Neolithic genomes reveal a distinct ancient HLA allele pool and population transformation in Europe

Alexander Immel[1], Christoph Rinne[2], John Meadows[3,4], Rodrigo Barquera[5], András Szolek[6], Federica Pierini[7], Julian Susat[1], Lisa Böhme[1], Janina Dose[1], Joanna Bonczarowska[1], Clara Drummer[2], Katharina Fuchs[1], David Ellinghaus[1], Jan Christian Kässens[1], Martin Furholt[8], Oliver Kohlbacher[6,9,10,11], Sabine Schade-Lindig[12], Iain Mathieson[13], Andre Franke[1], Johannes Krause[5], Johannes Müller[2], Tobias L. Lenz[7], Almut Nebel[1], Ben Krause-Kyora[1,*]

*** corresponding author**: b.krause-kyora@ikmb.uni-kiel.de

[1]Institute of Clinical Molecular Biology, Kiel University, Rosalind-Franklin-Straße 12, 24105 Kiel, Germany

[2]Institute of Pre- and Protohistoric Archaeology, Kiel University, Johanna-Mestorf-Straße 2-6, 24118 Kiel, Germany

[3]Leibniz Laboratory for AMS Dating and Isotope Research, Kiel University, Max-Eyth-Str. 11-13, 24118 Kiel, Germany

[4]Centre for Baltic and Scandinavian Archaeology (ZBSA), Schloss Gottorf, 24837 Schleswig, Germany

[5]Max Planck Institute for the Science of Human History, Khalaische Strasse 10, 07745 Jena, Germany

[6]Applied Bioinformatics, Dept. for Computer Science, University of Tübingen, Sand 14, 72076 Tübingen, Germany

[7]Research Group for Evolutionary Immunogenomics, Max Planck Institute for Evolutionary Biology, August-Thienemann-Str. 2, 24306 Plön, Germany

[8]Department of Archaeology, Conservation and History, University of Oslo, Blindernveien 11, 0371 Oslo

[9]Institute for Bioinformatics and Medical Informatics, University of Tübingen, Sand 14, 72076 Tübingen, Germany

[10]Institute for Translational Bioinformatics, University Hospital Tübingen, Hoppe-Seyler-Str. 9, 72076 Tübingen, Germany

[11]Biomolecular Interactions, Max Planck Institute for Developmental Biology, Max-Planck-Ring 5, 72076 Tübingen

[12]Landesamt für Denkmalpflege hessenArchäologie, Schloss Biebrich, 65203 Wiesbaden

[13]Department of Genetics, Perelman School of Medicine, University of Pennsylvania, 415 Curie Blvd, Philadelphia, PA 19104, USA

Abstract

The Wartberg culture (WBC, 3,500-2,800 BCE) dates to the Late Neolithic period, a time of important demographic and cultural transformations in western Europe. We perform a genome-wide analysis of 42 individuals who were interred in a WBC collective burial in Niedertiefenbach, Germany (3,300-3,200 cal. BCE). Our results highlight that the Niedertiefenbach population indeed emerged at the beginning of the WBC. This farming community was genetically heterogeneous and carried a surprisingly large hunter-gatherer ancestry component (40%). We detect considerable differences in the human leukocyte antigen gene pool between contemporary Europeans and the Niedertiefenbach individuals whose immune response was primarily geared towards defending viral infections.

Introduction

Over the last few years, large-scale ancient DNA (aDNA) studies have provided unprecedented insights into the peopling of Europe and the complex genetic history of its past and present-day inhabitants[1 2,3 4,5]. Recent research has particularly focused on the population dynamics during the Neolithic period. The first agriculturalists across central Europe, who are associated with the uniform Linear Pottery culture (Linearbandkeramik, LBK, 5,450-4,900 BCE) across central Europe, probably co-existed with local hunter-gatherers (HG) for about two thousand years[6]. Although these groups are thought to have lived in close proximity, initially only limited admixture occurred[2,3]. This situation changed later (4,400-2,800 BCE) when the gene-pool of the early farmers was transformed through the introgression of genomic components typical of HG populations[1,3,7].

The Late Neolithic period is archaeologically characterized by strong regional diversification and a patchwork of small units of classification (i.e. archaeological cultures)[8]. One of the western units that emerged at the beginning of the Late Neolithic period is associated with the Wartberg culture (WBC, 3,500-2,800 BCE), which most likely developed from the Late Michelsberg culture (MC, 3,800-3,500 BCE)[9,10]. WBC is mainly found in western central Germany (Fig. 1)[11,12]. It is known for its megalithic architecture of large gallery graves that is distinct from that in adjacent regions, but shows a striking resemblance to similar monuments in the Paris Basin and Brittany[13,14]. Despite the central geographical location of WBC that connects cultural influences from several directions, no genomic data of human remains from WBC sites have so far been investigated.

Here, we performed a genome-wide data analysis of 42 individuals who were buried in a WBC gallery grave near the township of Niedertiefenbach in Hesse, Germany (Fig. 1), dated between 3,300-3,200 cal. BCE (Supplementary Information). In contrast to other genome-wide aDNA studies, which usually include a small number of individuals from a specific site and period, we provided a snapshot of a burial community that used the collective grave for approximately 100 years[15] (Supplementary Information). In addition to population genetic and kinship analyses, we also investigated the human leukocyte antigen (HLA) region. This approach allowed us not only to reconstruct the genetic ancestry of the WBC-associated people from Niedertiefenbach, but also to gain insights into the makeup of immunity-related genes of a Late Neolithic group.

Results

In total, aDNA extracts obtained from 89 randomly selected individuals interred in the Niedertiefenbach grave[16, 17] were subjected to shotgun sequencing. Of these, we filtered out 47 who i) had fewer than 10,000 single-nucleotide polymorphisms (SNPs) benchmarked on a previously published dataset of 1,233,013 SNPs[1,2,4] or ii) showed evidence of X-chromosomal contamination (>=5%). Thus, after quality control, datasets from 42 individuals were available for subsequent comprehensive analyses (Supplementary Table 1). aDNA damage patterns[18] were consistent with an ancient origin of the isolated DNA fragments. When we screened the sequence data for known blood-borne pathogens such as *Yersinia pestis*, *Mycobacterium tuberculosis* and *Mycobacterium leprae* with MALT[19], no signs of an infection were detected. Ten individuals were genetically determined as females and 25 as males. Although of the remaining seven individuals one was osteologically identified as female and one as male.

First, we applied principal component analysis (PCA) to project the SNP information derived from the Niedertiefenbach collective together with previously published datasets of 122 ancient populations onto a basemap calculated from 59 modern-day West-Eurasian populations[1,2,4,20]. The Niedertiefenbach individuals formed a cluster that is mainly explained by genetic variation between HG and early farmers on the first principal component (Fig. 2). However, the Niedertiefenbach samples covered a wide genetic space which reflects a high intra-population diversity. Some of the individuals grouped closely with those from the Blätterhöhle, a cave site near Hagen, Germany (4,100-3,000 BCE)[6] (Fig. 2). ADMIXTURE analysis[21] with four to eight components suggested two main genetic

contributions to the Niedertiefenbach collective – one maximized in European HG and the other in Neolithic farmers from Anatolia (Fig. 3; Supplementary Fig. 1). Next we applied *f3* outgroup statistics[22] to calculate the amount of shared genetic drift between the Niedertiefenbach sample and another test population relative to an outgroup [f_3*(Niedertiefenbach; test; Mbuti)*]. The highest amount of shared genetic drift was observed between Niedertiefenbach and European HG from Sicily, Croatia, and Hungary (Supplementary Fig. 2). To estimate the amount of Neolithic farmer and HG genetic ancestry in the Niedertiefenbach group, we ran qpADM[22]. We obtained feasible models for Niedertiefenbach as a two-way mixture of Neolithic farmers from Anatolia and various European HG, which altogether gave on average ~60% farmer and ~40% HG ancestry (Supplementary Table 2). Another feasible two-way admixture model for Niedertiefenbach was as the combination of Anatolian Neolithic farmers (41%) and individuals from the Blätterhöhle. We then applied ALDER[23] to estimate the date of admixture. We observed significant results for a mixture of components associated with early farmers and Loschbour HG (Waldbillig, Luxembourg) in the Niedertiefenbach population 14.85 +/- 2.82 generations before the ^{14}C benchmark of 3,300-3,200 cal. BCE (Supplementary Note 1). Based on a generation time of 29 years[24], the date for the emergence of the genetic composition of the Niedertiefenbach community appears to be between 3,850-3,520 cal. BCE. However, the dates are based on the idealized model of a single wave of admixture between Anatolian Neolithic farmers and Loschbour HG. These populations were used as closest unadmixed genetic proxies for possible parental sources based on the qpADM results. The models do not take into consideration multiples waves, continuous admixture or admixture of populations that were already admixed[3]. Thus, the obtained dates reflect only the minimal number of generations.

For phenotype reconstruction, we investigated selected SNPs associated with skin pigmentation and hair color (rs16891982), eye color (rs12913832), starch digestion (rs11185098) and lactose tolerance (rs4988235)[25]. Not all of these SNPs were available for all of the investigated individuals due to poor sequence coverage. Fourteen of the 42 individuals carried only the rs16891982-C allele, which is associated with dark hair and increased skin pigmentation[26], while three had both alleles (C and G). Only three individuals carried the rs12913832-G allele associated with blue eye colour, seven had the A allele associated with brown eye colour, and eight had both alleles. The minor A allele of rs11185098 is positively associated with AMY1 (amylase 1) gene copies and high amylase activity, responsible for starch digestion[27]. Only one individual was found to be homozygous for the G allele and six had both alleles, while no homozygous carrier for the A allele was found. Interestingly, all individuals with enough coverage for the rs4988235 SNP carried the G-allele that tags an ancestral haplotype associated with lactose intolerance[28], which suggests that the Niedertiefenbach people could not digest dairy products.

To determine the HLA class I and II alleles of the Niedertiefenbach individuals, we applied a previously developed method[29]. In addition, we used OptiType, an automated HLA-typing tool[30]. Only alleles that were consistently called by both methods were considered for the analysis. We successfully genotyped HLA A, B, C, DPB1, DQB 1 and DRB1 alleles in 23 unrelated individuals (Supplementary Table 3). Among the HLA class I alleles, we noted the highest frequencies for A*02:01 (~63%), B*27:05 (~23%) and C*02:02 (~17%). For the HLA

class II alleles, we obtained the highest frequencies for DPB1*02:01 (~37%), DQB1*03:01 (~41%) and DRB1*08:01 (~28%).

We noted 28 different mitochondrial DNA (mtDNA) and 9 Y-chromosome haplotypes (Supplementary Table 1). Interestingly, 9 of 25 males carried the same Y-chromosome haplotype.

We performed kinship analyses using f3 statistics[22] and READ[34]. Both programmes identified one triplet consisting of a female and two males as first-degree relatives (Supplementary Fig. 5). Parent-child relationships can be ruled out as all three individuals died in infancy (at age 1-3 years) or early childhood (4-6 years). This leaves the sibling constellation as the only other possible explanation, which is supported by the respective mtDNA and Y-chromosome haplotypes as well as HLA allele profiles.

Discussion

It has clearly been established that the transformation from the LBK, which is characterised by a homogeneous material culture over a large area, to the later more diverse Neolithic cultures in Europe was accompanied by turnovers in the genomic record[3]. However, the population interactions underlying this transformation have not yet been fully analysed. The admixture events were geographically highly localized and involved various populations with different ancestry components[3]. These processes likely led to the increase in HG ancestry proportions and mtDNA lineages that were observed in Middle to Late Neolithic communities[1,7]. It is currently not known what might have influenced these wide-spread demographic and genomic processes in Europe, but climate change and/or social processes may be considered contributing factors[34].

Here, we investigated a community of 42 Late Neolithic farmers excavated from the WBC gallery grave in Niedertiefenbach, Germany[16,17,36]. As expected, the studied population exhibited a mixture of genomic components from western HG and early farmers. The continuous range (34-58%) of the relatively high genetic HG proportion in the Niedertiefenbach collective is surprising. Admixture dating indicated that the mixing of the two components started 14.85 +/- 2.82 generations before 3,300-3,200 cal. BCE. From these results, it cannot be inferred to what extent the contributing populations themselves were already admixed or which subsistence economy they practised. But interestingly, the estimated admixture date of 3,770-3,600 cal. BCE coincided with farming expansion phases and social changes during the Late MC (3,800-3,500 BCE)[37]. Archaeologically, there is a well-documented continuity from Late MC to WBC[9]. MtDNA data from two MC sites in France[38]

and Germany[39] indicate that the analysed individuals belonged to an already admixed population comprising haplotypes typical of both farmers and HG[38]. Human genomic datasets from clear archaeological MC contexts are not available as yet. A possible exception could be the genome-wide data of four individuals from the Blätterhöhle that may be chronologically (based on their radiocarbon dates of 4,100-3,000 BCE) and geographically linked with Late MC and/or WBC[6]. However, it has to be kept in mind that the remains were found in a cave without any definite cultural assignment. Our analyses showed that the Niedertiefenbach population appeared most closely related to the Blätterhöhle collective. In particular, their large HG components (39-72%)[3] fell into the range observed for Niedertiefenbach. Moreover, they were the best proxies for the HG and farmer components of the Niedertiefenbach sample. In addition, our ALDER admixture date is very similar to the one obtained for the Blätterhöhle that yielded 18-23 generations before the average sample date of 3,414 +/- 84 cal. BCE[3]. Thus, there is a probable genetic link between the people buried in the Blätterhöhle and those in the gallery grave of Niedertiefenbach.

The WBC-associated population in Niedertiefenbach represents a genetically diverse group with a very broad spectrum in their HG proportions. This finding suggests that the admixture was still in progress at that time or had taken place a few generations before. This scenario is tentatively supported by the ALDER analysis (admitting that its admixture date may be biased towards a more recent time point). Given the surprisingly large HG component, it seems conceivable that the admixture included also individuals who had an exclusive or near-exclusive genetic HG ancestry. Taking into account all available lines of evidence, we

hypothesize that the increase in the HG component likely occurred during the consolidation of the MC and/or the beginning of the WBC and could have involved also direct gene-flow from unadmixed local western HG into expanding farming populations.

The genetic data of the Niedertiefenbach sample, along with information obtained from archaeological and osteological analyses, sheds light on the community that used this gallery grave. In total, the skeletal remains of a minimal number of 177 individuals were recovered from the 7 m^2 site, reflecting a very high occupancy rate for a collective WBC burial[35]. The sex distribution in the sample, which was assessed based on diagnostic skull elements, was similar to those described for other prehistoric populations[40]. Regarding age, we did not observe a numerical deficit of children that is often recorded for Neolithic cemeteries in Germany[36,41,42]. Thus, it is likely that the skeletal population of Niedertiefenbach represented a demographic cross-section of the group that was associated with this gallery grave. The phenotype reconstruction revealed that the examined individuals had a predominantly dark complexion and were genetically not yet adapted to digest starch-rich food or lactose. These phenotypes have typically been described for HG and early farmers[3].

Overall, the genomic data indicate that the gallery grave was mainly used by not closely related people who may have lived in various neighbouring locations. This observation is supported by the large number of mtDNA (28) and Y-chromosome (9) haplotypes. However, also directly related individuals were interred. In one case, we observed inhumations of

first-degree relatives (Supplementary Fig. 4). In addition, the presence of one frequent Y-chromosome haplotype suggests a patrilineage.

In line with studies investigating the health status of Neolithic populations in central Europe[43], the Niedertiefenbach individuals showed numerous unspecific skeletal lesions that could be indicative of physical stress, including malnutrition, and infections[36]. Interestingly, we did not observe any pathogens. This observation is consistent with aDNA-based findings describing only relatively few sporadic cases of infectious diseases for the Neolithic period[44]. Noteworthy is the absence of *Yersinia pestis*, as lineages of this bacterium have already been postulated for the Late Neolithic and are reported in a Scandinavian case dated to 2900 cal. BCE[45]. If it were present, we would have expected to detect the pathogen, given the good preservation of endogenous DNA in the samples.

The HLA class I and II dataset generated for Niedertiefenbach was relatively small and thus precluded thorough statistical analysis. However, relative to contemporary European populations some striking shifts in allele frequencies could be seen (Supplementary Table 3). Interestingly, the majority of these alleles (e.g. B*51:01, DQB1*03:01) are today associated with higher resistance to viral pathogens (e.g. HIV, HCV, influenza A) and higher susceptibility to bacterial infections or complications thereof[46,47,48,49]. This observation strengthens the hypothesis that ancient epidemics influenced the present-day frequency of variants associated with modern inflammatory diseases[46,47]. Later on it may have lost its relative fitness advantage, for example because pathogens adapted to this most common

allele in a process of negative frequency-dependent selection[50], and was replaced by alleles beneficial against newly emerging human pathogenic bacteria, such as *Y. pestis*.

Further, notable difference concerns the HLA allele DRB1*15:01. It is widespread in present-day Europeans (ca. 15%), but absent in Niedertiefenbach samples. This allele predisposes to mycobacterial infections (tuberculosis, leprosy)[51]. In disease studies, the SNP allele rs3135388-T is often used as a marker for DRB1*15:01[52]. In the published aDNA datasets[25], rs3135388-T was also found to be absent in all European Palaeolithic, Mesolithic and Neolithic populations analysed. It seemed to appear for the first time only during the Bronze Age. Since then, its initially high frequency (approx. 20%) has decreased to the present levels (Supplementary Fig. 6). This finding raises the intriguing possibility that the allele might have been incorporated into the European gene pool as part of the steppe-related ancestry component in the Final Neolithic and Bronze Age.

The advent of farming and subsequent shifts in pathogen exposure are thought to have radically changed the immune genes in early agriculturalists[25]. The immune response of the Niedertiefenbach collective was primarily geared towards fighting viral agents. To what extent this antiviral profile was due to the specific demographic history of the Niedertiefenbach population or was typical of Neolithic communities in general remains to be clarified. Together, our study detected large differences in the HLA variation and immune responsiveness over the last 5300 years in Europe.

By applying a comprehensive genomics approach to individuals interred in the WBC-associated collective burial in Niedertiefenbach, we discovered that the community, which

used this site for about 100 years, was genetically heterogeneous and carried both Neolithic and HG ancestry components. The mixture of these two components likely occurred at the beginning of the 4th millennium, indicating important demographic and cultural transformations during that time in western Europe. This event may also have affected the immune status of the admixed population and its descendants for generations to come.

Methods

Samples

The archaeological site and anthropological characteristics are described elsewhere[16,17,36].

Radiocarbon dating

Collagen was dated from 25 human bone samples, originally collected for aDNA analysis, and each attributed to a different individual. Dating was performed following standard protocols at the Leibniz Laboratory for AMS Dating and Isotope Research, Kiel (details in Supplementary Information).

aDNA extraction and sequencing

Surface contaminations from petrous bones and teeth were removed with bleach solution. Partial uracil-DNA-glycosylase treated sequencing libraries were prepared from bone powder-derived DNA extracts following previously established protocols[29]. Sample-specific index combinations were added to the sequencing libraries[53]. Sampling, DNA extraction and the preparation of sequencing libraries were performed in clean-room facilities of the Ancient DNA Laboratory in Kiel. Negative controls were taken along for the DNA extraction and library generation steps. The libraries were paired-end sequenced using 2x75 cycles on an Illumina HiSeq 4000. Demultiplexing was performed by sorting all the sequences according to their index combinations. Illumina sequencing adapters were removed and paired-end reads were merged if they overlapped by at least 11bp. Merged reads were filtered for a minimum length of 30 bp.

Pathogen screening

All samples were screened with MEGAN[54] and the alignment tool MALT[19] for their metagenomic content using parameters as described in Krause-Kyora et al. 2018[29].

Mapping and aDNA damage patterns

Sequences were mapped to the human genome build hg19 (International Human Genome Sequencing Consortium, 2001) using BWA 0.7.12[55] with a reduced mapping stringency parameter "-n 0.01" to account for mismatches in aDNA. Duplicates were removed. C to T misincorporation frequencies were obtained using mapDamage 2.0[55] in order to assess the authenticity of the aDNA fragments[18]. After the validation of terminal damage, the first two positions from the 5' end of the fastq-reads were trimmed off.

Genotyping

Alleles were drawn at random from each of the 1,233,013 SNP positions[1,2,25] in a pseudo-haploid manner using a custom script as described in Lamnidis et al. 2018[57]. Datasets were filtered for at least 10,000 SNPs to be considered for further analysis[5].

Genetic sex determination

Sexes were determined based on the ratio of sequences aligning to the X and Y chromosomes compared to the autosomes[58]. Females are expected to have a ratio of 1 on

the X chromosome and 0 on the Y chromosome, whereas males are expected to have both X and Y ratios of 0.5.

Contamination estimation and authentication

Estimation of DNA contamination was performed on the mitochondrial level using the software Schmutzi[59], and in males additionally on the X-chromosomal level by applying ANGSD[60] to investigate the amount of heterozygosity on the X chromosome.

Principal component analysis

The genotype data of the Niedertiefenbach collective was merged with previously published genotypes of 5519 ancient and modern individuals genotyped on the aforementioned 1,233,013 SNPs using the program *mergeit* from the *EIGENSOFT* package[61]. Principal component analysis (PCA) was performed using the software *smartpca*[61] projecting the genotype datasets of the Niedertiefenbach and all other ancient individuals on the principal components calculated from genotype datasets of 59 West Eurasian populations by use of the 'lsqproject' option.

ADMIXTURE analysis

Prior to ADMIXTURE analysis, we used Plink (v1.90b3.29) to filter out SNPs with insufficient coverage (0.999) and a minor allele frequency (maf) below 5%. LD pruning was performed to filter out SNPs at an R^2 threshold of 0.4 using a window size of 200 and a step size of 25. We ran ADMIXTURE (version 1.3.0)[20] on the same populations as used in the PCA analysis

and a number of ancestral components ranging from 4 to 12. Cross-validation was performed for every admixture model.

Admixture dating

The source code of ALDER (v1.03)[23] was modified to decrease the minimal number of samples needed for the analysis, following a suggestion from this work: https://www.diva-portal.org/smash/get/diva2:945151/FULLTEXT01.pdf. In so doing also reference populations with only a single individual could be included.

The following reference populations were used for Niedertiefenbach: Anatolia_Neolithic, OrienteC_HG, Croatia_Mesolithic_HG, Bichon, Blatterhohle_MN, Koros_Hungary_EN_HG, Serbia_HG, Serbia_Mesolithic_Neolithic, Narva_LT, Iron_Gates_HG, Loschbour, Iberia_HG, Latvia_EN, Baalberge_MN France_MN, Latvia_HG.

To calculate calender dates of admixture we multiplied the obtained the average ALDER generation time for two-way admixture models with significant LD-decay curves with an assumed generation time of 29 years[24].

F3 outgroup statistics

f3 outgroup statistics were run as a part of the *Admixtools* package[22] in the form of f_3(*Niedertiefenbach; test, Mbuti*) using for *test* the same populations as in the PCA and ADMIXTURE analysis.

qpADM analysis

qpADM analysis was run on transition-filtered genotypes that were previously prepared for ADMIXTURE analysis as described above. We ran 48 different combination models of Niedertiefenbach as a two-way admixture, since three-way admixture models appeared to be less feasible indicating that the 3rd component was excessive. The following populations were used as outgroups: Mbuti, Ust Ishim, Kostenki14, MA1, Han, Papuan, Onge, Chukchi and Karitiana.

Kinship analysis

Kin relatedness was assessed using READ[34] and lcMLkin[62]. READ identifies relatives based on the proportion of non-matching alleles. lcMLkin infers individual kinship from calculated genotype likelihoods. A pair of individuals was regarded related only if evidence of relatedness was independently provided by both programs (Supplementary Fig. 5).

Determination of mitochondrial and Y chromosome haplotypes

Sequencing reads were mapped to the human mitochondrial genome sequence rCRS[63]. Consensus sequences were generated in Geneious (v. 9.1.3) using a default threshold of 85% identity among the covered positions and a minimum coverage of 3. HAPLOFIND[64] was applied to assess mitochondrial haplotypes from the consensus sequences and yHaplo[65] to determine Y chromosome haplotypes in male individuals.

Calling of phenotypic SNPs

We generated a pile-up of reads mapping to the positions of the selected phenotypic SNPs with samtools mpileup (v. 1.3) in order to see how many reads supported which allele for each individual.

HLA typing and analysis

We used a previously established HLA capture and HLA typing pipeline[29]. In addition, we applied OptiType[30] for automated HLA class I and II typing. We then removed one in a pair of datasets of directly related individuals (1st and 2nd degree relatives) based on the maximum number of reads supporting the HLA call in either of the related individuals. Samples with low coverage of the HLA region were also excluded. Only alleles that were consistently called by both methods were considered for the analysis. For comparing the ancient HLA allele pool with a representative modern allele pool, we used a cohort of 3,219 healthy northern German individuals and imputed HLA genotypes at 2nd field level of resolution from high-density SNP data following an established procedure[66].

Data availability. The aligned sequences are available through the European Nucleotide Archive under accession number XXXXXXXXX.

Acknowledgements

This study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through Projektnummer 2901391021 SFB 1266 to B.K.-K. and A.N., grant LE 2593/3-1 to T.L.L. and Germany`s Excellence Strategy – EXC 2167-390884018. AN was supported by the Dorothea Erxleben Female Investigator Award of the DFG Cluster of Excellence Inflammation at Interfaces (EXC306). We are grateful to Stephan Schiffels, Thiseas C. Lamnidis and Alissa Mittnik from the Max Planck Institute for the Science of Human History in Jena for sharing their expertise in ALDER, qpADM and kinship analyses and their advice on result verification. We acknowledge financial support by Land Schleswig-Holstein within the funding programme Open Access Publikationsfonds. J.B. and F.P. were funded by the International Max Planck Research School for Evolutionary Biology.

Author contributions

B.K.-K., A.N. and Ch.R. conceived and designed the research. K.F. analyzed the human skeletal remains. L.B. and B.K.-K. generated ancient DNA data. A.I., J.S. and B.K. analyzed the ancient DNA data. A.I., A.S., F.P., A.F., L.B., J.D., J.B., D.E., J.Ch.K., R.B., O.K., I.M., T.L., A.F., J.K. and B.K.-K. analyzed modern and ancient HLA data. S.Sch., J.M., Jo.M., Ch.R., C.D., M.F., A.N. and B.K.-K. interpreted the findings. A.N., A.I. and B.K.-K. wrote the manuscript with input from all other authors.

Conflict of Interests

The authors declare no conflict of interests.

References

1. Haak, W. *et al.* Massive migration from the steppe was a source of Indo-European languages in Europe. *Nature* **522**: 207-211 (2015).
2. Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**: 409-413 (2014).
3. Lipson, M. *et al.* Parallel paleogenomic transects reveal complex genetic history of early European farmers. *Nature* **551**: 368-372 (2017).
4. Mathieson, I. *et al.* The Genomic History of Southeastern Europe. *Nature* **555**: 197-203 (2018).
5. Olalde, I. *et al.* The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature* **555**: 190-196 (2018).
6. Bollongino, R. *et al.* Years of Parallel Societies in Stone Age Central Europe. *Science* **342**: 479-481 (2013).
6. Brandt, G. *et al.* Ancient DNA Reveals Key Stages in the Formation of Central European Mitochondrial Genetic Diversity. *Science* **342**: 257-261 (2013).
7. Siegmund, F. Schnelle Zeiten – langsame Zeiten: Archäologische Chronologiesysteme als Geschichtsquelle. *Arch Inf* **35**, 259–270 (2012).
8. Raetzel-Fabian, D. Revolution, Reformation, Epochenwechsel? Das Ende der Kollektivgrabsitte und der Übergang von der Wartberg- zur Einzelgrabkultur in Nordhessen und Westfalen. *J Neolith Archaeol* **4** (2002).
9. Schierhold, K. Studien zur Hessisch-Westfälischen Megalithik: Forschungsstand und -perspektiven im europäischen Kontext, VML, Rahden (2012).
11. Geschwinde, M. & Raetzel-Fabian, D. EWBSL: eine Fallstudie zu den jungneolithischen Erdwerken am Nordrand der Mittelgebirge (VML, 2009).

12. Raetzel-Fabian, D. Calden. Erdwerk und Bestattungsplätze des Jungneolithikums. Architektur - Ritual - Chronologie. Universitätsforschungen zur Prähistorischen Archäologie (70) (Dr. Rudolf Habelt, 2000).
13. Schrickel, W. Westeuropäische Elemente im neolithischen Grabbau Mitteldeutschlands und die Galeriegräber Westdeutschlands und ihre Inventare. Beiträge zur ur- und frühgeschichtlichen Archäologie des Mittelmeer-Kulturraumes (4 u. 5) (Dr. Rudofl HAbelt 1966).
14. Raetzel-Fabian, D. Absolut Chronology and Cultural Development of the Neolithic Wartberg Culture in Germany. *J Neolith Archaeol* **4** (2002).
15. Meadows, J. et al. High-precision Bayesian chronological modeling on a calibration plateau: the Niedertiefenbach gallery grave. Radiocarbon. (in review).
16. Czarnetzki, A. Die menschlichen Skelettreste aus vier neolithischen Steinkisten Hessens und Niedersachsens (Diss. Tübingen, 1966).
17. Wurm, K., Schoppa, H., Ankel, C. & Czarnetzki, A. Die westeuropäische Steinkiste von Niedertiefenbach, Oberlahnkreis. *Fundber Hessen* **3**, 46-78 (1963).
18. Briggs, A. *et al*. Patterns of damage in genomic DNA sequences from a Neandertal. *Proc Natl Acad Sci USA* **104**: 14616-14621 (2007).
19. Vågene, Å. J. *et al.* Salmonella enterica genomes from victims of a major sixteenth-century epidemic in Mexico. *Nat Ecol Evol* **2**: 520-528 (2018).
20. Jones, E. *et al.* The Neolithic Transition in the Baltic Was Not Driven by Admixture with Early European Farmers. *Curr Biol* **27**: 576-582 (2017).
21. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* **19**: 1655-1664 (2009).
22. Patterson, N. *et al.* Ancient Admixture in Human History. *Genetics* **192**: 1065 -1093 (2012).
23. Loh, PR. *et al.* Inferring admixture histories of human populations using linkage disequilibrium. *Genetics* **193**: 1233-1254 (2013).

24. Fenner, JN. Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *Am J Phys Anthropol* **128**: 415-423 (2005).

25. Mathieson, I. *et al.* Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**: 499-503 (2015).

26. Han, J. *et al.* A Genome-Wide Association Study Identifies Novel Alleles Associated with Hair Color and Skin Pigmentation. *PLoS Genet* **16**: e1000074 (2008).

27. Heianza, Y. *et al*. Starch Digestion–Related Amylase Genetic Variant Affects 2-Year Changes in Adiposity in Response to Weight-Loss Diets: The POUNDS Lost Trial. *Diabetes* **66**: 2416-2423 (2017).

28. Enattah, NS. *et al.* Evidence of Still-Ongoing Convergence Evolution of the Lactase Persistence T-13910 Alleles in Humans. *Am J Hum Genet* **81**: 615-625 (2007).

29. Krause-Kyora, B. *et al.* Ancient DNA study reveals HLA susceptibility locus for leprosy in medieval Europeans. *Nat Commun* **9**:1569 (2018).

30. Szolek, A. *et al.* OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics* **30**:3310-3316 (2014).

31. Mathieson, S. & Mathieson, I. FADS1 and the Timing of Human Adaptation to Agriculture. *Mol Biol Evol* **35**: 2957-2970 (2018).

32. Pierini, F. & Lenz, TL. Divergent allele advantage at human MHC genes: signatures of past and ongoing selection. *Mol Biol Evol* **35**: 2145-2158 (2018).

33. Arora, J. *et al.* HLA heterozygote advantage against HIV-1 is driven by quantitative and qualitative differences in HLA allele-specific peptide presentation. *Mol Biol Evol* doi:10.1093/molbev/msz249 (2019).

34. Kuhn, M., Jakobsson, M. & Günther, T. Estimating genetic kin relationships in prehistoric populations. *PLoS One* **13**: e0195491 (2018).

35. Warden, L. *et al.* Climate induced human demographic and cultural change in northern Europe during the mid-Holocene. *Sci Rep* **7**: 15251 (2017).

36. Rinne, C. *et al.* Niedertiefenbach. Ein Galeriegrab der spätneolithischen Wartberggruppe südwestlich von Niedertiefenbach (Landkreis Limburg-Weilburg, Hessen). *Praehistorische Zeitschrift* **91**: 284-316 (2016).

37. Seidel, U. *et al.* Die Zeit der großen Gräben: Modelle zur Chronologie des Michelsberger Fundplatzes von Heilbronn-Klingenberg „Schlossberg", Stadtkreis Heilbronn, Baden-Württemberg. *Praehistorische Zeitschrift* **91**: 225-283 (2016).

38. Beau, A. *et al.* Multi-scale ancient DNA analyses confirm the western origin of Michelsberg farmers and document probable practices of human sacrifice. *PLoS One* **12**: e0179742 (2017).

39. Keller, M. *et al.* United in death-related by blood? Genetic and archeometric analyses of skeletal remains from the neolithic earthwork bruchsal-aue. *Am J Phys Anthropol* **157**: 458-471 (2015).

40. H. Bach, H. & Bach, A. Paläanthropologie im Mittelelbe-Saale-Werra-Gebiet. Beiträge zur Rekonstruktion der biologischen Situation ur- und frühgeschichtlicher Bevölkerungen. Weimarer Monographien zur Ur- und Frühgeschichte (23) (Weimar 1989).

41. Alt, K. et al. Jechtingen – Anthropologie eines mittelneolithischen Gräberfeldes. *Fundberichte aus Baden-Württemberg* **34**: 177-298 (2014).

42. Grupe, G. & Herrmann, B. Die Skelettreste aus dem neolithischen Kollektivgrab von Odagsen, Stadt Einbeck, Ldkr. Northeim. Nachr Niedersachsens Urgesch **55**: 41–91 (1986).

43. Wittwer-Backofen, U. & Tomo, N. From health to civilization stress? In search for traces of a health transition during the early Neolithic in Europe. In: Bocquet-Appel, J-P. & Bar-Yosef, O. (eds). The Neolithic Demographic Transition and Its Consequences. 501-538 (Springer, 2008).

44. Fuchs, K. *et al.* Infectious diseases and Neolithic transformations: Evaluating biological and archaeological proxies in the German loess zone between 5500 and 2500 BCE: *Holocene* doi.org/10.1177/0959683619857230 (2019).

45. Rascovan, N. *et al.* Emergence and Spread of Basal Lineages of Yersinia pestis during the Neolithic Decline. *Cell* **176**: 295-305 (2019).

46. Bodis, G., Toth, V. &. Schwarting, A. Role of Human Leukocyte Antigens (HLA) in Autoimmune Diseases. *Rheumatol Ther* **5**: 5-20 (2018).

47. Ferreira, A., Monteiro, M., Vita, P., Marinho, A., & Vasconcelos, C. Post-Infectious Arthritis and Reactive Arthritis. In: Schoenfeld, Y., Agmon-Levin, N. & Rose, N. Infection and Autoimmunity (Second Edition), 807-828 (Elsevier, 2015).

48. Huang, J. *et al.* Association of HLA DQB1*03:01 and DRB1*11:01 with spontaneous clearance of hepatitis C virus in Chinese Li ethnicity, an ethnic group genetically distinct from Chinese Han ethnicity and infected with unique HCV subtype. *J Med Virol* **91**:1830-1836 (2019).

49. McLaren, PJ. & Carrington, M. The impact of host genetic variation on infection with HIV-1. *Nat Immunol* **16**: 577-583 (2015).

50. Lenz, TL. Adaptive value of novel MHC immune gene variants. *Proc Natl Acad Sci U S A* **115**: 1414-1416 (2018).

51. Zhang, FR. *et al.* Evidence for an association of HLA-DRB1*15 and DRB1*09 with leprosy and the impact of DRB1*09 on disease onset in a Chinese Han population. *BMC Med Genet* **10**: 133 (2009).

52. de Bakker, PIW. *et al.* A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat Genet* **38**, 1166-1172 (2006).

53. Kircher, M., Sawyer, S. & Meyer, M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res* **40**: e3 (2012).

54. Huson, D. *et al.* MEGAN Community Edition – Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLoS Comput Biol* **12**: e1004957 (2016).

55. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**: 589-595 (2010).

56. Jonsson, H. *et al.* mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**: 1682-1684 (2013).

57. Lamnidis, T. *et al.* Ancient Fennoscandian genomes reveal origin and spread of Siberian ancestry in Europe. *Nat Commun* **9**: 5018 (2018).

58. Fu, C. *et al.* The genetic history of Ice Age Europe. *Nature* **534**: 200-205 (2016).

59. Renaud, G. *et al.* Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol* **16**: 224 (2015).

60. Korneliussen, TS., Albrechtsen, A. & Nielsen, R. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**: 356 (2014)

61. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet* **2**: e190, (2006).

62. Lipatov, M. *et al.* Maximum Likelihood Estimation of Biological Relatedness from Low Coverage Sequencing Data. BioRxiv; doi: https://doi.org/10.1101/023374 (2015).

63. Andrews, R. M. *et al.* Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* **23**: 147 (1999).

64. Vianello, D. *et al.* HAPLOFIND: a new method for high-throughput mtDNA haplogroup assignment. *Hum Mutat* **34**: 1189-1194 (2013).

65. Poznik, G. Identifying Y-chromosome haplogroups in arbitraritly large samples of sequenced or genotyped men. BioRxiv doi: https://doi.org/10.1101/088716 (2016).

66. Degenhardt, F. *et al.* Construction and benchmarking of a multi-ethnic reference panel for the imputation of HLA class I and II alleles. *Hum Mol Genet* **28**: 2078-2092 (2019).

67. Jurtz, V. *et al.* NetMHCpan-4.0: Improved Peptide–MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. J Immunol 199: 3360-3368 (2017).

68. Oksanen, J. et al. vegan: Community Ecology Package (Version R package version 2.0-3). Retrieved from http://CRAN.R-project.org/package=vegan (2012).

69. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/ (2017).

70. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol Biol Evol* **35**:1547-1549 (2018).

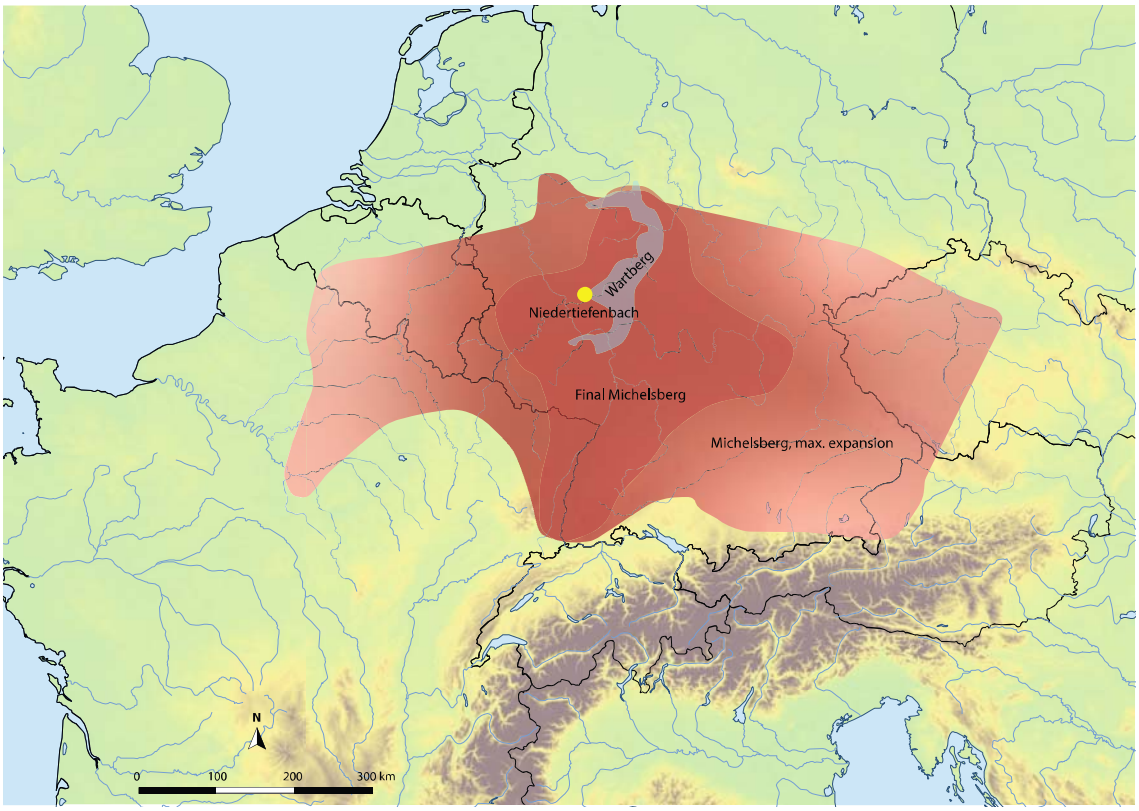

Figure 1: Map with the site Niedertiefenbach from where the individuals presented in this study were recovered. The temporal and geographic distributions of the archaeological small units mentioned in this study are shown.

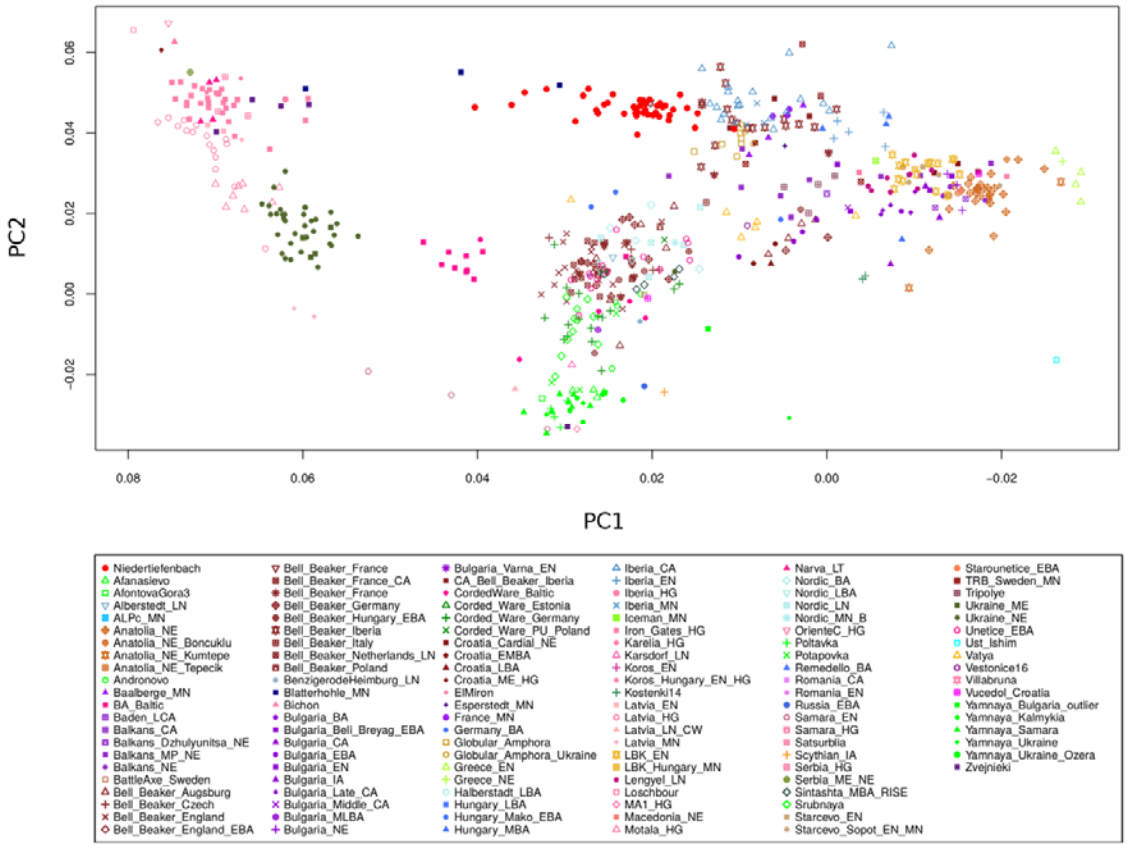

Figure 2: PCA of individuals from 123 ancient populations including Niedertiefenbach projected onto the first two principal components calculated from 59 present-day West-Eurasian populations (not shown for clarity). Niedertiefenbach individuals are depicted as red dots.

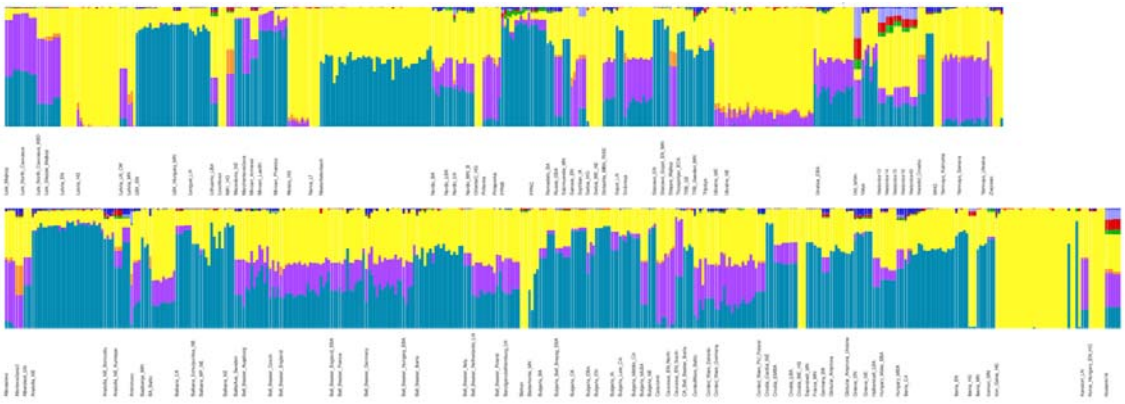

Figure 3: ADMIXTURE modeling of ancient populations including Niedertiefenbach with K=8 genetic components.