Highlights

- Population-level palaeogenetic inference on the evolution of lactase persistence from three Bronze Age regions in Europe
- Genomic data of Bronze Age warriors from Tollense, the oldest large-scale conflict site north of the Alps
- Novel method indicates that Bronze Age warriors represents an unstructured population of mainly male individuals
- Adult lactase persistence frequency in Tollense (7.1%) is significantly lower than today
- Expansion of eastern European steppe ancestry did not cause a large shift in LP allele frequency in other parts of Europe
- Selection coefficient estimate of 6% over the last 3.000 years

Genomic data from an ancient European battlefield indicates on-going strong selection on a genomic region associated with lactase persistence over the last 3,000 years

Joachim Burger*[a3], Vivian Link*[bc], Jens Blöcher*[a], Anna Schulz[d], Christian Sell[a], Zoé Pochon[bc], Yoan Diekmann[a], Aleksandra Žegarac[e], Zuzana Hofmanova[bc], Laura Winkelbach[a], Carlos S. Reyna-Blanco[bc], Vanessa Bieker[f], Jörg Orschiedt[g], Ute Brinker[h], Amelie Scheu[a], Christoph Leuenberger[i], Thomas S. Bertino[j], Ruth Bollongino[k], Gundula Lidke[l], Sofija Stefanović[e], Detlef Jantzen[h], Elke Kaiser[m], Thomas Terberger[n], Mark G. Thomas[o], Krishna R. Veeramah[j], Daniel Wegmann[bc3]

a Palaeogenetics Group Institute of Organismic and Molecular Evolution (iomE), Johannes Gutenberg University Mainz, 55099 Mainz, Germany.
b Department of Biology, University of Fribourg, 1700 Fribourg, Switzerland.
c Swiss Institute of Bioinformatics, 1700 Fribourg, Switzerland.
d Centre for the Study of Manuscript Cultures, Cluster of Excellence Understanding Written Artefacts, Hamburg University, 20354 Hamburg, Germany.
e University of Novi Sad, 21000 Novi Sad, Serbia/Laboratory of Bioarchaeology, Faculty of Philosophy, University of Belgrade, 11000 Belgrade, Serbia
f Department of Natural History, NTNU University Museum, Norwegian University of Science and Technology (NTNU), 7012 Trondheim, Norway
g Curt Engelhorn Zentrum für Archäometrie, Mannheim, Germany
h State Agency for Heritage Service of Mecklenburg-Vorpommern, Domhof 4, 19055 Schwerin, Germany
i Department of Mathematics, University of Fribourg, 1700 Fribourg, Switzerland
j Department of Ecology and Evolution, Stony Brook University, Stony Brook, New York,11794-5245, USA.
k 108a Central Road, Upper Moutere 7175, New Zealand
l Schloßstraße 50, 14059 Berlin, Germany
m Institut für Prähistorische Archäologie, Freie Universität Berlin; Fabeckstr. 23-25, 14195 Berlin, Germany
n University of Göttingen, Seminar for Pre- and Protohistory, Nikolausberger Weg 15, 37073 Göttingen, Germany
o UCL Genetics Institute, Department of Genetics, Evolution and Environment, University College London, London WC1E 6BT, UK.

Authors list footnotes
[*] These authors contributed equally to this work
[3] Corresponding authors: jburger@uni-mainz.de, daniel.wegmann@unifr.ch

Summary

Lactase persistence (LP), the continued expression of lactase into adulthood, is the most strongly selected single gene trait over the last 10,000 years in multiple human populations. It has been posited that the primary allele causing LP among Eurasians, rs4988235-A (Enattah et al. 2008), only rose to appreciable frequencies during the Bronze and Iron Ages (Mathieson et al. 2015; Olalde et al. 2018), long after humans started consuming milk from domesticated animals. This rapid rise has been attributed to an influx of peoples from the Pontic-Caspian steppe that began around 5,000 years ago (Allentoft et al. 2015; Furholt et al. 2016). We investigate the spatiotemporal spread of LP through an analysis of 14 warriors from the Tollense Bronze Age battlefield in northern Germany (~3,200 BP), the oldest large-scale conflict site north of the Alps. Genetic data indicate that these individuals represent a single unstructured Central/Northern European population. We complemented these data with genotypes of 18 individuals from the Bronze Age site Mokrin in Serbia (~4,100 to ~3,700 BP) and 37 individuals from eastern Europe and the Pontic-Caspian Steppe region, predating both Bronze Age sites (~5,980 to ~3,980 BP). We infer low LP in all three regions, i.e. in northern Germany, south-eastern, and eastern Europe, suggesting that the surge of rs4988235 in central and northern Europe was unlikely caused by Steppe expansions. We estimate a selection coefficient of 0.06, and conclude that the selection was on-going in various parts of Europe over the last 3,000 years.

Results and Discussion

The majority of the Tollense samples are male and unrelated

We enriched DNA from 21 samples from the Tollense battlefield for 5 MB of putatively neutral regions and 487 phenotypically informative loci associated with metabolic syndrome, adult lactase persistence (LP), non-infectious and inflammatory diseases, and eye, skin and hair pigmentation (Veeramah et al. 2018). Of the 21 samples, two showed evidence of high contamination (> 9% based on mtDNA) and one was not involved in the battle but rather dated to the Neolithic period. We analyzed the data of 14 of the remaining individuals for which we successfully enriched DNA at targeted regions to a mean depth > 4x (Table 1 and Supplemental Table 1). We detected no close relatives among these 14 individuals (Supplemental Figure 1). Surprisingly, two of these individuals were women, consistent with a male-dominated but not exclusively male battle.

Tollense sample shows no structure

Projected onto a principle component analysis (PCA) plot trained on modern samples, all Tollense individuals fall within the range of central and northern European variation, suggestive of little or no genetic substructure (Figure 1). Multiple lines of evidence corroborate that impression: First, the spread of he Tollense samples on the PCA matches that of other ancient population samples of similar age, namely the Lech valley in Bavaria (~4,700 to ~3,250 BP; Mittnik et al. 2019) and Mokrin in Serbia (~4,100 to ~3,700 BP; Žegarac et al. 2020). Second, only seven (0.4%) out of the 1,638 D-statistics (Green et al. 2010) of the form D (Tollense1, Tollense2; Test_Population, YRI), were statistically significant with an absolute Z-score cut-off of 3, and none were significant after correcting for multiple testing (Supplemental Table 2). Among the ten highest D-statistic values, six indicated that WEZ35 was more distant from various

European population samples than the other battlefield samples. Third, F_{st} between random partitions of the Tollense sample were generally lower than those between random partitions of a modern 1000 Genomes CEU sample of the same size ($P(F_{st}$ Tollense $< F_{st}$ CEU) = 65.57%, 65.42%, 65.65% and 65.35% for partitions of 4/10, 5/9, 6/8 or 7/7 individuals). Finally, we failed to reject Hardy-Weinberg equilibrium (HWE) using a novel method that accounts for genotyping uncertainty in ancient samples (see Methods). Specifically, we tested for a deficit of heterozygous genotypes as quantified by the inbreeding coefficient F, which is expected to be positive among samples from a structured population due to the Wahlund effect, but for which we estimated F = 0.0 (MAP, 90% CI 0.0 - $1.6 \cdot 10^{-4}$, Figure 2a). Contrasting this model with strict HWE, the latter received 99.94% posterior support. The Tollense individuals thus conform to a sample from a single, unstructured population.

Table 1. Samples analyzed and frequency of the LP allele (rs4988235-A) in different cultures. Frequencies incl. 90% Credible Intervals (CI) for NGS data were estimated based on genotype likelihoods using ATLAS. Frequencies based on pseudo-haploid (PH) calls and PCR data were estimated from allele counts taken from genotypes and a 90% CI was calculated based on the Beta distribution. Sample location as well as replication status of the LP genotype are given in Supplemental Table 3. N: number of individuals, *: family graves with high degrees of relatedness.

	Sample	Reference	Sample age [calBP]	N	13.910*T allele frequency	90% CI	Genotyping
East European (Steppe)	Eneolithic	This study	5,200-5,978	2	0	0.000-0.602	PCR
	Usatovo	This study	4,950-5,450	1	0	0.000-0.842	PCR
	Yamnaya	This study	4,480-5,378	13	0	0.000-0.132	PCR
	Early Catacomb Culture	This study	4,341-4,674	8	0	0.000-0.206	PCR
	Developed Catacomb Culture	This study	4,250-4,564	2	0	0.000-0.602	PCR
	Yamnaya-Poltavkinskaja	This study	4,450-4,884	6	0	0.000-0.265	PCR
	Late Catacomb Culture	This study	3,975-4,706	5	0	0.000-0.3085	PCR
	East European Steppe: 3600-2300 BCE	Allentoft et al. 2015, Järve et al. 2019, Lipson et al. 2017, Mathieson et al. 2015, Mathieson et al. 2018, Narasimhan et al. 2019, Olalde et al. 2018, Schroeder et al. 2019, Wang et al. 2019	4,277-5,471	37	0	0.000-0.040	NGS (PH)
	Area of Corded Ware Culture: 2900-2300 BCE	Allentoft et al. 2015, Fernandes et al. 2018, Gamba et al. 2014, Jones et al. 2017, Mathieson et al. 2015, Mathieson et al. 2018, Mittnik et al. 2019, Narasimhan et al. 2019, Olalde et al. 2018, Sanchez-Quinto et al. 2019, Schroeder et al. 2019, Skoglund et al. 2014	4,250- 4,833	55	0,018	0.000-0.027	NGS (PH)
Rest of the continent	Mokrin (Bronze Age Serbia)	Žegarac et al. 2020	4,100- 3,700	18*	0,046	0.001-0.145*	NGS
	Tollense	This study	3,100-3,200	14	0,071	0.009-0.235	NGS
	Prague (Jinonice, Zahradnictví and Kobylisy, Ke Stírce Street)	Olalde et al. 2018	4,200-3,700	14	0,101	0.031-0.227	NGS
	Bedfordshire, Biddenham Loop	Olalde et al. 2018	3,130-3,206	6	0,140	0.028-0.358	NGS
	Lichtenstein Cave, Late Bronze Age in Germany	Schilz 2006; Seidenberg 2016	3,000-2,700	34*	0,294	0.2 – 0.42*	PCR
	Kivutkalns, Baltic Bronze Age in Latvia	Mittnik et al. 2018	2,730-2,560	8	0,575	0.364-0.77	NGS

	Early Medieval Bavarians	Veeramah et al. 2018	~1,500	21	0,524	0.4-0.647	NGS
	Szolad (northern ancestry), Early Medieval in Hungary	Amorim et al. 2018	1,500	13	0,727	0.567-0.856	NGS

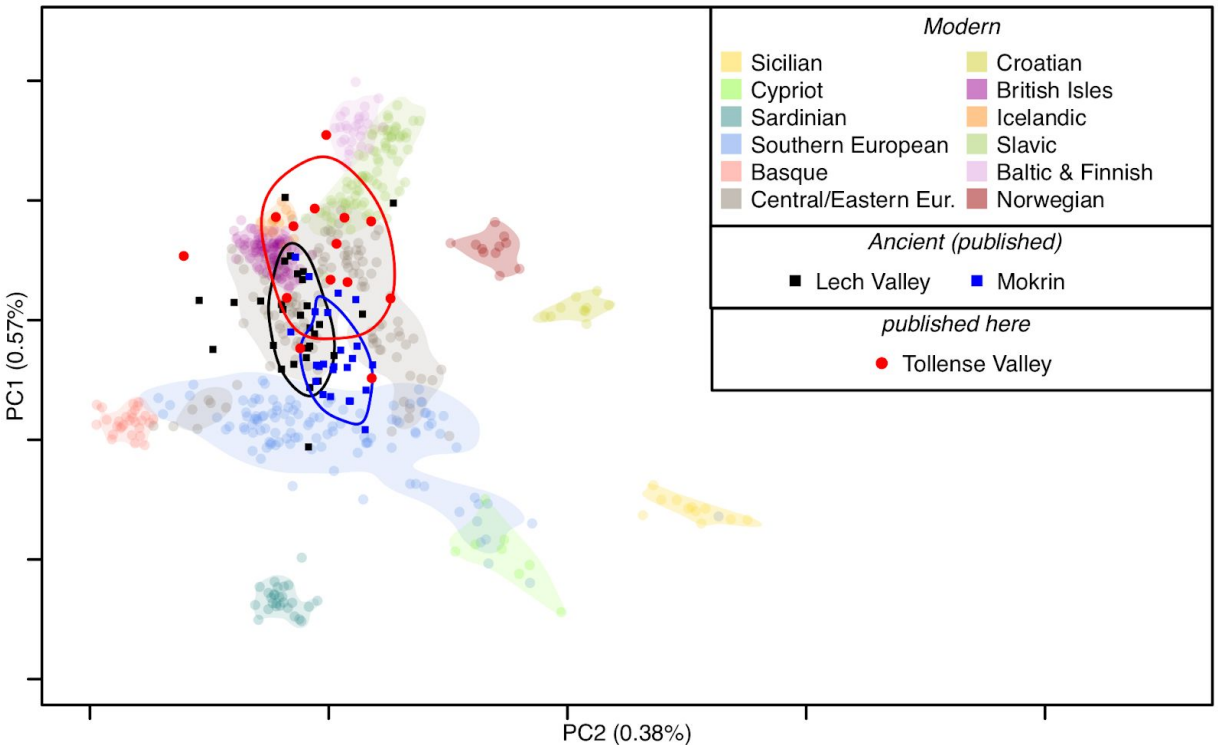

Figure 1: PCA of modern and ancient Europeans. Black squares: ancient samples from the Lech Valley in Bavaria, blue squares: ancient samples from Mokrin in Serbia, red circles: Tollense samples from this study. Ancient samples are projected onto modern reference sample space. Colored circles correspond to the 70% density contour lines of the ancient groups.

Tollense population is from Central / Northern Europe

The ancestry of the 14 Tollense individuals is central to northern European, as indicated by several analyses. First, the 70% density contour in the PCA shows a slightly more northern center for the Tollense sample than for those of the Bavarian Lech Valley and Serbian Mokrin individuals (Figure 1). Second, the Tollense sample appears closest to 5th century Bavarians (Veeramah et al. 2018) and modern central and northern Europeans when quantified with F_{ST}, although all European populations have overlapping confidence intervals (Figure 2b). Finally, the same clustering was obtained with a Treemix analysis (Figure 2c, Supplemental Figure 2) at the maximum likelihood estimate of five migration events, of which an event between the Finnish and an east Asian population was the only one involving Europeans.

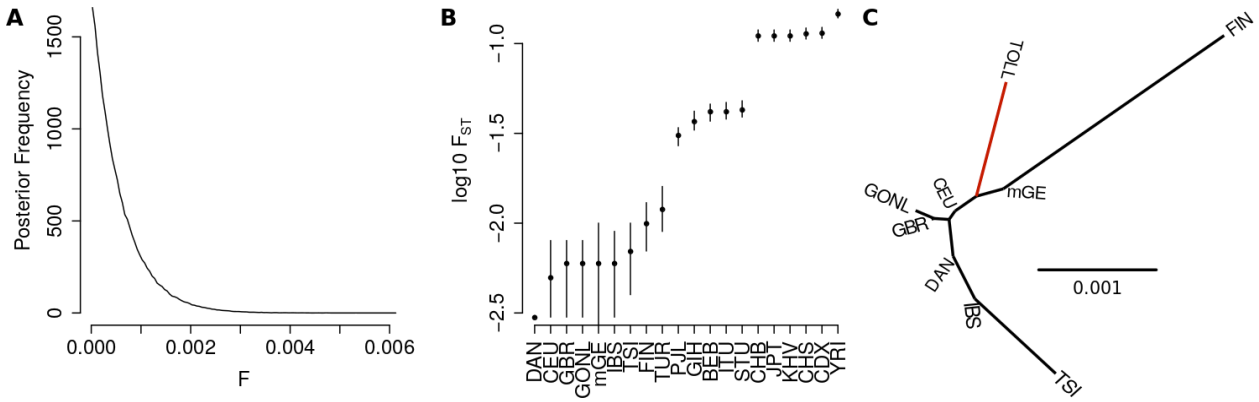

Figure 2: Tollense samples show no population structure. A) Posterior density of inbreeding coefficient F. B) F_{ST} between Tollense samples at 5K neutralome against 5th century Bavarians (mGE), Dutch (GONL), Danish, Turkish (TUR) and Eurasian+YRI 1000 Genomes populations, excluding transition SNPs. C) TreeMix analysis with zoom on European populations allowing five migration events with an unrooted tree. CEU: Utah residents with northern and western European ancestry

Based on archeological finds at the battlefield, it has been hypothesized that some of the warriors came from southern central Europe, i.e. south-eastern Germany and Bohemia (Uhlig et al. 2019; Terberger et al. 2014; Jantzen 2011). Carbon and Strontium isotope ratios further suggest a mixture of locals and non-locals (Price et al. 2019). While it is possible that our data are insufficient to resolve genetic affinities at such a small geographic scale, it appears likely that the Tollense individuals were sampled from a relatively homogeneous population with a high degree of continuity to those living in the same broad region today. But we caution that a strong correspondence between the material culture and genetic ancestry is only expected under conditions of strict and enduring population separation.

Major allele frequency increase before the Medieval period

To compare LP frequencies temporally and spatially, we determined rs4988235-A allele frequencies (also known as -13,910*T) using a Bayesian estimator in various ancient, post-Neolithic population samples with more than 5 individuals per site. We estimated low frequencies for both Tollense (7.1%; N=14) and Mokrin (4.6%; N=18) samples ($P(f_{Mokrin} < f_{Tollense})=0.889$). However, we caution that various individuals at the Mokrin site are close relatives. These estimates are consistent with those of other Bronze Age populations, albeit with a lack of high quality data for comparison (Table 1). For instance, data from somewhat older sites in the UK (N=6; Olalde et al. 2018) and Czechia (N=14; Olalde et al. 2018) indicate allele frequencies of 17% and 10%, respectively, which are not significantly different from those of Tollense ($P(f_{UK} < f_{Tollense})=0.225$; $P(f_{Czechia} < f_{Tollense})=0.33$). However, these samples do not originate from a single time point or location, and are based on extremely low allelic depth, few individuals, or both. The highest Bronze Age frequency (29%; N=34; Schilz 2006, Seidenberg 2016) was observed for the Lichtenstein Cave in Germany, a family grave site a few centuries younger than Tollense.

Markedly higher frequencies are observed from more recent samples, suggesting ongoing strong selection on the rs4988235-A allele during the intermediate period. A sample from Latvia, dating to about 2,730-2,560 BP ($P(f_{Latvia} < f_{Tollense})<10^{-3}$; N=8; Mittnik et al. 2018), and a Medieval sample from southern Germany (~1,500 BP; $P(f_{Germany} < f_{Tollense})<10^{-3}$; N=21; Veeramah et al. 2018) both have a frequency of roughly 57%, and a sample with northern European ancestry from an Early Medieval cemetery in Hungary has a frequency of 73% ($P(f_{Latvia} < f_{Tollense})<10^{-3}$; N=13; Amorim et al. 2018).

These frequency estimates indicate that while the rs4988235-A allele had reached frequencies that make it detectable in small sample size ancient DNA studies by the beginning of the Bronze Age in various parts of Europe, it had not reached the frequencies observed in population samples dating from the Iron Age or later. Such a pattern is consistent with selection starting as early as the Neolithic, but also indicates continuing strong selection during and particularly after the Bronze Age.

Selection is ongoing and strong after the Neolithic period

To quantify the selection strength on the rs4988235-A allele since the Bronze Age, we treated the Tollense individuals as sampled from an unstructured population with broad continuity to modern central/northern Europeans and examined allele frequency changes in the intervening 130 generations for all phenotypically relevant loci genotyped. We used the unrelated 1,000 Genomes CEU samples as a modern reference population sample and employed two complementary methods: i) a Bayesian approach (ApproxWF, Ferrer-Admetlla et al. 2016) under an additive model (dominance coefficient of h=0.5) and a constant population, the size of which we estimated based on allele frequencies at neutral loci, and ii) forward simulations under a model of exponential population growth (Wilde et al. 2014).

Both methods identify two alleles influenced by strong positive selection: the LP causing allele rs4988235-A, and rs7570971-C, which occurs almost exclusively on the same haplotype as rs4988235-A, and additionally plays a role in total serum cholesterol reduction (Zhang et al. 2013). Persuasive evidence for a functional role in LP strongly indicates that selection is acting on the rs4988235-A allele, rather than the linked rs7570971-C allele (Lewinsky et al. 2005; Leseva et al. 2018). Within 3,000 years, sample counts for these alleles increased from 2/26 to 144/192 and 2/28 to 134/190, respectively. ApproxWF estimates a selection coefficient of around 5% for rs7570971-C and 6% for rs4988235-A, also when assuming dominance (h=1, Figure 3b and c). Consistent with the low frequency we report here for the Bronze Age, these estimates are a bit higher than recent estimates from modern (e.g. Stern et al. 2019, 1.6%) and ancient data (e.g. Mathieson & Mathieson 2018, 1.8%).

The forward simulation approach additionally rejected genetic drift as the only explanation for the inferred frequency change of the derived allele at rs5743810 (p-value<0.001). This locus is located in the toll-like receptor gene complex (TLR6) and is associated with pathogen pattern recognition and innate immune response. Notably, none of the other 440 tested loci showed a significant signal of selection between the period of the Tollense site (3,200 BP) and today.

Figure 3: Lactase persistence distribution and selection in Europe. A) Map of lactase persistence in the Bronze Ages sites of Tollense and Mokrin, in Eneolithic and Early Bronze Age samples from eastern Europe and the Pontic-Caspian Steppe area, and an Early Medieval sample from Bavaria. Red corresponds to the ancestral and yellow to the derived allele associated with LP. The shaded shape encompasses the samples from the east European Steppe. B) For each phenotypic locus we show the posterior selection coefficient and the false discovery rate (FDR) as inferred with ApproxWF. Loci identified as under selection by forward simulations approach are colored. Open circles show estimates obtained under a dominant model (h=1). The dashed line indicates the 5% FDR cutoff used. C) Posterior densities estimated by ApproxWF under the additive (h=0.5, solid) and dominant (h=1, dashed) model for the loci identified as under selection by both methods.

Eastern Europeans steppes are not the source for lactase persistence

Based on imputed data, Allentoft et al. (2015) reported a low allele frequency (5%) of the rs4988235-A allele in Bronze Age Europeans, similar to that reported here, but a much higher frequency (~25%) among Bronze Age samples from the Pontic-Caspian Steppe, indicating a possible steppe origin of lactase persistence. Since imputing allele presence in ancient samples using modern reference individuals may be problematic in regions of strong recent selection, we investigated this hypothesis by genotyping the rs4988235 locus using PCR in Eneolithic and Early Bronze Age samples from eastern Europe and the Steppe region. The majority of the sampled individuals were buried in barrow graves dating from the end of 4th mill. BCE to the end of the 2nd mill. BCE and are representative for the Early Bronze Age in Eastern Europe. (Figure 3a, Table1). We could not detect the rs4988235-A allele among any of these samples (N=37), suggesting that the frequency of this allele was very low, possibly close to zero, and almost certainly lower than the 5.4% previously reported for a geographically, culturally and temporally diverse sample with “Steppe ancestry” (Mathieson et al. 2015). Additionally, we re-analyzed published data from the Eastern European steppe area (5600-4300 BP) and that of the Corded Ware Culture in Central and NE Europe (4900-4300 BP) -based on pseudo-haploid random allele picking- obtaining frequencies of 0% and 1,8% respectively; this corresponds to a single LP-associated allele in 92 individuals (Table 1). While these estimates are not directly informative about the origin of the rs4988235-A allele, they appear inconsistent with a major contribution of the Steppe-associated expansion to the high frequencies observed after the Bronze Age in Europe.

The time course of selection

While the LP-causing rs4988235-A allele has been under very strong natural selection at the broad geographic scale of western Eurasia, considerable uncertainty and debate remains concerning the underlying drivers and spatiotemporal distribution of that selection, and the role demographic processes played in shaping allele frequencies (Walker and Thomas 2019). When considering the underlying drivers of selection, it is important to recognize that the date or origin of the rs4988235-A allele, the timing of selection on LP (whether constant from some point in the past, or episodic), and the first observation of that allele in ancient DNA data, are distinct and probably all separated by thousands of years. Linkage disequilibrium studies (Bersaglieri *et* al. 2004; Gallego Romero et al. 2011; Liebert et al. 2017) indicate an allele origin in the last 20,000 years, perhaps in the Holocene. Evidence of milk consumption from analysis of fatty acids deposited on pot sherds (Craig et al. 2005; Evershed et al. 2008; Salque et al. 2013), and from archaeological herd kill-off profiles (Vigne and Helmer 2007) dates back to the Early Neolithic in Anatolia, the Levant, and south-eastern Europe. However, even if strong natural selection favoring LP has been operating since the Early Neolithic, we would expect sigmoidal allele frequency trajectories, and so low and barely detectable frequencies for up to several thousand years. Thus, low frequency estimates until the later Bronze Age (Burger et al. 2007; Mittnik et al. 2019; Olalde et al. 2018; Allentoft et al. 2015; Mathieson et al. 2015; Stern et al*.* 2019; Segurel et al. 2020; this study) are not necessarily inconsistent with selection on LP starting when or shortly after milk became a significant dietary component.

To illustrate this, we inferred the age of the rs4988235-A allele from the Tollense data using a novel Bayesian method and assuming a selection coefficient of 6% since the mutation occurred (Figure 3). The inferred allele age is highly dependent on the unknown effective population size between the time of mutation and the Bronze age, as well as the dominance coefficient h. Assuming dominance (h=1), the maximum a posteriori (MAP) estimate varied between 3550 (90% CI 3200-5100) and 7140 (90% CI 5525-7650) years BP for population sizes $2N=10^2$ and $2N=10^5$, respectively, and was about 20% older when assuming additivity (h=0.5). In line with the low frequency for the Tollense sample, these estimates are considerably younger than recent estimates from modern data (e.g. Stern et al. 2019, 17500 BP). However, we note that all such estimates make strong assumptions about the demographic history, the strength of selection, as well as the origin of the selected allele.

The Bronze Age therefore represents an important waypoint at which the rs4988235-A allele became sufficiently common to be detectable in reasonably sized ancient DNA samples. Data from the Bronze Age can then serve as a starting point to quantify selection reliably in subsequent times. The Tollense sample, consisting of individuals that most likely died in a battle within a short period of time, possibly on the same day, is well suited for this. The inferred selection coefficient of 6% between Tollense (~3,200 BP) and today is high, especially considering that the advancement of agricultural skills and the increase in dietary breadth since the Neolithic should have produced alternatives to milk consumption.

Drivers of Selection

Beyond milk being a nutrient-dense and relatively nutrient-balanced food, various explanations have been offered for the strong selection inferred for LP. These include improved calcium absorption by supplementing vitamin D-poor diets at high latitudes (Flatz and Rotthauwe 1973), the supply of a relatively pathogen-free fluid (Cook & Al-Torki 1975; Cook, 1978; Gerbault et al. 2011), the suppression of malaria symptoms through a reduction of p-aminobenzoic acid consumption (Cordain et al. 2012; Lokki et al. 2011), improvements in gut health through galactose and galacto-oligosaccharides reshaping the colonic microbiome (Cederlund et al. 2013; Gibson et al. 2017; Wahlqvist 2015; Knol et al. 2005; Walker and Thomas, 2019; Segurel

et al. 2020), avoidance of diarrhea under famine conditions (Sverrisdottir et al. 2014), and increased economic efficiency of calorie production for dairy farming (Ingold 1980). It is unlikely that any single one of these factors has acted alone over the whole period from the Neolithic to modern times. However, the inference of an on-going, and possibly increased selection coefficient from the Bronze Age to at least Medieval times could be interpreted as favoring factors related to increases in population and settlement density, such as those concerning pathogen loads (Sverrisdottir et al. 2014; Walker and Thomas 2019). In this context, it is interesting to note that among the other >400 functional loci we examined, the only other allele reaching significance in terms of a signature of selection is the derived allele at rs5743810 in the toll-like receptor gene complex (TLR6), which is associated with pathogen pattern recognition and innate immune response. Although difficult to test explicitly, there remains a strong possibility that selection on the rs4988235-A allele is modulated by other genetic factors through a pleiotropic network, possibly in relation to epidemic disease resistance in the context of milk drinking.

Demography and selection

Aside from strong natural selection, demographic processes such as migration and population range expansion – leading to allele surfing (Klopfstein et al. 2005; Edmonds et al. 2004) – will have shaped the distribution of the rs4988235-A allele (Itan et al. 2009). It is not clear if the northwestern distribution of this allele in Europe is primarily the result of these processes, or spatially structured selection strengths shaped by, for example, different traditions of milk use, different climates and ecologies, or different levels of incident ultraviolet radiation. However, it is noteworthy that there has been little genetic turnover between the Tollense population sampled in this study and populations in the same region today, yet we still infer a high selection coefficient. The same applies to the Mokrin site in Serbia, where a major population shift is equally unlikely – leaving natural selection as the main explanation for the observed allele frequency change over the last few millennia.

Final Remark

The somewhat similar inferred rs4988235-A allele frequencies between the early Middle Ages and today at different locations does not exclude the possibility of ongoing selection during that period (Montalva et al. 2019), since this likely represents the transitional phase of the sigmoid frequency trajectory expected for a dominant trait. For LP, it now seems likely that the phase of most rapid frequency rise was between 4,000 BP and 1,500 BP. We contend that research should be focused on this phase to better understand the evolutionary history of the most strongly selected single gene trait in Holocene western Eurasia, and many other parts of the world.

AUTHOR CONTRIBUTIONS

JBu, DW, TT and RB initiated this research. ASchu, AZ, LW, and RB produced data. CS, JBl, KV, VL, JBu, ZH, JO, ASche, VB, ZP, YD and CL analyzed data. JBu, VL, DW, KV, and MGT wrote the article with the help of all co-authors.

Acknowledgements

This work was funded by grants BO4119/1-1 of the German Science foundation to RB and JBu, and 31003A_173062 of the Swiss National Science foundation to DW. We thank the Genome of the Netherlands Consortium for access to BAM files. Parts of this research were conducted using the supercomputer Mogon and/or advisory services offered by Johannes Gutenberg University Mainz (hpc.uni-mainz.de), which is a member of the AHRP (Alliance for High Performance Computing in Rhineland Palatinate, www.ahrp.info) and the Gauss Alliance e.V..

Genomic data are available at the European Nucleotide Archive under the accession no. PRJEB38406 in BAM and fastq format.

REFERENCES

Auton, A., Brooks, L., D., Durbin, R., M., Garrison, E., P., Kang, H., M., Korbel, J., O., Marchini, J., L., McCarthy, S., McVean, G., A., and Abecasis, G., R. 1000 Genomes Project Consortium, A global reference for human genetic variation. Nature 526, 68–74, 2015.

Alkan, C., Kavak, P., Somel, M., Gokcumen, O., Ugurlu, S., Saygi, C., Dal, E., Bugra, K., Güngör, T., Sahinalp, S., Özören, N., and Bekpen C. Whole genome sequencing of Turkish genomes reveals functional private alleles and impact of genetic interactions with Europe, Asia and Africa. BMC Genomics, 15(1):963, 2014.

Allentoft, M., E., Sikora, M., Sjögren KG., Rasmussen, S., Rasmussen, M., Stenderup, J., Damgaard, P., B., Schroeder, H., Ahlström, T., Vinner, L., et al. Population genomics of bronze age eurasia. Nature, 522(7555):167, 2015.

Amorim, C., E., G., Vai, S., Posth, C., Modi A., Koncz, I., Hakenbeck, S., La Rocca, M., C., Mende, B., Bobo, D., Pohl W., et. al. Understanding 6th-century barbarian social organization and migration through paleogenomics. Nature communications, 9(1):3547, 2018.

Aronesty E. ea-utils: "Command-line tools for processing biological sequencing data"; https://github.com/ExpressionAnalysis/ea-utils, 2011.

Bersaglier,i T., Sabeti, P., C., Patterson, N., Vanderploeg, T., Schaffner, S., F., Drake, J., A., Rhodes, M., Reich, D., E., and Hirschhorn, J., N. Genetic signatures of strong recent positive selection at the lactase gene. The American Journal of Human Genetics, 74(6): 1111–1120, 2004.

Besenbacher, S., Liu, S., Izarzugaza, J., M., Grove, J., Belling, K., Bork-Jensen, J., Huang, S., Als, T., D., Li, S., Yadav, R. et al. Novel variation and de novo mutation rates in population-wide de novo assembled Danish trios. Nature communications, 6(1), pp.1-9, 2015.

Bhatia G., Patterson N., Sankararaman S., Price A., L. Estimating and interpreting FST: the impact of rare variants. Genome Research, 23(9):1514-21, 2013.

Bollongino R., Nehlich O., Richards M., P., Orschiedt J., Thomas M., G., Sell C., Fajkošová Z., Powell A., and Burger J. 2000 Years of Parallel Societies in Stone Age Central Europe. Science 342(6157):479-481, 2013.

Bramanti B., Thomas M., G., Haak W., Unterländer M., Jores P., Tambets K., Antanaitis-Jacobs I., Haidle M., N., Jankauskas R., .Kind C., J. et al. Genetic discontinuity between local hunter-gatherers and Central Europe's first farmers. Science 326(5949):137-140, 2009.

Brinker, U., Flohr, S., Piek, J. and Orschiedt, J. Human remains from a Bronze Age site in the Tollense Valley: victims of a battle?, in: Knüsel, C. and Smith, M.J. (eds.), The Routledge Handbook of the Bioarchaeology of Human Conflict. London: Routledge, 146-160, 2015.

Brinker, U., Schramm, A., Flohr, S., Jantzen, D., Piek, J., Hauenstein, K. and Orschiedt, J. The Bronze Age battlefield in the Tollense Valley, Mecklenburg-Western Pomerania, northeast Germany – Combat marks on human bones as evidence of early warrior societies in northern Middle Europe?, in: Coimbra, F., Delfino, D., Sirbu, V. and Schuster, C. (eds.), Later Prehistory to the Bronze Age. 1. The emergence of warrior societies and its economic, social and environmental consequences. 2. Imports and Aegeo-Mediterranean influences on the continental European tombs in the Bronze and Iron Ages. *Proceedings of the XVII UISPP World Congress* (1–7 September 2014, Burgos, Spain) 9. Oxford: Archaeopress, 39-56. 2016.

Brinker, U., Harten-Buga, H., Staude, A., Jantzen, D. and Orschiedt, J. Perimortem lesions on human bones from the Bronze Age battlefield in the Tollense Valley: An interdisciplinary approach, in: Dolfini, A., Crellin, R., Horn, C., and Uckelmann, M. (eds.), Prehistoric Warfare and Violence. Quantitative and Qualitative Approaches. Cham: *Springer*, 39-60. 2018.

Burger, J., Kirchner, M., Bramanti, B., Haak, W., and Thomas, M., G. Absence of the lactase-persistence-associated allele in early neolithic europeans. Proceedings of the National Academy of Sciences, 104(10):3736–3741, 2007.

Busing, F., M., Meijer, E., & Van Der Leeden, R. Delete-m jackknife for unequal m. *Statistics and Computing*, *9*(1), 3-8, 1999.

Cederlund, A., Kai-Larsen, Y., Printz, G., Yoshio, H., Alvelius, G., Lagercrantz, H., Strömberg, R., Jörnvall, H., Gudmundsson, G., H., and Agerberth, B. Lactose in human breast milk an inducer of innate immunity with implications for a role in intestinal homeostasis. PLoS One, 8(1):e53876, 2013.

Chang, C., C., Chow, C., C., Tellier, L., C., A., M., Vattikuti, S., Purcell, S., M., and Lee, J., J. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience, 4(1):7, 2015.

Cook GC. Breath hydrogen concentrations after oral lactose and lactulose in tropical malabsorption and adult hypolactasia. Transactions of the Royal Society of Tropical Medicine and Hygiene, 72(3):277–281, 1978.

Cook GC and Al-Torki MT. High intestinal lactase concentrations in adult arbs in saudi arabia. Br Med J, 3(5976):135–136, 1975.

Cordain, L., Hickey, M., S., & Kim, K. Malaria and rickets represent selective forces for the convergent evolution of adult lactase persistence. In *Biodiversity in Agriculture: Domestication, Evolution, and Sustainability* (pp. 299-308). Cambridge University Press, 2012.

Craig, O., E., Chapman, J., Heron, C., Willis, L., H., Bartosiewicz, L., Taylor, G., Whittle, A., and Collins, M. Did the first farmers of central and eastern europe produce dairy foods? Antiquity, 79(306):882–894, 2005.

Church, D., M., Schneider, V., A., Graves, T., Auger, K., Cunningham, F., Bouk, N., Chen, H., C., Agarwala, R., McLaren, W., M., Ritchie, G., R. et al. Modernizing reference genome assemblies. PLoS Biol, 9(7), p.e1001091, 2011.

Danecek P., Auton A., Abecasis G., Albers C. Banks A., E., DePristo M., A., R., E., Handsaker, G.Lunter, Marth G. T., Sherry S. T., McVean G., et al. The variant call format and VCFtools. Bioinformatics, 27(15):2156–2158, 2011.

Dimitrieva S. and Bucher P. UCNEbase database of ultraconserved non-coding elements and genomic regulatory blocks. Nucleic Acids Research, 41(D1):D101–D109, 2013.

Edmonds, C., A., Lillie, A., S., Cavalli-Sforza L., L. Mutations arising in the wave front of an expanding population. Proceedings of the National Academy of Sciences 101, no. 4:975-979, 2004.

Enattah, N., S., Jensen, T., G., K., Nielsen, M., Lewinski, R., Kuokkanen, M., Rasinpera, H., El-Shanti, H., Seo, J., K., Alifrangis, M., Khalil, I., F., et al. Independent introduction of two lactase-persistence alleles into human populations reflects different history of adaptation to milk culture. The American Journal of Human Genetics, 82(1):57–72, 2008.

Evershed, R., P., Payne, S., Sherratt, A., G., Copley, M., S., Coolidge, J., Urem-Kotsu, D., Kotsakis, K., Özdoğan, M., Özdoğan, A., E., Nieuwenhuyse, O., et al. Earliest date for milk use in the near east and southeastern europe linked to cattle herding. Nature, 455(7212):528, 2008.

Fernandes, D., M., Strapagiel, D., Borówka, P., Marciniak, B., Żądzińska, E., Sirak, K., Siska, V., Grygiel, R., Carlsson, J., Manica, A. et al. A genomic Neolithic time transect of hunter-farmer admixture in central Poland. Scientific reports, 8(1), pp.1-11, 2018.

Ferrer-Admetlla, A., Leuenberger, C., Jensen, J., D., and Wegmann, D. An approximate markov model for the wright–fisher diffusion and its application to time series data. Genetics, 203(2):831–846, 2016.

Flatz, G., and Rotthauwe, H., W. Lactose nutrition and natural selection. The Lancet, 302(7820):76–77, 1973.

Francioli, L., C., Menelaou, A., Pulit, S., L., Van Dijk, F., Palamara, P., F., Elbers, C., C., Neerincx, P., B., Ye, K., Guryev, V., Kloosterman, W.P. et al. Whole-genome sequence variation, population structure and demographic history of the Dutch population. Nature genetics, 46(8), p.818, 2014.

Fu, Q., Li, H., Moorjani, P., Jay, F., Slepchenko, S., M., Bondarev, A., A., Johnson, P., L., Aximu-Petri, A., Prüfer, K., de Filippo, C., Meyer, M., et al. Genome sequence of a 45,000-year-old modern human from western Siberia. Nature, 514(7523), pp.445-449, 2014.

Furholt, M., Grossmann R., and Szmyt M. Corded ware and bell beakers: A practice-based perspective on local communities, transregional interaction and social heterogeneity in late neolithic europe. Transitional Landscape, pages 117–131, 2016.

Gamba C., Jones E., R., Teasdale M., D., McLaughlin R., L., Gonzalez-Fortes G., Mattiangeli V., Domboróczki L., Kovári I., Pap I., et al. Genome flux and stasis in a five millennium transect of European prehistory. Nat Commun 5:5257, 2014.

Gallego, R., I., Basu, M., C., Liebert, A., Crivellaro, F., Chaubey, G., Itan, Y., Metspalu, M., Eaaswarkhanth, M., Pitchappan, R., Villems, R., et al. Herders of indian and european cattle share their predominant allele for lactase persistence. Molecular biology and evolution, 29(1): 249–260, 2011.

Gerbault, P., Liebert, A., Itan, Y., Powell, A., Currat, M., Burger, J., Swallow, D., M., and Thomas, M., G. Evolution of lactase persistence: an example of human niche construction. Philosophical Transactions of the Royal Society B: Biological Sciences, 366(1566): 863–877, 2011.

Gibson, G., R., Hutkins, R., Sanders, M., E., Prescott, S., L., Reimer, R., A., Salminen, S., J., Scott, K., Stanton, C., Swanson, K., S., Cani, P., D., et al.. Expert consensus document: The international scientific association for probiotics and prebiotics (isapp) consensus statement on the definition and scope of prebiotics. Nature reviews Gastroenterology & hepatology, 14(8):491, 2017.

Green, R., E., Krause, J., Briggs, A., W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M. HY., et al. A draft sequence of the neandertal genome. Science, 328(5979):710–722, 2010.

Ingold T. Hunters pastoralists and ranchers Reindeer economies and their transformations (Cambridge university press, Cambridge) p 336, 1980.
Itan, Y., Powell, A., Beaumont, M., A., Burger, J., and Thomas M., G. The origins of lactase persistence in europe. PLoS computational biology, 5(8):e1000491, 2009.
Jantzen, D., Brinker, U., Orschiedt, J., Heinemeier, J., Piek, J., Hauenstein, K., Krüger, J., Lidke, G., Lübke, H., Lampe, R., et al. A bronze age battlefield? weapons and trauma in the tollense valley, north-eastern germany. Antiquity, 85(328): 417–433, 2011.
Järve, M., Saag, L., Scheib, C., L., Pathak, A., K., Montinaro, F., Pagani, L., Flores, R., Guellil, M., Saag, L., Tambets, K., et al. Shifts in the Genetic Landscape of the Western Eurasian Steppe Associated with the Beginning and End of the Scythian Dominance. Current Biology, 29(14), pp.2430-2441, 2019.
Jones, E., R., Zarina, G., Moiseyev, V., Lightfoot, E., Nigst, P., R., Manica, A., Pinhasi, R. and Bradley, D., G. The Neolithic transition in the Baltic was not driven by admixture with early European farmers. Current Biology, 27(4), pp.576-582, 2017.
Kachalova, N., K. O sootnoshenii Yamnoi i Poltavkinskoi kultur [On the relation between Yamnaya and Poltavkinskaya Cultures] In: P. M. Munchaev (ed.) Problemy arkheologii Evrazii. K 80-letiyu N. Ya. Merperta. Moscow, 117-126, 2002.
Klopfstein, S., Currat, M., Excoffier, L. The fate of mutations surfing on the wave of a range expansion. Molecular biology and evolution, 23. Jg., Nr. 3, S. 482-490, 2005.
Knol, J., Scholtens, P., Kafka, C., Steenbakkers, J., Gro, S., Helm, K., Klarczyk, M., Schöpfer, H., Böckler, H., M., and Wells, J. Colon microflora in infants fed formula with galacto-and fructo-oligosaccharides: more like breast-fed infants. Journal of pediatric gastroenterology and nutrition, 40(1):36–42, 2005.
Kousathanas A., Leuenberger C., Link V., Sell C., Burger J. Wegmann D. Inferring Heterozygosity from Ancient and Low Coverage Genomes. Genetics, 205(1):317-332, 2017.
Lazaridis, I., Nadel, D., Rollefson, G., Merrett, D., C., Rohland, N., Mallick, S., Fernandes, D., Novak, M., Gamarra, B., Sirak, K., Connell, S., et al. Genomic insights into the origin of farming in the ancient Near East. Nature, 536(7617), pp.419-424, 2016.
Leseva, M., N., Grand, R., J., Klett, H., Boerries, M., Busch, H., Binder, A., M., and Michels, K., B. Differences in DNA methylation and functional expression in lactase persistent and non-persistent individuals. Scientific reports, 8(1), pp.1-14, 2018.
Lewinsky, R., H., Jensen, T., G., Møller, J., Stensballe, A., Olsen, J., and Troelsen, J., T. T-13910 DNA variant associated with lactase persistence interacts with Oct-1 and stimulates lactase promoter activity in vitro. Human Molecular Genetics, 14(24), 3945!3953, 2005
Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v1. 2013.
Li, H. and Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 25(14):1754–1760, 2009.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. The Sequence Alignment/Map format and SAMtools. Bioinformatics (Oxford, England), 25(16):2078–9, 2009.
Liebert, A., López, S., Jones, B., L., Montalva, N., Gerbault, P., Lau, W., Thomas, M., G., Bradman, N., Maniatis, N., and Swallow, D., M. World-wide distributions of lactase persistence alleles and the complex effects of recombination and selection. Human genetics, 136(11-12):1445–1453, 2017.
Lidke, G., Jantzen, D., Lorenz, S., Terberger, T. The Bronze Age battlefield in the Tollense Valley, Northeast Germany. Conflict scenario research. In: M. Fernández-Götz/N. Roymans (eds.), Conflict Archaeology. Materialities of Collective Violence from Prehistory

to Late Antiquity. EAA Themes in Contemporary Archaeology 5. Abingdon: Routledge, 61-68, 2018.
Link, V., Kousathanas, A., Veeramah, K., Sell, C., Scheu, A., and Wegmann, D. ATLAS: Analysis Tools for Low-depth and Ancient Samples. bioRxiv, 2017.
Lipson, M., Szécsényi-Nagy, A., Mallick, S., Pósa, A., Stégmár, B., Keerl, V., Rohland, N., Stewardson, K., Ferry, M., Michel, M. et al. Parallel palaeogenomic transects reveal complex genetic history of early European farmers. Nature, 551(7680), pp.368-372, 2017.
Lokki, A., I., Järvelä, I., Israelsson, E., Maiga, B., Troye-Blomberg, M., Dolo, A., Doumbo, O., K., Meri, S., and Holmberg, V. Lactase persistence genotypes and malaria susceptibility in fulani of mali. Malaria journal, 10(1):9, 2011.
Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S., A., Harney, E., Stewardson, K., Fernandes, D., Novak, M., et al. Genome-wide patterns of selection in 230 ancient eurasians. Nature, 528(7583):499, 2015.
Mathieson S. and Mathieson I. *FADS1* and the Timing of Human Adaptation to Agriculture. Molecular Biology and Evolution 35(12):2957–2970, 2018.
Mathieson, I., Alpaslan-Roodenberg, S., Posth, C., Szécsényi-Nagy, A., Rohland, N., Mallick, S., Olalde, I., Broomandkhoshbacht, N., Candilio, F., Cheronet, O., et al. The genomic history of southeastern Europe. Nature, 555(7695), pp.197-203, 2018.
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M., A. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research, 20(9):1297–1303, 2010.
Meyer M., and Kircher M. Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing. Cold Spring Harb Protoc 6:pdb.prot5448, 2010.
Mittnik, A., Wang, C., C., Pfrengle, S., Daubaras, M., Zariņa, G., Hallgren, F., Allmäe, R., Khartanovich, V., Moiseyev, V., Tõrv, M., et al. The genetic prehistory of the baltic sea region. Nature communications, 9(1):442, 2018.
Mittnik, A., Massy, K., Knipper, C., Wittenborn, F., Friedrich, R., Pfrengle, S., Burri, M., Carlichi-Witjes, N., Deeg, H., Furtwängler, A., et al. Kinship-based social inequality in bronze age europe. Science, 366(6466):731–734, 2019.
Montalva, N., Adhikari, K., Liebert, A., Mendoza-Revilla, J., Flores, S., V., Mace, R., Swallow, D., M. Adaptation to milking agropastoralism in Chilean goat herders and nutritional benefit of lactase persistence. Ann Hum Genetics 83: 11–22, 2019.
Narasimhan, V., M., Patterson, N., Moorjani, P., Rohland, N., Bernardos, R., Mallick, S., Lazaridis, I., Nakatsuka, N., Olalde, I., Lipson, M., et al. The formation of human populations in South and Central Asia. Science, 365(6457), p.eaat7487, 2019.
Olalde, I., Brace, S., Allentoft, M., E., Armit, I., Kristiansen, K., Booth, T., Rohland, N., Mallick, S., Szécsényi-Nagy, A., Mittnik, A., et al. The beaker phenomenon and the genomic transformation of northwest europe. Nature, 555(7695):190, 2018.
Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. Ancient admixture in human history. Genetics, 192(3):1065–1093, 2012.
Pickrell, J., K., and Pritchard, J., K. Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. PLoS Genetics, 8(11):e1002967, 2012.
Poznik, G., D., Xue, Y., Mendez, F., L., Willems, T., F., Massaia, A., Sayres, M., A., W., & Chen, Y. Punctuated bursts in human male demography inferred from 1,244 worldwide Y-chromosome sequences. Nature genetics, 48(6), 593-599, 2016.
Price, T., D., Frei, R., Brinker, U., Lidke, G., Terberger, T., Frei, K., M., and Jantzen, D. Multi-isotope proveniencing of human remains from a bronze age battlefield in the

Tollense valley in northeast Germany. Archaeological and Anthropological Sciences, 11(1):33–49, 2019.
Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M., A., Bender D., Maller J., Sklar P., de Bakker P., I., Daly M., J., Sham P., C., PLINK: a tool set for whole-genome association and population-based linkage analyses. American Journal of Human Genetics, 81(3):559-75, 2007
Salque, M., Bogucki, P., I., Pyzel, J., Sobkowiak-Tabaka, I., Grygiel, R., Szmyt, M., and Evershed, R., P. Earliest evidence for cheese making in the sixth millennium bc in northern europe. Nature, 493(7433):522, 2013.
Sánchez-Quinto, F., Malmström, H., Fraser, M., Girdland-Flink, L., Svensson, E., M., Simões, L., G., George, R., Hollfelder, N., Burenhult, G., Noble, G., et al. Megalithic tombs in western and northern Neolithic Europe were linked to a kindred society. Proceedings of the National Academy of Sciences, 116(19), pp.9469-9474, 2019.
Schilz, F., Molekulargenetische Verwandtschaftsanalysen am prähistorischen Skelettkollektiv der Lichtensteinhöhle. Dissertation, Göttingen, 2006 (http://webdoc.sub.gwdg.de/diss/2006/schilz/schilz.pdf)
Schroeder, H., Margaryan, A., Szmyt, M., Theulot, B., Włodarczak, P., Rasmussen, S., Gopalakrishnan, S., Szczepanek, A., Konopka, T., Jensen, T., Z., et al. Unraveling ancestry, kinship, and violence in a Late Neolithic mass grave. Proceedings of the National Academy of Sciences, 116(22), pp.10705-10710, 2019.
Segurel, L., Guarino-Vignon, P., Marchi, N., Lafosse, S., Laurent, R., Bon, C., Fabre, A., Hegay, T. and Heyer, E. Why and when was lactase persistence selected for? Insights from Central Asian herders and ancient DNA. Plos Biology, 18(6), p.e3000742, 2020.
Seidenberg, V.: Ein bronzezeitlicher Familienclan als genetisches Archiv – Morphologisch-paläogenetische Bearbeitung des Skelettkollektivs aus der Lichtensteinhöhle. Dissertation, Göttingen, 2016 .(https://ediss.uni-goettingen.de/bitstream/handle/11858/00-1735-0000-002B-7D55-C/Dissertation_Seidenberg.pdf?sequence=1)
Skoglund, P., Stora, J., Götherström, A., Jakobsson, M. Accurate sex identification of ancient human remains using DNA shotgun sequencing. Journal of Archeological Science, 40(12):4477-4482, 2013.
Skoglund, P., Malmström, H., Omrak, A., Raghavan, M., Valdiosera, C., Günther, T., Hall, P., Tambets, K., Parik, J., Sjögren, K.G. et al. Genomic diversity and admixture differs for Stone-Age Scandinavian foragers and farmers. Science, 344(6185), pp.747-750, 2014.
Stern, A., J., Wilton, P.,R., and Nielsen, R. An approximate full-likelihood method for inferring selection and allele frequency trajectories from DNA sequence data. PLoS genetics, 15(9), p.e1008384, 2019.
Sverrisdóttir, O., O., Timpson, A., Toombs, J., Lecoeur, C., Froguel, P., Carretero, J., M., Arsuaga, F., J., L., Götherström, A., and Thomas, M., G.,. Direct estimates of natural selection in iberia indicate calcium absorption was not the only driver of lactase persistence in europe. Molecular biology and evolution, 31(4):975–983, 2014.
Terberger, T. and Heinemeier, J. Die Ernährungsweise der bronzezeitlichen Menschen aus dem Tollensetal im Spiegel ihrer 13C und 15N Isotopie - erste Ergebnisse. Tod im Tollensetal—-Forschungen zu den Hinterlassenschaften eines bronzezeitlichen Gewaltkonfliktes in Mecklenburg-Vorpommern, 1:209–214, 2014.
Terberger, T., Jantzen, D., Krüger, J., Lidke, G. Das bronzezeitliche Kampfgeschehen im Tollensetal – ein Großereignis oder wiederholte Konflikte? In: S. Hansen/R. Krause (eds.), Bronzezeitliche Burgen zwischen Taunus und Karpaten / Bronze Age Hillforts between Taunus and Carpathian Mountains. UPA 319. Bonn: Habelt, 103-122, 2018.

Uhlig, T., Krüger, J., Lidke, G., Jantzen, D., Lorenz, S., Ialongo, N., and Terberger, T. Lost in combat? A scrap metal find from the bronze age battlefield site at tollense. antiquity, 93 (371):1211–1230, 2019.

Veeramah, K., R., Rott, A., Groß, M., Dorp, L., López, S., Kirsanow, K., Sell, C., Blöcher, J., Wegmann, D., Link, V., Hofmanová, Z., Peters, J., Trautmann, B., Gairhos, A., Haberstroh, J., Päffgen, B., Hellenthal, G., Haas, G., B., Harbeck, M., and Burger, J. Population genomic analysis of elongated skulls reveals extensive female-biased imigration in early medieval bavaria. Proceedings of the National Academy of Sciences, 115(13):3494–3499, 2018.

Vigne, JD., and Helmer, D. Was milk a ”secondary product” in the old world neolithisation process? its role in the domestication of cattle, sheep and goats. Anthropozoologica, 42(2):9–40, 2007.

Wahlqvist, M.L. Lactose nutrition in lactase nonpersisters. Asia Pacific journal of clinical nutrition, 24(Supplement), p.S21, 2015.

Walker, C. and Thomas, M., G. The evolution of lactose digestion. In Lactose: Lactose: Evolutionary Role, Health Effects, and Applications, Editors: Paques Marcel, Lindner Cordula, pages 1–48. Academic Press, SBN 9780128117200, https://doi.org/10.1016/B978-0-12-811720-0.00001-5, 2019.

Wang, C., Zhan, X., Bragg-Gresham, J., Kang, H., M., Stambolian, D., Chew, E., Y., Branham, K., E., Heckenlively, J., Fulton, R., Wilson, R., K., Mardis, E., R., et al. Ancestry estimation and control of population stratification for sequence-based association studies. Nature Genetics, 46(4):409–415, 2014.

Wang, C.C., Reinhold, S., Kalmykov, A., Wissgott, A., Brandt, G., Jeong, C., Cheronet, O., Ferry, M., Harney, E., Keating, D. et al. Ancient human genome-wide data from a 3000-year interval in the Caucasus corresponds with eco-geographic regions. Nature communications, 10(1), pp.1-13, 2019.

Wang, C., C., Reinhold, S., Kalmykov, A., Wissgott, A., Brandt, G., Jeong, C., Cheronet, O., Ferry, M., Harney, E., Keating, D. et al. Ancient human genome-wide data from a 3000-year interval in the Caucasus corresponds with eco-geographic regions. Nature communications, 10(1), pp.1-13, 2019.

Waples R., Albrechtsen A., Moltke I. Allele frequency-free inference of close familial relationships from genotypes or low-depth sequencing data. Molecular Ecology, 28(1):35-48, 2019.

Wilde, S., Timpson, A., Kirsanow, K., Kaiser, E., Kayser, M., Unterländer, M., Hollfelder, N., Potekhina, I., D., Schier, W., Thomas, M., G., et al. Direct evidence for positive selection of skin, hair, and eye pigmentation in Europeans during the last 5,000 years. Proceedings of the National Academy of Sciences, page 201316513, 2014.

Zhang, G., Muglia, L., J., Chakraborty, R., Akey, J., M., and Williams, S., M. Signatures of natural selection on genetic variants affecting complex human traits. Applied & translational genomics, 2:78–94, 2013.

Žegarac, A., Winkelbach, L., Blöcher, J., Diekmann, Y., Gavrilovic, M., K., Porcic, M., Stojkovic, B., Milasinovic, L., Schreiber, M., Wegmann, D., et al. 2020. Kinship, acquired and inherited status, and population structure at the Early Bronze Age Mokrin necropolis in northern Serbia. bioRxiv.

Star Resources

Deposited Data		
Tollense capture raw data	This study	PRJEB38406
PCR genotype estimates of 37 individuals from the following cultures: Eneolithic, Usatovo, Yamnaya, Early Catacomb Culture, Developed Catacomb Culture, Yamnaya-Poltavkinskaja, Late Catacomb Culture	This study	Supplementary Table 1
Human reference genome NCBI build 37	Church et al. 2011	GRCh37
Unrelated CEU from 1000 Genomes	1000 Genomes Project Auton et al. 2015	NA06984, NA06985, NA06986, NA06989, NA06994, NA07000, NA07037, NA07048, NA07051, NA07056, NA07346, NA07347, NA07357, NA10847, NA10851, NA11829, NA11830, NA11831, NA11832, NA11840, NA11843, NA11881, NA11892, NA11893, NA11894, NA11918, NA11919, NA11920, NA11930, NA11931, NA11932, NA11933, NA11992, NA11993, NA11994, NA11995, NA12003, NA12004, NA12005, NA12006, NA12043, NA12044, NA12045, NA12046, NA12058, NA12144, NA12154, NA12155, NA12156, NA12234, NA12249, NA12273, NA12275, NA12282, NA12283, NA12286, NA12287, NA12340, NA12341, NA12342, NA12347, NA12348, NA12383, NA12399, NA12400, NA12413, NA12414, NA12489, NA12546, NA12716, NA12717, NA12718, NA12748, NA12749, NA12750, NA12751, NA12760, NA12763, NA12775, NA12778, NA12813, NA12814, NA12815, NA12827, NA12828, NA12829, NA12830, NA12842, NA12843, NA12872, NA12873, NA12874, NA12889, NA12890, NA12891, NA12892
1000 Genomes Eurasian Yoruba	1000 Genomes Project Auton et al. 2015	NA18486, NA18488, NA18489, NA18498, NA18499, NA18501, NA18502, NA18504, NA18505, NA18507, NA18508, NA18510, NA18511, NA18516, NA18517, NA18519, NA18520, NA18522, NA18523, NA18853, NA18856, NA18858, NA18861, NA18864, NA18865, NA18867, NA18868, NA18870, NA18871, NA18873,

		NA18874, NA18876, NA18877, NA18878, NA18879, NA18881, NA18907, NA18908, NA18909, NA18910, NA18912, NA18915, NA18916, NA18917, NA18923, NA18924, NA18933, NA18934, NA19092, NA19093, NA19095, NA19096, NA19098, NA19099, NA19102, NA19107, NA19108, NA19113, NA19114, NA19116, NA19117, NA19118, NA19119, NA19121, NA19129, NA19130, NA19131, NA19137, NA19138, NA19141, NA19143, NA19144, NA19146, NA19147, NA19149, NA19152, NA19153, NA19159, NA19160, NA19171, NA19172, NA19175, NA19184, NA19185, NA19189, NA19190, NA19197, NA19198, NA19200, NA19201, NA19204, NA19206, NA19207, NA19209, NA19210, NA19213, NA19214, NA19222, NA19223, NA19225, NA19235, NA19236, NA19238, NA19239, NA19247, NA19248, NA19256, NA19257
Turkish genomes	Alkan et al. 2014	06A010111, 08P210611, 24D220611, 25A220611, 31P140611, 32A140611, 33M140611, 34S291210, 35C240511, 38I220611, 42S291210, 48S210611, 50G301210, 52C130611, 57M220611, 65A220611
5th century Bavarian individuals	Veeramah et al. 2018	AED_106, AED_1119, AED_1135, AED_204, AED_249, AED_432, AED_92, ALH_10, ALH_1, ALH_2, ALH_3, NW_255, STR_241, STR_248, STR_266, STR_300, STR_316, STR_393, STR_480, STR_486, STR_502
Prague (Jinonice, Zahradnictví and Kobylisy, Ke Stírce Street)	Olalde et al. 2018	I7195, I7196, I7197, I7198, I7199, I7200, I7201, I7202, I7203, I4886, I4888, I4889, I4890, I4891, I4945
Bedfordshire, Biddenham Loop	Olalde et al. 2018	I7575, I7576, I7577, I7578, I7580, I7626, I7628
Kivutkalns, Baltic Bronze Age in Latvia	Mittnik et al. 2018	Kivutkalns153, Kivutkalns164, Kivutkalns194, Kivutkalns19, Kivutkalns207, Kivutkalns215, Kivutkalns222, Kivutkalns25, Kivutkalns42
Szolad (northern ancestry), Early Medieval in Hungary	Amorim et al. 2018	SZ2, SZ3, SZ4, SZ5, SZ9, SZ11, SZ13, SZ15, SZ16, SZ18, SZ30, SZ38, SZ45

Genome DK (Danish)	Besenbacher et al. 2015	only allele frequencies available
GONL (Dutch)	Francioli et al. 2014	gonl-100c, gonl-101c, gonl-102c, gonl-103c, gonl-104c, gonl-105c, gonl-106c, gonl-107c, gonl-108c, gonl-109c, gonl-10c, gonl-110c, gonl-111c, gonl-112c, gonl-113c, gonl-114c, gonl-115c, gonl-116c, gonl-117c, gonl-118c, gonl-119c, gonl-11c, gonl-120c, gonl-121c, gonl-122c, gonl-123c, gonl-124c, gonl-125c, gonl-126c, gonl-127c, gonl-128c, gonl-129c, gonl-12c, gonl-130c, gonl-131c, gonl-132c, gonl-133c, gonl-134c, gonl-135c, gonl-136c, gonl-137c, gonl-138c, gonl-139c, gonl-13c, gonl-140c, gonl-141c, gonl-142c, gonl-143c, gonl-144c, gonl-145c, gonl-146c, gonl-147c, gonl-148c, gonl-149c, gonl-14c, gonl-150c, gonl-151c, gonl-152c, gonl-153c, gonl-154c, gonl-155c, gonl-156c, gonl-157c, gonl-158c, gonl-159c, gonl-15c, gonl-160c, gonl-161c, gonl-162c, gonl-163c, gonl-164c, gonl-165c, gonl-166c, gonl-167c, gonl-168c, gonl-169c, gonl-16c, gonl-170c, gonl-171c, gonl-172c, gonl-174c, gonl-175c, gonl-176c, gonl-177c, gonl-178c, gonl-179c, gonl-17c, gonl-180c, gonl-181c, gonl-182c, gonl-183c, gonl-184c, gonl-185c, gonl-186c, gonl-187c, gonl-188c, gonl-189c, gonl-18c, gonl-190c, gonl-191c, gonl-192c, gonl-193c, gonl-194c, gonl-195c, gonl-196c, gonl-197c, gonl-198c, gonl-199c, gonl-19c, gonl-1c, gonl-200c, gonl-201c, gonl-202c, gonl-203c, gonl-204c, gonl-205c, gonl-206c, gonl-207c, gonl-208c, gonl-209c, gonl-20c,

		gonl-210c, gonl-212c, gonl-213c, gonl-214c, gonl-215c, gonl-216c, gonl-217c, gonl-218c, gonl-219c, gonl-21c, gonl-220c, gonl-221c, gonl-222c, gonl-223c, gonl-224c, gonl-225c, gonl-226c, gonl-227c, gonl-228c, gonl-229c, gonl-22c, gonl-230c, gonl-231c, gonl-232c, gonl-233c, gonl-234c, gonl-235c, gonl-236c, gonl-237c, gonl-238c, gonl-239c, gonl-23c, gonl-240c, gonl-241c, gonl-242c, gonl-243c, gonl-244c, gonl-245c, gonl-246c, gonl-247c, gonl-248c, gonl-249c, gonl-24c, gonl-250c, gonl-25c, gonl-26c, gonl-27c, gonl-28c, gonl-29c, gonl-2c, gonl-30c, gonl-31c, gonl-32c, gonl-33c, gonl-34c, gonl-35c, gonl-36c, gonl-37c, gonl-38c, gonl-39c, gonl-3c, gonl-40c, gonl-41c, gonl-42c, gonl-43c, gonl-44c, gonl-45c, gonl-46c, gonl-47c, gonl-48c, gonl-49c, gonl-4c, gonl-50c, gonl-51c, gonl-52c, gonl-53c, gonl-54c, gonl-55c, gonl-56c, gonl-57c, gonl-58c, gonl-59c, gonl-5c, gonl-60c, gonl-61c, gonl-62c, gonl-63c, gonl-64c, gonl-65c, gonl-66c, gonl-67c, gonl-68c, gonl-69c, gonl-6c, gonl-70c, gonl-71c, gonl-72c, gonl-73c, gonl-74c, gonl-75c, gonl-76c, gonl-77c, gonl-78c, gonl-79c, gonl-7c, gonl-80c, gonl-81c, gonl-82c, gonl-83c, gonl-84c, gonl-85c, gonl-86c, gonl-87c, gonl-88c, gonl-89c, gonl-8c, gonl-90c, gonl-91c, gonl-92c, gonl-93c, gonl-94c, gonl-95c, gonl-96c, gonl-97c, gonl-98c, gonl-99c, gonl-9c
Lech Valley Bavarians	Mittnik et al. 2019	AITI_119, AITI_2, AITI_40, AITI_43, AITI_50, AITI_72, AITI_78, AITI_98, AITI_65adult, AITI_66, AITI_92, AITI_95, OBKR_80, OBKR_117, OBKR_2, OBKR_50, OBKR_67, OBKR_9A, POST_28, POST_44, POST_50, POST_6, POST_35,

		POST_99, UNTA58_153, UNTA85_1412, WEHR_1192SkB, WEHR_1415adult, WEHR_1586
Mokrin	Žegarac et al. 2020	MOK10B, MOK12, MOK13, MOK14, MOK15, MOK16A, MOK17A, MOK18A, MOK19A, MOK20, MOK21A, MOK22, MOK23, MOK24A, MOK25A, MOK26A, MOK27, MOK28A, MOK29A, MOK30, MOK31, MOK32, MOK33, MOK9B
East European Steppe: 3600-2300 BCE	Allentoft et al. 2015, Järve et al. 2019, Lipson et al. 2017, Mathieson et al. 2015, Mathieson et al. 2018, Narasimhan et al. 2019, Olalde et al. 2018, Schroeder et al. 2019, Wang et al. 2019	RISE547.SG, RISE552.SG, I0374, I1917, I2105, I4110, I5118, I5119, ILK001, ILK002, ILK003, I5882, I8745, I11501, I11531, I11732, I11735, I11736, I11734, RISE1166.SG, AY2001, AY2003, RK1001, RK1003, RK4001, RK4002, SA6001, SA6002, SA6003, SA6004, SA6010, SA6013, SIJ002, SIJ003, ZO2002, I2791_published, I5884_publishedRISE547.SG, RISE552.SG, I0374, I1917, I2105, I4110, I5118, I5119, ILK001, ILK002, ILK003, I5882, I8745, I11501, I11531, I11732, I11735, I11736, I11734, RISE1166.SG, AY2001, AY2003, RK1001, RK1003, RK4001, RK4002, SA6001, SA6002, SA6003, SA6004, SA6010, SA6013, SIJ002, SIJ003, ZO2002, I2791_published, I5884_published
Area of Corded Ware Culture: 2900-2300 BCE	Allentoft et al. 2015, Fernandes et al. 2018, Gamba et al. 2014, Jones et al. 2017, Mathieson et al. 2015, Mathieson et al. 2018, Mittnik et al. 2019, Narasimhan et al. 2019, Olalde et al. 2018, Sanchez-Quinto et al. 2019, Schroeder et al. 2019, Skoglund et al. 2014	I7207, I7208, I7209, I7278, I7289, I7212, I6695, I6696, I7040, I7044, I7045, I7195, I7196, I7200, I7201, I7202, I7203, I7213, I7286, I0805, Ajvide52.SG, Ajvide58.SG, RISE1.SG, RISE94.SG, I3529, I2365, I4178, I2741, I4629, I5015, I5043, I5520, I5521, I5523, I5525, I5529, I6534, I6581, I2786, N44.SG, N45.SG, N47.SG, N49.SG, RISE1159.SG, RISE1162.SG, RISE1163.SG, RISE1164.SG, RISE1167.SG, RISE1170.SG, RISE1171.SG, RISE1172.SG, RISE1173.SG, RISE1248.SG,

		RISE1249.SG, RISE1250.SG, ans016.SG

Oligonucleotides		
MYBait kit	Arbor biosciences; https://arborbiosci.com/genomics/targeted-sequencing/mybaits/mybaits-custom/	
P5 and P7	Meyer and Kircher (2010); IDT, Leuven, Belgium	
IS4, IS5, IS6 and IS7	Meyer and Kircher (2010); IDT, Leuven, Belgium	
Software and Algorithms		
ATLAS	Link et al. 2017	
TreeMix	Pickrell and Pritchard 2012	
plink2	Chang et al. 2015	
approxWF	Ferrer-Admetlla et al. 2016	
laser2	Wang et al. 2015	
ADMIXTOOLS	Patterson et al. 2012	
yHaplo*	Poznik 2016	
scipy.stats python library	https://docs.scipy.org/doc	
numpy python library	https://numpy.org/	
bwa aln	Li and Durbin 2009	
GATK	McKenna et al. 2010	
Samtools	Li et al. 2009	
VCFtools	Danecek et al. 2011	
picard-tools	https://broadinstitute.github.io/picard/	
contamMix	Fu et al. 2014	
Other		
Agilent 2100 Expert Bioanalyzer System and High Sensitivity DNA Analysis Kit	Agilent Technologies	
Amicon Ultra-15 Centrifugal Filter Units	Merck Millipore, Darmstadt, Germany	
AmpliTaq Gold ® Buffer II (10x)	Life Technologies ™	
AmpliTaq Gold ® DNA Polymerase	Life Technologies ™	
ATP Solution (100 mM)	Life Technologies ™	
Bovine Serum Albumin (BSA) (20 mg/ml)	Roche Diagnostics	
Bst Polymerase, Large Fragment (8 U/µl)	New England Biolabs GmbH	
dNTPs (each 10 mM)	Qiagen, Hilden, Germany	
dNTPs (each 25 mM)	Agilent Technologies	
EDTA (0.5 M), pH 8.0	Ambion/Applied Biosystems, Life Technologies ™, Darmstadt, Germany	
Herculase II Fusion ® DNA Polymerase	Agilent Technologies	
Herculase II Reaction Buffer	Agilent Technologies	

$MgCl_2$ (25 mM)	Life Technologies ™	
MinElute ® PCR Purification Kit	Qiagen, Hilden, Germany	
MSB ® Spin PCRapace	Invitek, Stratec Molecular, Berlin, Germany	
Sodium N-lauryl sarcosinate	Merck Millipore, Darmstadt, Germany	
Nuclease-free H_2O	Life Technologies ™	
PEG-4000	Thermo Scientific ™	
Proteinase K	Roche Diagnostics, Mannheim, Germany	
Phenol/chloroform/isoamylalcohol (25:24:1)	Roth, Karlsruhe, Germany	
Qubit ® Fluorometric quantitation and dsDNA HS Assay Kit	Invitrogen ™	
T4 DNA Ligase (5 U/µl)	Thermo Scientific ™	
T4 DNA Ligase Buffer (10X)	Thermo Scientific ™	
T4 DNA Polymerase (5 U/µl)	Thermo Scientific ™	
T4 Polynucleotide Kinase	Invitrogen ™	
Tango Buffer (10x)	Life Technologies ™	
ThermoPol Buffer (10X)	New England Biolabs GmbH	
Trichlormethan/Chloroform	Roth, Karlsruhe, Germany	