Shared chromosomal segments connect ancient human societies

Anders Bergström
School of Biological Sciences, University of East Anglia, Norwich, UK

Long segments of the genome shared identical by descent (IBD) demonstrate recent relatedness between individuals. A new computational method robustly identifies shared IBD segments in ancient DNA data, providing insights into mobility and demography of prehistoric human societies.

The stochastic nature of meiotic recombination means that the similarity between the chromosomes of relatives varies greatly as one moves along the genome. In many parts of their chromosomes, relatives will not be more similar to each other than any other two people. But in other parts, demarcated by recombination events that occurred in the last few generations, relatives will have inherited the very same segment of DNA from the same ancestor. Such segments are said to be identical by descent (IBD), and the more recent the common ancestor from which it was inherited, the longer the IBD segment. Because they unambiguously demonstrate recent relatedness, IBD segments can inform greatly on the kinship structure, demography and migration history of populations. This is also what personal ancestry companies use to identify relatives within their customer base.

The ability to analyse DNA from ancient skeletal remains is revolutionising several areas of the genetic sciences, with ancient genomic data now being available from more than 10,000 past humans as well as more than 1,000 individuals from other species. However, the low-quality data typically associated with ancient DNA, reflecting low endogenous DNA retrieval and post-mortem damage, makes the calling of IBD segments challenging. Even a small number of erroneous genotypes inside a long IBD segment might make it look like multiple, short segments. In this issue of Nature Genetics, Ringbauer et al.[1] develop a new computational method, ancIBD, to identify IBD segments in ancient genomic data.

Past analyses of IBD sharing in high-quality, modern genetic datasets highlight how this information can inform on human population history. A study of more distant IBD sharing among present-day Europeans found that on a timescale of 500 years most genetic relatives tend to live within the same country or geographical region, but that on a timescale of 1,500 years or more there are genetic relatives between all parts of Europe[2]. Methods have also been developed to use the degree of IBD sharing among large numbers of genomes to infer effective population sizes histories over the quite recent past—from the last few generations to the last couple of hundred or so generations (i.e. ~100-5,000 years ago)—exploiting how IBD segments will on average be shared more recently within a smaller population[3,4].

The detection of close genetic relatives is already an important component of ancient human genomics, but previous studies have mainly used simpler methods to identify very close relatives. By estimating relatedness coefficients based on the degree of allele sharing across the genome as a whole, individuals related up to third degree can be identified. For example, the identification of a pair of second degree relatives who lived during the Viking Age, one in Denmark and the other in England, directly demonstrated long-distance movement within the preceding two generations[5]. Tombs from the European Neolithic have revealed very large,

multi-generational pedigrees in which social structure and burial practices are reflected, including in who was buried where within the tomb[6]. Because the limits of human age at reproduction and death impose constraints on how far apart in time relatives could have lived, kinship can also be used to infer approximate dates of undated individuals, or to reduce the uncertainty associated with radiocarbon dates[7].

Kinship detection methods based on genome-wide allele sharing, because they average the signal across IBD and non-IBD parts of the genome, are inherently less powerful than methods calling the actual IBD segments. The ancIBD method can identify relatives up to sixth degree, which thus greatly extends the reach of the approach. It also employs various computational tricks to keep runtimes low and thereby enable very large numbers of pairwise genome comparisons. The large-scale application of IBD analyses to ancient genomes now promises to add a whole new layer to our understanding of the genetic past.

Most commonly used methods for inferring relationships between populations rely on allele frequency similarities, and this has its limitations. It might be that a given population has allele frequencies that make it the best available proxy for an observed episode of gene flow, relative to others that have been sampled—but that doesn't necessarily mean that the given population is the actual source. This concern has been raised at the conclusion that the Yamnaya herders of the eastern European steppe were the source of large-scale migrations across Europe in the Late Neolithic and Bronze Age—perhaps some other, genetically unsampled group was actually the source[8]. Ringbauer et al. identify very strong signals of recent IBD sharing between Yamnaya and people in central and northern Europe (specifically people from the so-called Corded Ware culture). Such IBD sharing can only be explained by direct genealogical relationships within at most the last few hundred years. The study also detects kinship between members of steppe cultures that migrated east, including a pair of 5th degree relatives separated by 1,410 km in central Asia—a striking demonstration of how mobile people in these cultures must have been.

A development that greatly helps Ringbauer et al. identify IBD segments is genotype imputation, which is performed before running the ancIBD software. By allowing missing or low-quality genotypes in the ancient genomes to be filled in using information from large, high-quality reference panels of present-day genomes, imputation has emerged as an increasingly popular approach. While there are concerns about how accurately ancient genomes are imputed, especially very old ones with ancestries poorly represented by reference panels, it's increasingly clear that imputation works well at least for most ancient genomes from Eurasia and the last 10,000 years[9]. IBD analyses in principle also hold great promise for ancient genomics in other species—for example by uncovering human-driven changes in mating and movement in domesticated species—but lesser availability of resources such as imputation panels and high-resolution recombination maps means application to non-human species will be more challenging.

The ancIBD method has already been used to study large, multi-generational pedigrees at a Neolithic site in France[10], and kinship patterns among Bronze Age people in the Ural mountains[11]. The layer of information provided by IBD sharing is likely to become integral to research on the human genetic past, especially as many studies are increasingly shifting emphasis from big-picture migrations to more local questions, right down to the structure of communities and families.

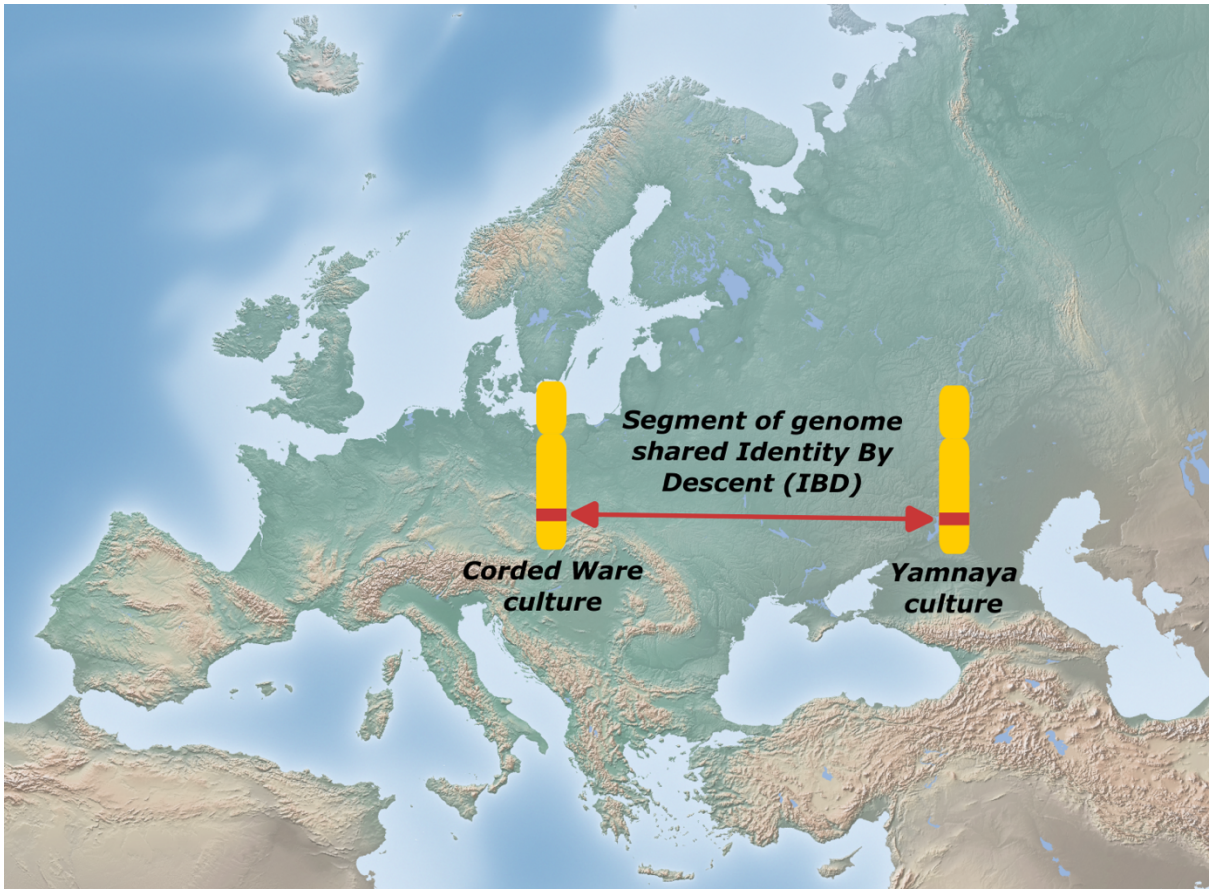

Fig 1. Long-distance IBD sharing. Individuals from the Yamnaya culture on the eastern European steppe, and individuals from the Corded Ware culture in central and northern Europe, share long chromosomal segments IBD. This reflects relatedness on a timescale of at most a few hundred years, providing unambiguous evidence of recent connectivity.

Competing Interests

The author declares no competing interests.

References

1. Ringbauer, H. *et al.* Accurate detection of identity-by-descent segments in human ancient DNA. *Nat. Genet.* **56**, 143–151 (2024).
2. Ralph, P. & Coop, G. The geography of recent genetic ancestry across Europe. *PLoS Biol.* **11**, e1001555 (2013).
3. Browning, S. R. & Browning, B. L. Accurate Non-parametric Estimation of Recent Effective Population Size from Segments of Identity by Descent. *Am. J. Hum. Genet.* **97**, 404–418 (2015).
4. Palamara, P. F., Lencz, T., Darvasi, A. & Pe'er, I. Length distributions of identity by descent reveal fine-scale demographic history. *Am. J. Hum. Genet.* **91**, 809–822 (2012).

5. Margaryan, A. *et al.* Population genomics of the Viking world. *Nature* **585**, 390–396 (2020).
6. Fowler, C. *et al.* A high-resolution picture of kinship practices in an Early Neolithic tomb. *Nature* **601**, 584–587 (2022).
7. Sedig, J. W., Olade, I., Patterson, N., Harney, É. & Reich, D. Combining ancient DNA and radiocarbon dating data to increase chronological accuracy. *J. Archaeol. Sci.* **133**, (2021).
8. Klejn, L. S. *et al.* Discussion: Are the Origins of Indo-European Languages Explained by the Migration of the Yamnaya Culture to the West? *European Journal of Archaeology* **21**, 3–17 (2018).
9. Sousa da Mota, B. *et al.* Imputation of ancient human genomes. *Nat. Commun.* **14**, 3660 (2023).
10. Rivollat, M. *et al.* Extensive pedigrees reveal the social organization of a Neolithic community. *Nature* **620**, 600–606 (2023).
11. Blöcher, J. *et al.* Descent, marriage, and residence practices of a 3,800-year-old pastoral community in Central Eurasia. *Proc. Natl. Acad. Sci. U. S. A*. **120**, e2303574120 (2023).