

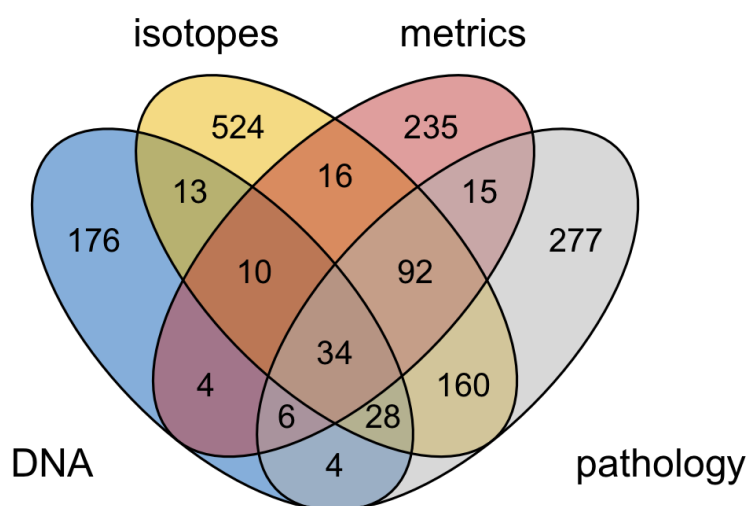
Socio-cultural practices may have affected sex differences in stature in Early Neolithic Europe

In the format provided by the
authors and unedited

List of Supplementary Figures

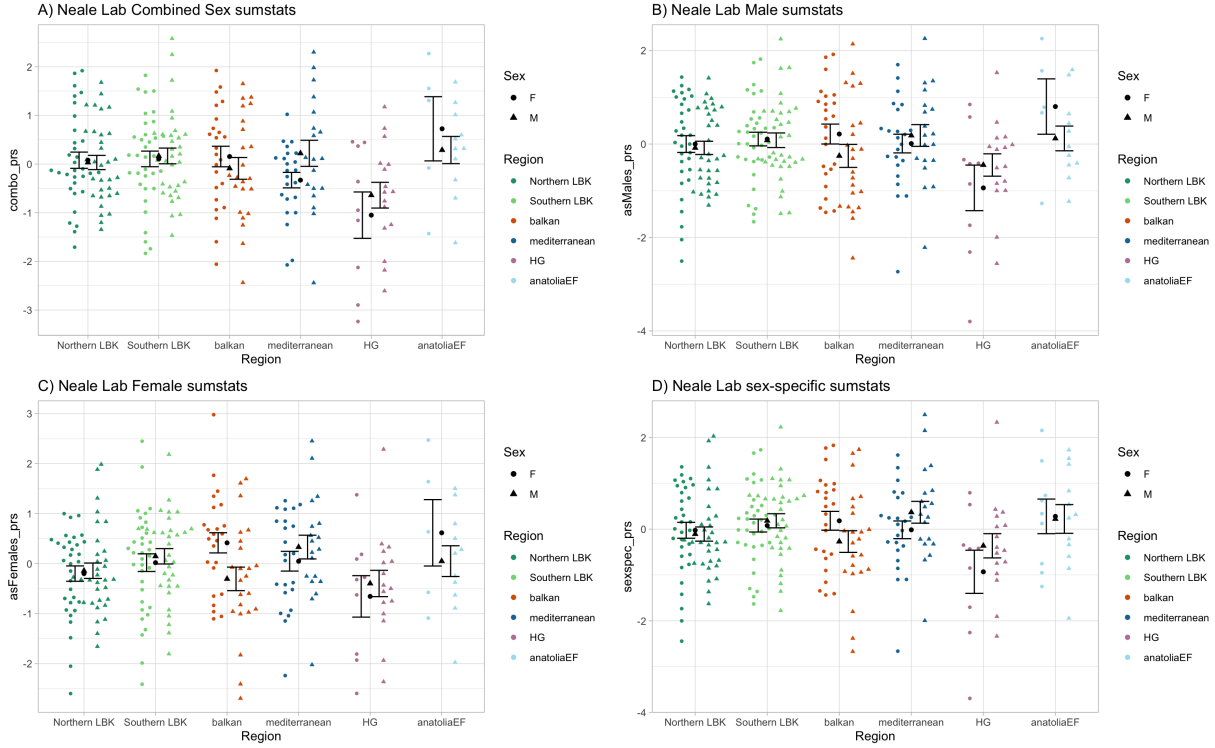
1	Samples by data type	2
2	Sex-specific clumping/thresholding PRS.	3
3	Sibling-based PRS and femoral lengths	4
4	PCA for unimputed DNA.	4
5	LDpred2 PRS.	5
6	Herbivore baseline.	6
7	Date restricted plots.	6

Supp.Fig. 1: Data set counts



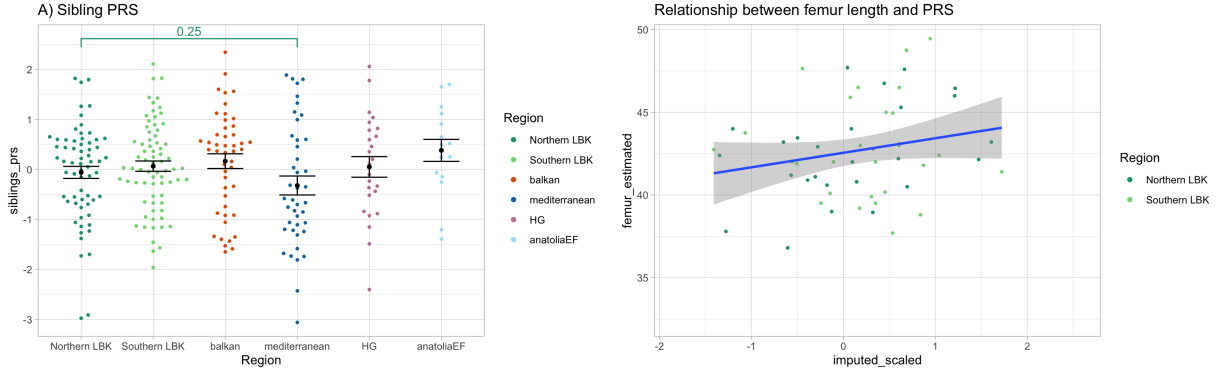
Supplementary Figure 1: Samples by data type

Venn diagram illustrating the overlaps in data. All individuals have at least one type of data, but only subsets have more than one type.



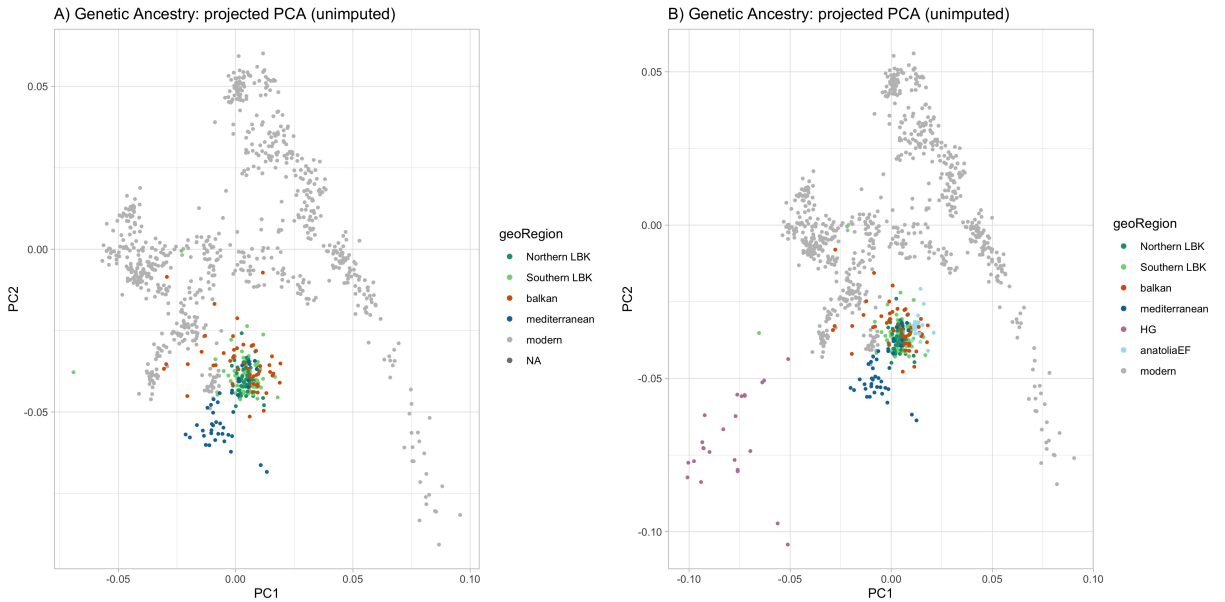
Supplementary Figure 2: Sex-specific clumping/thresholding PRS.

In each plot, black points are the group mean and black bars are the 95% confidence intervals. Differences between groups was assessed based on two-tailed pairwise t -tests adjusted for multiple comparisons; sample sizes are the same for all plots, from left to right, females: 32, 37, 24, 23, 9, 9; males: 32, 35, 25, 19, 16, 12. A) PRS calculated using the combined sex summary statistics from the Neale Lab. There are no significant differences. B) PRS calculated for all individuals using the male-specific summary statistics from the Neale Lab. There are no significant differences. C) PRS calculated for all individuals using the female-specific summary statistics from the Neale Lab. There are no significant differences. D) PRS calculated for males and females separately using their respective summary statistics. There are no significant differences. (Summary stats: Neale Lab (2018))



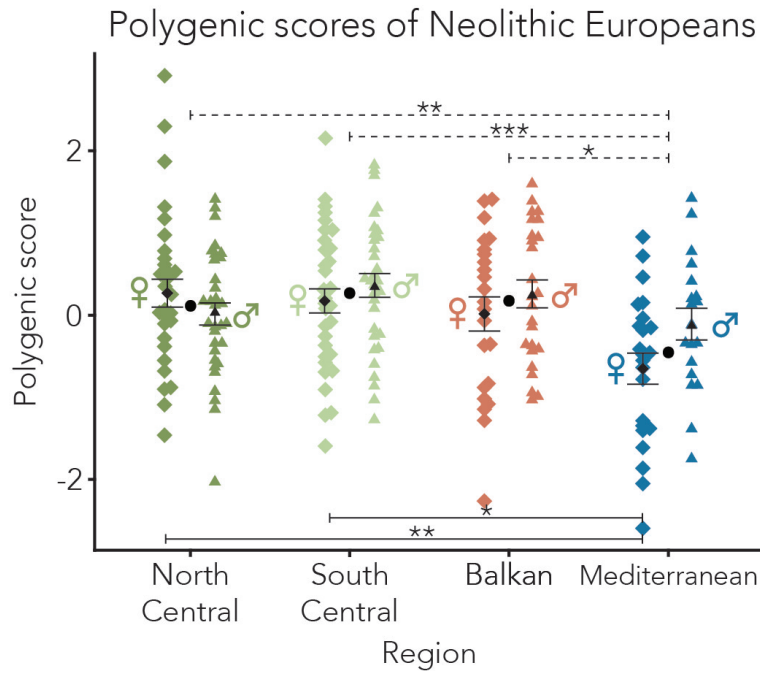
Supplementary Figure 3: Sibling-based PRS and femoral lengths

A) PRS calculated with between-sibling effect sizes using the LDpred2 model. Based on two-tailed pairwise *t*-tests adjusted for multiple comparisons, there are no differences between groups. Black points indicate group means. Black bars show the 95% confidence intervals. left to right, $n=64$, 72, 49, 42, 25, 21. B) Femur length (*y*-axis) increases with clumping/thresholding PRS (*x*-axis) though the effect is not significant likely due to the small sample size ($n=55$). Blue line is the regression line and grey area is the 95% confidence interval.



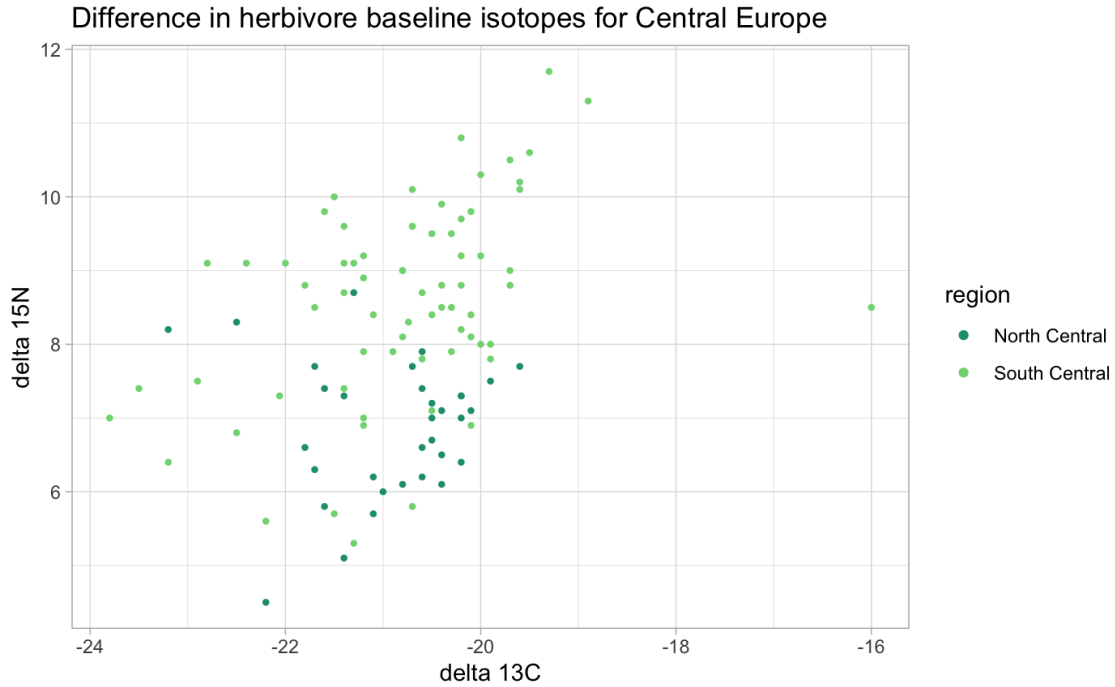
Supplementary Figure 4: PCA for unimputed DNA.

A) Unimputed PCA for Neolithic Europe, equivalent to Figure 2C in the main text. B) Unimputed PCA for all samples, equivalent to Figure 5A in the main text.



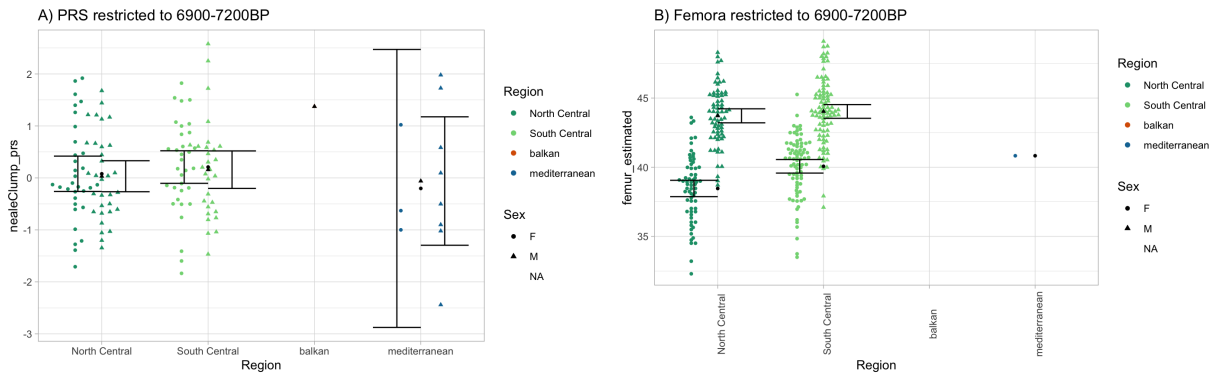
Supplementary Figure 5: LDpred2 PRS.

Polygenic scores calculated using infinitesimal LDpred 2 model (Privé, 2020). There is no difference between North Central, South Central or Balkans populations. Mediterraneans PRS are significantly lower than the other groups (linear regression, $p=0.002$). Similarly, the difference between males and females is only significant in the Mediterranean (two-sided pairwise t -test, adjusted for multiple comparisons, $p=0.04$). Sample sizes, left to right, males: 32, 35, 25, 19; females: 32, 37, 24, 23. Black points are group means, black bars represent 95% confidence intervals.



Supplementary Figure 6: Herbivore baseline.

Isotopic data from herbivores to establish a baseline for interpretation of human isotopic results. See citation in main text for data sources.



Supplementary Figure 7: Date restricted plots.

Black points are group means, black bars represent 95% confidence intervals. A) Plot of PRS data restricted to the same narrow date range used in the Central European population samples and illustrating the subsequent loss of power to detect effects. Sample sizes left to right, females: 32, 37, 0, 3; males: 32, 35, 1, 8. B) Plot of femora data restricted to the same date range as used in the Central European sample and illustrating the subsequent loss of power to detect effects. Sample sizes, left to right, females: 65, 88, 0, 1; males: 68, 100, 0, 0.