others</li> </ul>	10
<b>Haber, M. et al. 2017</b> Continuity and Admixture in the Last Five Millennia of Levantine History from Ancient Canaanite and Present-Day Lebanese Genome Sequences. <i>Am. J. Hum. Genet.</i> 101, 274–282.	<b>mixed system (a)</b>	<ul style="list-style-type: none"> <li>• Natufians</li> <li>• Neolithic Levant, Levant_N</li> <li>• Neolithic Anatolia, Anatolia_N</li> <li>• Chalcolithic Iran, Iran_ChL</li> <li>• Chalcolithic/Bronze Age Armenia</li> <li>• Armenia_ChL</li> <li>• Armenia_EBA</li> <li>• Armenia_MLBA</li> <li>• *Bronze Age Levant, Sidon_BA</li> </ul>	5
Continued			

Publication	Nomenclature used	main Genetic Clusters *newly encountered	newly reported Individuals
Lipson, M. <i>et al.</i> 2017 Parallel palaeogenomic transects reveal complex genetic history of early European farmers. <i>Nature</i> 551, 368–372.	<b>mixed system (a)</b> with many archaeological cultural affiliations used distinguishing LBK in Transdanubia (LBKT MN) from LBK in Germany (LBK EN)	<ul style="list-style-type: none"> <li>● Anatolia_Neolithic</li> <li>● *Körös_EN</li> <li>● *Starcevo_EN</li> <li>● *ALPc_MN (Alföld Linear Pottery culture Middle Neolithic)</li> <li>● *LBKT_MN (Linearbandkeramik in Transdanubia)</li> <li>● *Vinca_MN</li> <li>● *Tisza_LN</li> <li>● *TDLN (Transdanubian Late Neolithic)</li> <li>● *Tiszapolgár_CA</li> <li>● *(Balaton_)Lasinja_CA</li> <li>● *Protoboleraz_CA</li> <li>● *Baden_CA</li> <li>● LBK_EN</li> <li>● Germany_MN</li> <li>● *Blatterhöhle_MN</li> <li>● Iberia_EN</li> <li>● Iberia_MN</li> <li>● Iberia_CA</li> <li>● WHG</li> </ul>	130
Sikora, M. <i>et al.</i> 2017 Ancient genomes show social and reproductive behavior of early Upper Paleolithic foragers. <i>Science</i> 358, 659–662.	<b>mixed system (a)</b>	<ul style="list-style-type: none"> <li>● *Sungir_UP</li> <li>● Motala_M</li> <li>● Barcin_EN</li> <li>● Hungary_MN</li> <li>● LBK_EN</li> <li>● Iberia_EN</li> <li>● Iberia_MN</li> <li>● Iberia_CA</li> <li>● BellBeaker_EBA</li> <li>● Central_EBA</li> <li>● YamnayaSamara_EBA</li> <li>● Potavka_EBA</li> <li>● Srubnaya_LBA</li> </ul>	5
Mittnik, A. <i>et al.</i> 2018 The genetic prehistory of the Baltic Sea region. <i>Nat. Commun.</i> 9, 442.	<b>mixed system</b> keeping many already established names adding new ones that consist of 3 elements: ● geography + relative time + archaeological culture	<ul style="list-style-type: none"> <li>● Baltic Mesolithic</li> <li>● WHG</li> <li>● EHG</li> <li>● SHG</li> <li>● *Baltic EMN Narva</li> <li>● Baltic MN CCC</li> <li>● *Baltic LN</li> <li>● *Baltic BA</li> </ul>	38
Mathieson, I. <i>et al.</i> 2018 The genomic history of southeastern Europe. <i>Nature</i> 555, 197–203.	<b>mixed system (a)</b>	<ul style="list-style-type: none"> <li>● WHG</li> <li>● *Ukraine_Mesolithic</li> <li>● *Ukraine_Neolithic</li> <li>● *Ukraine_Eneolithic</li> <li>● *Iron_Gates_HG</li> <li>● *Romania_HG, *Latvia_HG</li> <li>● *Latvia_MN</li> <li>● *Balkans_Neolithic</li> <li>● *Malak_Preslavets</li> <li>● *Peloponnese_Neolithic</li> <li>● *Balkans_Chalcolithic</li> <li>● *Varna</li> <li>● *Trypillia</li> <li>● *Vučedol</li> <li>● *Balkans_Bronze Age</li> <li>● LBK_Austria</li> <li>● *Globular_Amphora</li> </ul>	225
Olalde, I. <i>et al.</i> 2018 The Beaker phenomenon and the genomic transformation of northwest Europe. <i>Nature</i> 555, 190–196.	<b>mixed system</b>	<ul style="list-style-type: none"> <li>● Steppe Early Bronze Age</li> <li>● Corded Ware</li> <li>● Anatolia Neolithic</li> <li>● Europe Early Neolithic</li> <li>● Europe Middle Neolithic and Copper Age</li> <li>● Beaker-associated <ul style="list-style-type: none"> <li>● Central Europe <ul style="list-style-type: none"> <li>● *Britain</li> <li>● *Southern France</li> <li>● *Northern Italy</li> </ul> </li> <li>● Iberia</li> <li>● *The Netherlands</li> <li>● *Sicily</li> </ul> </li> </ul>	400

**Table 1.** Major publications of prehistoric ancient DNA data from Western Eurasia in chronological order detailing the naming principles used. NE, Neolithic; CO, Copper Age; BR, Bronze Age; IR, Iron Age; EN/MN/LN, Early/Middle/Late Neolithic; ChL, Chalcolithic; CA, Chalcolithic; E/M/LBA, Early/Middle/Late Bronze Age; EF, Early Farmer; HG, hunter-gatherer; CCC, Comb Ceramic Culture; CWC, Corded Ware Culture; EHG, Eastern European hunter-gatherer; CHG, Caucasus hunter-gatherer; SHG, Scandinavian hunter-gatherer; ANE, Ancient North Eurasian; WHG, Western European hunter-gatherer. The term *mixed system* means, that all 4 elements (geographical terms, relative time, subsistence practices and archaeological cultural names) are used in the publication to name genetic clusters. The (a) indicates that these publications use the type of mixed system first established in Haak *et al.*<sup>7</sup>.

**Stability:** Despite the need to be flexible, to some extent the names used in the genetic literature should be ones that can be reused in subsequent studies to help readers compare papers. Finding the appropriate line between flexibility and stability may be a challenge in some cases. An example is *Western European hunter-gatherer (WHG)*. This term was originally coined based on Mesolithic hunter-gatherers who lived about 8,000 years ago in present-day Luxembourg and Spain<sup>19</sup>. We now know that genetically similar people were also present in Sicily, the Balkans, and the Baltic region<sup>21,26,29</sup>. The authors of these subsequent papers decided to maintain the term and consequently used *WHG* also to refer to those newly sampled individuals that do technically not originate from Western Europe.

The ‘mixed system (a)’ developed organically over the past few years and has not been the subject of systematic criticism. It has been used in a considerable number of studies and gives a first impression of how much a uniform nomenclature can increase the readability across different publications. It makes clear that once a name for a certain cluster has been established, it should not be altered again and again without compelling grounds. The only remaining question is whether the ‘mixed-system (a)’ is the best suited choice for naming genetic clusters.

In contrast to the ‘mixed system (a)’ Fu *et al.*<sup>30</sup> introduced an entirely ‘genetic nomenclature’<sup>30</sup>. The authors of the study (including some of the authors of the present paper) highlighted five genetic clusters of individuals that share substantial amounts of genetic variants with each other. They named each of these clusters after the site where the oldest individual with high quality data originated. The five clusters were named *Satsurblia*, *Věstonice*, *Mal'ta*, *El Mirón*, and *Villabruna* cluster. The nomenclature borrows its system, but not the names itself, from archaeology, where groups are often named after so called ‘type-sites’ (see the following section on archaeological cultures), despite the fact that in the case of Fu *et al.*<sup>30</sup> the genetic clusters in question correspond strikingly well to material cultures (e.g. *Věstonice – Gravettian*, *El Mirón – Magdalenian*, *Villabruna – Mesolithic/Epipaleolithic/Azilian*). Accordingly, a genetic cluster first encountered by Jones *et al.*<sup>21</sup> and named there *CHG* (short for *Caucasian Hunter-gatherer*)<sup>21</sup>, a designation later taken over by Lazaridis *et al.*<sup>22</sup>, was renamed in Fu *et al.*<sup>30</sup> as the *Satsurblia Cluster*<sup>30</sup>. The authors themselves explain why they opted in favour of a new nomenclature: “In order not to prejudice any association between genetic and archaeological groupings among the individuals studied, we first allowed the genetic data alone to drive the groupings of the specimens, and only afterward examined their associations with archaeological cultural complexes” (p. 202)<sup>30</sup>. With this statement, Fu *et al.* touch a sensitive point on adopting archaeological cultural designations in genetic studies, which we would like to outline here.

## Archaeological Cultures and the Problem of Ethnicity

In the absence of the eye-witness information about present-day societies that sociologists use to define cultures, or the written information about ancient societies that historians use, the main corpus of information that archaeologists use to describe and define ancient cultures is confined to the material evidence. The term ‘material evidence’ refers to all traces that are left by past people. This includes artefacts (portable objects, e.g., pottery, weapons) and features (non-portable objects) like postholes, hearths, floors, walls, and ditches. Environmental and organic remains, like soils, sediments and human skeletons are also part of the material evidence. Archaeologists focus on the arrangement and relation among these different material remnants in order to detect patterns in the behaviour of past humans and interpret these practices. For their interpretation, the comparison with and the use of concepts from other disciplines like sociology, ethnology, and history play a major role. History is especially important for archaeological research during historic times when the written evidence and the archaeological evidence temporally overlap<sup>31–33</sup>.

The general observation that single objects and features as well as whole assemblages differ from each other in execution and style not only over time, but also on a geographical scale led to their classification into groups which are called ‘cultures’<sup>31,32</sup>. Usually the definition of an archaeological culture is based on only parts of the material object assembly. For prehistory stone tools and pottery are of crucial importance in this respect, since they constitute the most abundant and the most rapidly changing artefact groups<sup>33</sup>. Some archaeological cultures have been completely constructed on the basis of pottery typology and only afterwards the rest of the material assemblage was loosely fitted into this framework. A good example is the Early Neolithic in the Balkans. Here small changes in the quantitative composition of painted pottery as well as national borders have led to the construction of various archaeological cultures, like the *Starčevo Culture* in Serbia or the *Karanovo Culture* and *Western-Bulgarian-Painted-Pottery Culture* in Bulgaria with their numerous regional variants and sub-groups<sup>34,35</sup>. The case example of the Balkans illustrates that the distinction into different archaeological cultures or groups is relative and that there exists no universal explanation for what stands behind an archaeological culture. In particular, the question of whether a distinct material group/culture as defined by archaeologists also represented a meaningful entity in the past is a matter of debate and needs to be evaluated on a case-by-case basis.

Archaeological cultures are predominantly named after two systems: 1. Either they are named after a ‘type-site’, like the *Michelsberger Culture*<sup>36</sup>, or 2. their designation derives from a ‘type-artefact’ or ‘index-fossil’ as is the case with the *Bell Beaker Complex*<sup>37</sup>. Both systems are based on the assumption that the given diagnostic part of the evidence suffices to name the phenomenon as a whole. A third option for naming an archaeological culture consists in adopting a historically attested name, by transferring it to a materially discernible group, like *Philistine*<sup>38,39</sup> or *Viking*<sup>40</sup>. This “translation” can be based on written sources that are contemporaneous with the archaeological culture or postdate it (or both). The general idea that both the written sources and the material evidence depict the same group of people is driven mainly by their spatial equivalence. If and how the historical records about a certain group of people and the archaeological record are connected with each other needs again to be investigated on a case-by-case basis. These questions about identities and ethnicities in the past are one of the major fields of research in ancient studies<sup>40–46</sup>. It also needs to be taken into account here that most identifications of a historically attested group and an archaeological one and the corresponding naming of the one after the other date back to the 19<sup>th</sup> and early 20<sup>th</sup> centuries<sup>42,47</sup>.

The concept of defining cultures in archaeology was an integral part of the development of cultural history by archaeologists and anthropologists alike in North America and Europe in the 19<sup>th</sup> century<sup>33</sup>. Politically its origins are connected with the formation of modern national states and national identities in Europe during the 19<sup>th</sup> and 20<sup>th</sup> centuries<sup>47</sup>. The intellectual construction of a common history shared by the inhabitants of a country and putatively stretching back into prehistoric times was part of this process. Within this context of emerging national consciousness, which aimed at setting *our own* apart from *the other*, the German archaeologist Gustaf Kossinna established his ‘settlement archaeology’ (original: ‘Siedlungsarchäologie’)<sup>42,47</sup>. In his 1911 book *The Origin of the Germans* (original: *Die Herkunft der Germanen*) he stated: “Sharply defined archaeological culture-provinces coincide at all times with quite definite peoples or tribes.” (as translated in Childe 1956, 28<sup>48</sup>; original: Kossinna 1911, 3<sup>49</sup>: “Scharf umgrenzte archäologische Kulturprovinzen decken sich zu allen Zeiten mit ganz bestimmten Völkern oder Völkerstämmen.”). As a consequence, cultures defined via the archaeological record were perceived as being the material remnants of closed ethnic groups. For Kossinna (and others) these ethnic groups were distinct peoples who shared the same blood (or genes). During the 1930s and 1940s, fascist archaeology instrumentalised this notion to justify racial ideologies by adding a prehistoric perspective. By tracing *Germans* as an archaeological culture back into the Neolithic and over vast regions of Central and Eastern Europe, National Socialist archaeologists argued for ethnic cleansing and expansionist warfare. They were convinced of the biological superiority of their own *race* over other people and traced this superiority far back into prehistoric times. Kossinna’s ‘settlement archaeology’ not only shaped German archaeology up to the 1950s, but continued to exercise its influence well beyond. The general method of equating material culture with ethnicity became commonplace. In Britain, the main figure connected with this method is Vere Gordon Childe<sup>42</sup>. His work contributed greatly to cementing the static thinking that archaeological cultures equate to ethnic groups with shared ancestries.

In the decades following World War II, cultural studies moved away from the idea of the association between culture and ‘blood’. On the contrary, culture and ethnicity were perceived as dynamic, subjective and artificial. At the same time, Central European archaeological research became rather antiquarian. Archaeologists tended to collect and classify artefacts, but avoided far-reaching interpretations. Consequently, archaeological cultures became an abstract and mostly academic concern<sup>50</sup>. They continued to be used as a tool for classifying the material evidence and the undertone of equating material culture with ethnicity was never entirely dispelled – in part due to the lack of broad and open debate about the concept of culture in archaeology<sup>51</sup>.

With the advent of the ‘New Archaeology’ or ‘Processual Archaeology’ in the 1960s and 1970s, Anglo-American archaeological research shifted to functionalist questions of past social-cultural systems that understood culture (not cultures) as a means of adaptation to external factors and conditions imposed by the surrounding environment<sup>42,52,53</sup>.

The debate on prehistoric cultures and ethnicity only re-entered archaeology in the 1990s and early 2000s. By openly discussing how to trace ethnicity or identity in the archaeological record and by reviewing previous research, archaeologists developed a more nuanced understanding of the old argument about whether *pots equal people*. Whether and under which conditions an archaeologically discernible group can be viewed as the material remnants of a once living group of people who were connected via the same beliefs, social practices, ancestry, or in any other way needs to be investigated individually and will rarely stand without contradiction. In spite of all inherent problems, material groups in archaeology are often still called cultures. Yet the concept of cultures, including its whole history, seems to be so deeply rooted in archaeology that what it means (material group) and what it may or may not imply is widely if tacitly understood<sup>33,42–44,47,50,54</sup>. In some cases the term culture has been replaced by terms that the respective researchers believe to be a better description of the observed archaeological patterns<sup>55</sup>, as for example with the Bell Beaker *Complex* or *Phenomenon*<sup>37,56</sup>. The problematic nature of archaeological groupings has already been pointed out in archaeogenetic publications, for example, in the supplementary section of *The Beaker phenomenon and the genomic transformation of northwest Europe* by Olalde *et al.*<sup>28</sup>.

**The use of archaeological cultural designations in genetic studies.** If geneticists use names originating from the field of archaeology for genetic clusters, an inevitable by-product is to transfer at least part of the name’s archaeological meaning into the genetic study. Giving groups that have been identified through a completely different line of evidence – in this case material culture and genomics – the same or related names results in their conflation and the archaeological designations risk becoming reified in genetic terms (and vice versa). Even if there is a striking correspondence between a genetically defined cluster and the archaeological culture that the sampled individuals are associated with, they remain two different phenomena, identified by different methods with different criteria, which may or may not have some connections to each other. In principle, there does not need to be any clear relationship between the two at all.

This scenario has been shown in a recent genetic study. Here individuals from Iberia and Central Europe displaying, archaeologically speaking, the same cultural affiliations – both would be grouped under the term *Bell Beaker Complex* – are genetically rather heterogeneous<sup>28</sup>. Another interesting case is the so-called *Srubnaya outlier*: An individual who, archaeologically speaking, belongs to the *Srubnaya* culture turned out to be genetically very distinct from the other individuals with the same archaeological cultural affiliation<sup>17</sup>.

These two examples illustrate that only further analysis and discussion on a case-by-case basis can clarify the likelihood and nature of the association between a material and a genetic group. Even a specific pattern of relatedness between an archaeological culture and a genetic cluster does not in itself prove that this archaeological culture was a meaningful human entity in the past, but invites further reflection and investigation.

Despite these concerns, we do not feel it is appropriate to abandon the practice of comparing a genetic cluster and an archaeological culture. Correlations between these, if they exist, are of profound interest, as it is not irrelevant if people who shared one archaeological culture (considered as the material expression of parts of shared social practices and traditions of particular groups of actors) also shared a similar genetic makeup, and therefore

ancestry. Such convergences might help shed light on the nature of an archaeological culture, how it spread, and what may lie hidden behind a group of similar objects and/or practices. They can help to build a bridge from the material record back to the people who created and used these objects. The opposite is of course equally true, when people who did not share a similar genetic makeup chose to adopt similar cultural practices. We only insist here on a clear distinction between the two kinds of evidence, the genetic and the archaeological.

In most cases, scientists publishing genetic studies are fully aware of the general difference between a genetic and an archaeological group. This may explain why there has so far been no public discussion about the nomenclature of genetic clusters. Although it may simply not cause major misunderstandings at the source, where the data are generated, the casual conflation of archaeological and genetic classifications has considerable potential to confuse subsequent use of the data and its interpretation by third parties.

The young field of archaeogenetics needs to be as deliberate as possible in terminology to avoid falling into the pitfalls of earlier approaches to archaeological cultures and related interpretations. Even as an increasing number of archaeologists is moving away from the concept of material cultures in archaeology as a whole, it may seem unwise to expand it now to other disciplines<sup>33,57</sup>. Another related issue needs to be taken into consideration here, but lies beyond the scope of this paper, that is the political instrumentalisation of names for archaeological groups, especially those derived from historical (written) sources. Varying political agents still (ab)use archaeological group designations as ethnic labels with genealogical implications to underpin their respective agendas<sup>58</sup>.

### Three Possibilities for an Archaeogenetic Nomenclature

Since the costs of ancient DNA analysis are dropping rapidly and may soon become a routine part of archaeological practice, the time seems ripe to propose a systematisation in nomenclature that minimises the risks of falling into the various traps discussed above. The advantage of a universal naming system for clusters in ancient genetics lies in its user-friendliness and clarity. The ability to compare as well as integrate results will significantly improve. We have already listed five criteria for the applicability of nomenclatures in general: brevity, coherence, accessibility, flexibility, and stability. The following three nomenclature systems we propose all respect these five aspects to different degrees.

**Numeric Nomenclature.** The simplest possibility for a nomenclature would be to give each genetic cluster a number. In such a system, the observed genetic clusters could be called *population 1*, *population 2*, etc. or even *cluster 1*, *cluster 2*, and so on. Sub-divisions in regional in-depth studies could then be called *cluster 1a*, *cluster 1b*, *cluster 2a*, and so on. To avoid major confusions it would be important that the label for a cluster is maintained in subsequent studies by the same authors and others. The major advantage of such a system lies in its neutrality concerning the meaning of a genetically distinguished group. It completely avoids dilemmas arising from controversial cultural designations. On the other hand, the greatest disadvantage of a numeric system stems exactly from this neutrality. A numeric nomenclature does not provide any landmarks or clues to readers about the groups to which it refers. Its accessibility seems to be rather low in comparison to systems that use labels that have meanings beyond the pure function of denomination. A numeric nomenclature runs the risk of producing confusion since the connection between label (e.g. *cluster 1*) and what the label designates (e.g. a genetic cluster detected in the Bronze Age in Central Europe) are difficult to remember. This is why we would in general not recommend a numeric nomenclature.

**Mixed Nomenclature (a).** A second option is to use a modification of the ‘mixed system (a)’, where all names have a format like <1: geographic term (country or region or site)\_relative time span> or <2: geographic term\_ subsistence practice> or <3: archaeological culture (\_relative time span)>. The advantage is that the ‘mixed nomenclature (a)’ is already in use and the names that have been established in different publications would not need to be changed. However, slight alterations in terms of systematisation and the accentuation of archaeological cultural designations could improve its applicability.

Systematisation: in the first few years of the genome-wide ancient DNA era, the number of samples of individuals associated with any particular archaeological culture was limited, which has had the consequence that small numbers of samples from a limited number of sites were used to make generalisations about potentially much broader groups. In other words, a few individuals from only one or two sites were used to impute the genetic makeup of whole territories or archaeological groups during the relevant time periods.

There are two possible adjustments that could address this shortcoming. One would be to use a location name in addition to the archaeological culture name, as Lazaridis *et al.*<sup>23</sup> did with Bronze Age samples from Crete, naming them *Minoan\_Lasithi* and *Minoan\_Odigitria* instead of only using the generalising term *Minoan* for both<sup>23</sup>. In that paper, the authors made the statement that the individuals associated with the material culture that has been called *Minoan* are genetically very close, suggesting some degree of genetic homogeneity corresponding to the archaeological culture.

The other possibility for making the geographic distribution of samples clearly visible in the naming of genetic clusters would be to use regions and archaeological cultural designations only if the scale of sampling allows it. In this approach a genetic cluster would only be named after a whole region or archaeological culture if the data derives from a number of different sites distributed over a corresponding geographical area. If a genetic cluster is based on data from a single site, it would be named accordingly with no further qualifiers (e.g., this would make the examples above *Hagios Charalambos* and *Odigitria*).

Another general recommendation we would like to add here concerns the synoptic visualisation of the individual samples. Regardless of which approach is used to decide on the names of the clusters, each paper should include a table that gives an overview of the individual samples. The table should list each new sample of the study including the archaeological identifier (inventory or reference number of the individual), the site name, the



relative, and the absolute dating, and the attribution to the relevant cluster (cf. Haak *et al.*<sup>7</sup> extended data Table 2<sup>7</sup>). The goal should always be to make the data and results as easily accessible to the reader as possible.

The accentuation of archaeological cultural designations: If the ‘mixed system (a)’ becomes common practice in the future, archaeological cultural designations will frequently be used to name genetic clusters. In these cases, it might be advisable to invent a scheme to clarify whether the term used is referring to the archaeological group or its genetic counterpart. We have outlined earlier that a similar distribution of an archaeological group and a genetic cluster still does not make them the same. One possibility to make this general difference apparent to the reader would be to mark cultural terms when they are designating a genetic concept rather than an archaeological one. For example, an asterisk could refer to a genomic group that is being identified/associated with an archaeological group, e.g., \*Michelsberg\_Culture and \*Bell\_Beaker(\_Complex). The practice of asterisking is already established in linguistics and philology. “\*Word” means that the form ‘word’ is not positively attested in ancient or modern sources but can be deduced to have existed from other words known in ancient and modern languages. If applied to the case of genetics, the system would visibly establish the possible lines that might be drawn between a genetic cluster and the archaeological culture with a similar distribution, without equating them. Another possibility would be to use *Italic Font* rather than asterisks, since asterisks may be confusing to general readers, and also could be problematic for computer programmes that process data and for which asterisks can have specific meanings. Both markings, asterisks and *Italic Font*, are offered only as possibilities. The authors of this manuscript had differing opinions in their regard. Therefore, while we are not in a position to recommend either system as a new standard, we suggest that there should be further discussion about them in the archaeological, genetic, and other communities.

**Geographic-Temporal Nomenclature.** A third possible strategy is to avoid cultural terms altogether, and to use only geographic designations such as *Levant* or *North\_Pontic*, in combination with a rough indication of the time period, e.g., *BA* for Bronze Age – being conscious of the fact that these chronological nomenclatures are always spatially conditioned and that, for instance, the Bronze Age in Mesopotamia refers to a different time period in absolute chronology than the Bronze Age in Scandinavia. In this respect the system can be understood as a reduction of the ‘mixed system (a)’ to only one principle: <1: geographic term (country or region or site)\_relative time span>. A genetic cluster in Central Europe roughly spreading over the same area and time as the Bell Beaker Complex in archaeological terms might thus be called *C\_Europe\_LN* (Central-European Late Neolithic).

In general, we view geographic names as valuable because they are empirically grounded and highly flexible, especially when zooming in and out of regions in subsequent studies. Geographic designations can be easily sub-divided, if studies of smaller regions are conducted to gain a higher resolution. Moreover, they can be conveniently merged and translated into a larger, more global scale when looking at the genetic history of whole continents. For example, if a study focuses exclusively on the Upper Rhine Valley and encounters two distinct genetic clusters, they could be called *UpperRhine\_N\_LN* and *UpperRhine\_S\_LN*, with the *N* standing for North, the *S* standing subsequently for South and *LN* being the common abbreviation for the Late Neolithic. These two micro-clusters could then be integrated with other data and together translated into a Central European Late Neolithic cluster (*C\_Europe\_LN*), if a study focuses on the genetic history of Europe as a whole.

This particular system of naming separates genetic clusters and archaeological groups more clearly than the ‘mixed system (a)’. A slight variation on this option for a nomenclature might make it more user-friendly. If it seems advantageous for the questions a paper deals with, the name of the archaeological culture could be added to the spatial-time denomination in parentheses, when a genetic cluster is mentioned in a text or table for the first time. For example this would change *C\_Europe\_LN* into *C\_Europe\_LN\_(Bell Beaker)* or *C\_Europe\_LN (\*Bell-Beaker)*.

The question remains open whether the names of present-day countries should be used as geographic labels for naming genetic clusters. Countries are dynamic entities that are political constructs. Today’s national states with their clear borders did not exist in ancient times, but have arisen mainly since the 18<sup>th</sup> century. Each country has its own history of how it came to be. The correspondence between ‘nation’ and ‘genetic cluster’ is a broad and interesting field to investigate<sup>59</sup>. In some cases the two may correspond, in others not. Another aspect that deserves consideration here is that the names of present-day countries are also politically and ethno-nationalistically charged. Such modern ethnic or national resonance could spill over into archaeogenetic observations, if the same names are employed for naming genetic clusters<sup>60</sup>. It is important not to assume *a priori* strong correlations between clusters revealed by genetics and nationality or other categorical labels based on ethnicity, geography, language, religion, or any other cultural attribute. Neither can we assume that they are entirely distinct. The degree of correspondence (if any) should be evaluated on a case-by-case basis.

The main obstacle in reading a genetic nomenclature based on a combination of geographic terms and relative time periods is the temporal component, since relative periods such as the Late Neolithic or the Late Bronze Age date to vastly different absolute time frames depending on the geographical region in question. In macro-scale studies this can be solved for example by adding a suffix with an approximate date, e.g. *C\_Europe\_2800\_2200BCE* for Late Neolithic individuals from Central Europe, roughly spanning from 2800 to 2200 BCE. We think this approach can be useful, and could lead to a modified naming template such as <1a: geographic term\_relative date> or <1b: geographic term\_absolute date>.

The ‘geographic-temporal nomenclature’ holds yet another challenge. What happens if two distinct genetic clusters are encountered that spread during roughly the same time period over the same area or at least partly overlap spatially, as is the case in 3<sup>rd</sup> millennium BCE Central Europe? Would we then again include archaeological cultural designations, like *Bell Beaker* and *Corded Ware*? Moreover, there is the possibility that two co-existing genetic clusters have no equivalence in the archaeological record. Another option would be to combine the ‘geographic-temporal system’ in such cases with a numerical one that would result in <1a: geographic term\_relative time span\_number> or <1b: geographic term\_absolute time span\_number>.

## Conclusions

Genetic clusters are as flexible and dynamic as archaeological groupings. Both are theoretical constructions and result from our epistemological need to create space-time-entities as aids for further understanding. We need these entities in order to communicate, compare and integrate our results, but they remain artificial and to some extent arbitrary. This means inevitably that the names we give genetic groups are technical terms. Still, the nature of names is that they have or rather convey meanings. Those meanings are by no means static, but dynamic. The names we give genetic clusters and the way we use these names in the future also contribute to what we see in a genetic cluster. In this respect, the creation of a carefully reflected nomenclature for genetic clusters is of crucial importance for the future development of archaeogenetics and cooperation with affiliated fields like archaeology and history. We all need a common vocabulary when referring to a certain group, irrespective of whether we are dealing with an archaeological or a genetic group. Clear and stable names can help to avoid misunderstandings.

For user-friendliness and clarity a genetic nomenclature should respect the general qualities we have commented on in connection with the ‘mixed system (a)’: brevity, coherence, accessibility, flexibility, and stability. The ‘mixed system (a)’ incorporates all these features, but has shortcomings with regard to unsystematic choices of names in the past and the disputed use of archaeological cultural designations to denominate genetic clusters.

We therefore introduced with the ‘numeric system’ (option 1) and the ‘geographic-temporal system’ (option 3), two alternatives to the ‘mixed system (a)’ (option 2). While all of the proposed systems have advantages and challenges, the authors of this study agree that a numeric system should not be the first choice for future research. Its potential to confuse readers and hinder comparisons between papers outweighs the advantage of neutrality in naming.

With regard to the remaining two options, the archaeologists generally favour the ‘geographic-temporal system’, because it most visibly separates terms for archaeological cultures from those of genetic clusters. Several geneticists involved in this paper have pointed to the unsolved challenges presented by it and argue for a more coherent ‘mixed nomenclature (a)’.

In the end, it is up to each work group to decide how to name clusters. It is our aim to raise awareness of the issues involved and their consequences for subsequent investigation and interpretation. We hope that this article will serve as a basis to promote further reflection on the topic of naming conventions in archaeogenetics.

**Data availability statement.** No datasets were generated or analysed during the current study.

## References

- Gamba, C. *et al.* Genome flux and stasis in a five millennium transect of European prehistory. *Nat. Commun.* **5**, 5257 <https://doi.org/10.1038/ncomms6257> (2014).
- Pickrell, J. K. & Reich, D. Toward a new history and geography of human genes informed by ancient DNA. *TIG* **30**, 377–389, <https://doi.org/10.1016/j.tig.2014.07.007> (2014).
- Krause, J. & Pääbo, S. Genetic Time Travel. *Genetics* **203**, 9–12, <https://doi.org/10.1534/genetics.116.187856> (2016).
- Leonardi, M. *et al.* Evolutionary Patterns and Processes. Lessons from Ancient DNA. *Syst. Biol.* **66**, e1–e29, <https://doi.org/10.1093/sysbio/syw059> (2017).
- Callaway, E. Divided by DNA. The uneasy relationship between archaeology and ancient genomics. *Nature* **555**, 573–576, <https://doi.org/10.1038/d41586-018-03773-6> (2018).
- Allentoft, M. E. *et al.* Population genomics of Bronze Age Eurasia. *Nature* **522**, 167–172, <https://doi.org/10.1038/nature14507> (2015).
- Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211, <https://doi.org/10.1038/nature14317> (2015).
- Vander Linden, M. Population history in third-millennium-BC Europe. Assessing the contribution of genetics. *World Archaeol.* **48**, 714–728, <https://doi.org/10.1080/00438243.2016.1209124> (2017).
- Anthony, D. W. Migration in Archaeology. The Baby and the Bathwater. *Am. Anthropol.* **92**, 895–914 (1990).
- Furholt, M. Massive Migrations? The Impact of Recent aDNA Studies on our View of Third Millennium Europe. *Eur. j. archaeol* **21**, 159–191, <https://doi.org/10.1017/ea.2017.43> (2018).
- Heyd, V. Kossinna’s smile. *Antiquity* **91**, 348–359, <https://doi.org/10.15184/aqy.2017.21> (2017).
- Pääbo, S. Human evolution. *TIG* **15**, M13–M16, [https://doi.org/10.1016/S0168-9525\(99\)01904-6](https://doi.org/10.1016/S0168-9525(99)01904-6) (1999).
- Pääbo, S. *et al.* Genetic analyses from Ancient DNA. *Annu. Rev. Genet* **38**, 645–679, <https://doi.org/10.1146/annurev.genet.37.110801.143214> (2004).
- Hagelberg, E., Hofreiter, M. & Keyser, C. Ancient DNA: the first three decades. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **370**, 20130371, <https://doi.org/10.1098/rstb.2013.0371> (2015).
- Stoneking, M. & Krause, J. Learning about human population history from ancient and modern genomes. *Nat. Rev. Genet.* **12**, 603–614, <https://doi.org/10.1038/nrg3029> (2011).
- Margulies, M. *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376–380, <https://doi.org/10.1038/nature03959> (2005).
- Mathieson, I. *et al.* Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503, <https://doi.org/10.1038/nature16152> (2015).
- Haak, W. *et al.* Ancient DNA from European Early Neolithic Farmers Reveals Their Near Eastern Affinities. *PLoS Biol.* **8**, 1–16, <https://doi.org/10.1371/journal.pbio.1000536> (2010).
- Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413, <https://doi.org/10.1038/nature13673> (2014).
- Skoglund, P. *et al.* Genomic Diversity and Admixture Differs for Stone-Age Scandinavian Foragers and Farmers. *Science* **344**, 747–750, <https://doi.org/10.1126/science.1253448> (2014).
- Jones, E. R. *et al.* Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat. Commun.* **6**, 8912, <https://doi.org/10.1038/ncomms9912> (2015).
- Lazaridis, I. *et al.* Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419–424, <https://doi.org/10.1038/nature19310> (2016).
- Lazaridis, I. *et al.* Genetic origins of the Minoans and Mycenaeans. *Nature* **548**, 214–218, <https://doi.org/10.1038/nature23310> (2017).
- Saag, L. *et al.* Extensive Farming in Estonia Started through a Sex-Biased Migration from the Steppe. *Curr. Biol.* **27**, 2185–2193.e6, <https://doi.org/10.1016/j.cub.2017.06.022> (2017).
- Haber, M. *et al.* Continuity and Admixture in the Last Five Millennia of Levantine History from Ancient Canaanite and Present-Day Lebanese Genome Sequences. *Am. J. Hum. Genet.* **101**, 274–282, <https://doi.org/10.1016/j.ajhg.2017.06.013> (2017).

26. Lipson, M. *et al.* Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* **551**, 368–372, <https://doi.org/10.1038/nature24476> (2017).
27. Sikora, M. *et al.* Ancient genomes show social and reproductive behavior of early Upper Paleolithic foragers. *Science (New York, N.Y.)* **358**, 659–662, <https://doi.org/10.1126/science.aao1807> (2017).
28. Olalde, I. *et al.* The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature* **555**, 190–196, <https://doi.org/10.1038/nature25738> (2018).
29. Mathieson, I. *et al.* The genomic history of southeastern Europe. *Nature* **555**, 197–203, <https://doi.org/10.1038/nature25778> (2018).
30. Fu, Q. *et al.* The genetic history of Ice Age Europe. *Nature* **534**, 200–205, <https://doi.org/10.1038/nature17993> (2016).
31. Renfrew, C. & Bahn, P. G. *Archaeology. Theories, Methods and Practice*. 5th ed. (Thames & Hudson, 2008).
32. Eggert, M. K. H. & Samida, S. *Ur- und Frühgeschichtliche Archäologie*. 1st ed. (Francke, 2009).
33. Roberts, B. W. & Vander Linden, M. Chapter 1 in *Investigating Archaeological Cultures. Material Culture, Variability, and Transmission* (eds Roberts, B. W. & Vander Linden, M.) 1–21 (Springer, 2011).
34. Boyadzhiev, Y. Early Neolithic Cultures on the territory of Bulgaria in *The first Neolithic sites in Central/South-East European Transect, Volume 1* (eds Gatsov, I. & Boyadzhiev, Y.) 7–43 (Archaeopress, 2009).
35. Özdoğan, M. Neolithic cultures at the contact zone between Anatolia and the Balkans - Diversity and homogeneity at the Neolithic frontier in *Aegean - Marmara - Black Sea: The Present State of Research on the Early Neolithic* (eds. Gatsov, I. & Schwarzberg, H.) 21–28 (Beier & Beran, 2006).
36. Jeunesse, C. Die Michelsberger Kultur in *Jungsteinzeit im Umbruch. Die "Michelsberger Kultur" und Mitteleuropa vor 6000 Jahren* (ed. Badisches Landesmuseum Karlsruhe) 46–55 (Badisches Landesmuseum, 2010).
37. Strahm, C. (ed.) *Das Glockenbecher-Phänomen. Ein Seminar* (Institut für Ur- und Frühgeschichte der Universität Freiburg, 1995).
38. Killebrew, A. E. & Lehmann, G. Introduction in *The Philistines and Other "Sea Peoples" in Text and Archaeology* (eds Killebrew, A. E. & Lehmann, G.) 1–17 (Society of Biblical Literature Atlanta, 2013).
39. Maeir, A. M. & Hitchcock, L. A. Rethinking the Philistines: A2017 Perspective in *Rethinking Israel. Studies in the History and Archaeology of Ancient Israel in Honor of Israel Finkelstein* (eds. Lipshits, O., Gadot, Y. & Adams, M. J.) 247–266 (Eisenbrauns, 2017).
40. Hofmann, K. P. With vikingr into the Identity Trap: When Historiographical Actors Get a Life of Their Own in *The Genetic Challenge to Medieval History and Archaeology* (eds Pohl, W. & Gingrich, A.) 91–122 [https://doi.org/10.1553/medievalworlds\\_no4\\_2016s91](https://doi.org/10.1553/medievalworlds_no4_2016s91) (Austrian Academy of Sciences Press, 2016).
41. Burmeister, S. Archaeological Research on Migration as a Multidisciplinary Challenge in *The Genetic Challenge to Medieval History and Archaeology* (eds Pohl, W. & Gingrich, A.) 42–64; [https://doi.org/10.1553/medievalworlds\\_no4\\_2016s42%20](https://doi.org/10.1553/medievalworlds_no4_2016s42%20) (Austrian Academy of Sciences Press, 2016).
42. Jones, S. *The Archaeology of Ethnicity. Constructing identities in the past and present*. 1st ed. (Routledge, 1997).
43. Hall, J. M. *Ethnic identity in Greek antiquity* (Cambridge University Press, 1997).
44. McInerney, J. (ed.) *A Companion to Ethnicity in the Ancient Mediterranean* (Wiley, 2014).
45. Rieckhoff, S. & Sommer, U. (eds) *Auf der Suche nach Identitäten: Volk – Stamm – Kultur – Ethnos. Internationale Tagung der Universität Leipzig vom 8.-9. Dezember 2000* (Archaeopress, 2007).
46. Lucy, S. Ethnic and Cultural Identities in *The Archaeology of Identity* (eds. Díaz-Andreu García, M., Lucy, S., Babić, S. & Edwards, D. N.) 86–109 (Routledge, 2005).
47. Brather, S. *Ethnische Interpretationen in der frühgeschichtlichen Archäologie. Geschichte, Grundlagen und Alternativen*. 1st ed. (de Gruyter, 2004).
48. Childe, V. G. *Piecing together the Past. The Interpretation of Archaeological Data* (Routledge & Kegan Paul, 1956).
49. Kossinna, G. *Die Herkunft der Germanen. Zur Methode der Siedlungsarchäologie*. 1st ed. (Kabitzsch, 1911).
50. Hofman, K. P. & Stockhammer, P. W. Beyond Antiquarianism: Current Theoretical Issues in German-speaking Archaeology. (with comments by U. Veit, T. Meier, R. Bernbeck and K. Kristiansen and a response by the authors). *Arch Dial* **24**, 1–87 (2017).
51. Fehr, H. *Germanen und Romanen im Merowingerreich. Frühgeschichtliche Archäologie zwischen Wissenschaft und Zeitgeschehen* (de Gruyter, 2010).
52. White, L.A. *The Evolution of Culture. The Development of Civilization to the Fall of Rome* (MacGraw-Hill, 1959).
53. Binford, L. R. Archaeology as Anthropology. *Am. antiq.* **28**, 217–225 (1962).
54. Roberts, B. W. & Vander Linden, M. (eds) *Investigating Archaeological Cultures. Material Culture, Variability, and Transmission* (Springer, 2011).
55. Siegmund, F. Kulturen, Technokomplexe, Völker und Identitätsgruppen. Eine Skizze der archäologischen Diskussion. *Arch. Inf.* **37**, 53–65, <https://doi.org/10.11588/ai.2014.0.18191> (2014).
56. Vander Linden, M. *Le phénomène campaniforme dans l'Europe du 3ème millénaire avant notre ère. Synthèse et nouvelles perspectives* (Archaeopress, 2006).
57. Maran, J. Later Balkan prehistory: A transcultural perspective in *Balkan Dialogues. Negotiating Identity between Prehistory and the Present* (eds Gori, M. & Ivanova, M.) 17–37 (Routledge, 2017).
58. Kohl, P. L. & Fawcett, C. (eds) *Nationalism, politics, and the practice of archaeology* (Cambridge University Press, 1995).
59. Novembre, J. *et al.* Genes mirror geography within Europe. *Nature* **456**, 98–101, <https://doi.org/10.1038/nature07331> (2008).
60. Nash, C. Genome geographies. Mapping national ancestry and diversity in human population genetics. *Trans. Inst. Br. Geogr.* **38**, 193–206, <https://doi.org/10.1111/j.1475-5661.2012.00512.x> (2013).

## Acknowledgements

This paper is the result of a discussion that began during the first workshop of the *Max Planck–Harvard Research Center for the Archaeoscience of the Ancient Mediterranean* (short: *MHAAM*), held on February 17<sup>th</sup>–18<sup>th</sup> 2017, at the Harvard University, Cambridge, MA. We would like to thank all participants for their contributions regarding this topic. The research of SE, JK, MM, DR and PWS was generously funded by the *Max Planck – Harvard Research Center for the Archaeoscience of the Ancient Mediterranean*. DR is an investigator of the Howard Hughes Medical Institute. PWS' research is also part of his ongoing ERC Starting Grant project *FoodTransforms: Transformations of Food in the Eastern Mediterranean Late Bronze Age* (ERC-2015-StG 678901-FoodTransforms) funded by the European Research Council.

## Author Contributions

S.E. and P.W.S. wrote the main manuscript text with input from all co-authors. The manuscript was reviewed at several stages by all co-authors.

## Additional Information

**Competing Interests:** The authors declare no competing interests.

**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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018