Supplementary information

Contents

Table S1

Number of high-quality biallelic variants in different datasets before and after LD pruning.

dataset	unpruned	biallelic variants			
		pruned			
		sliding window		difference between pruning windows	
		50k	10000k	count	percent (%)
GENOME	483100	232194	206253	25941	11.17%
EXOME	405496	184258	149453	34805	18.89%
BEADCHIP	567187	79492	72701	6791	8.54%
HUN-EXOME	408790	186984	151259	35725	19.11%

The unpruned datasets were LD pruned by Plink with two different sliding windows (50k and 10,000k respectively), and the difference of pruned markers between the two parameter settings is also summarized.

Table S2

f_3-statistics of Hungarian and selected South Asian populations.

Pop A	Pop B	Pop C	F3 value	Std Err	Z score	SNP count
STU	HUN	PJL	-0.003761	0.000140	-26.782	76834
ITU	HUN	PJL	-0.003641	0.000141	-25.752	76845
BEB	HUN	PJL	-0.002865	0.000138	-20.743	80622

Negative F_3 values with significant Z score indicate complex relations (admixture) between the Hungarian and South Asian populations.

Table S2

Figure S1

F_4(TSI, X; CHB; YRI) values of the analyzed datasets
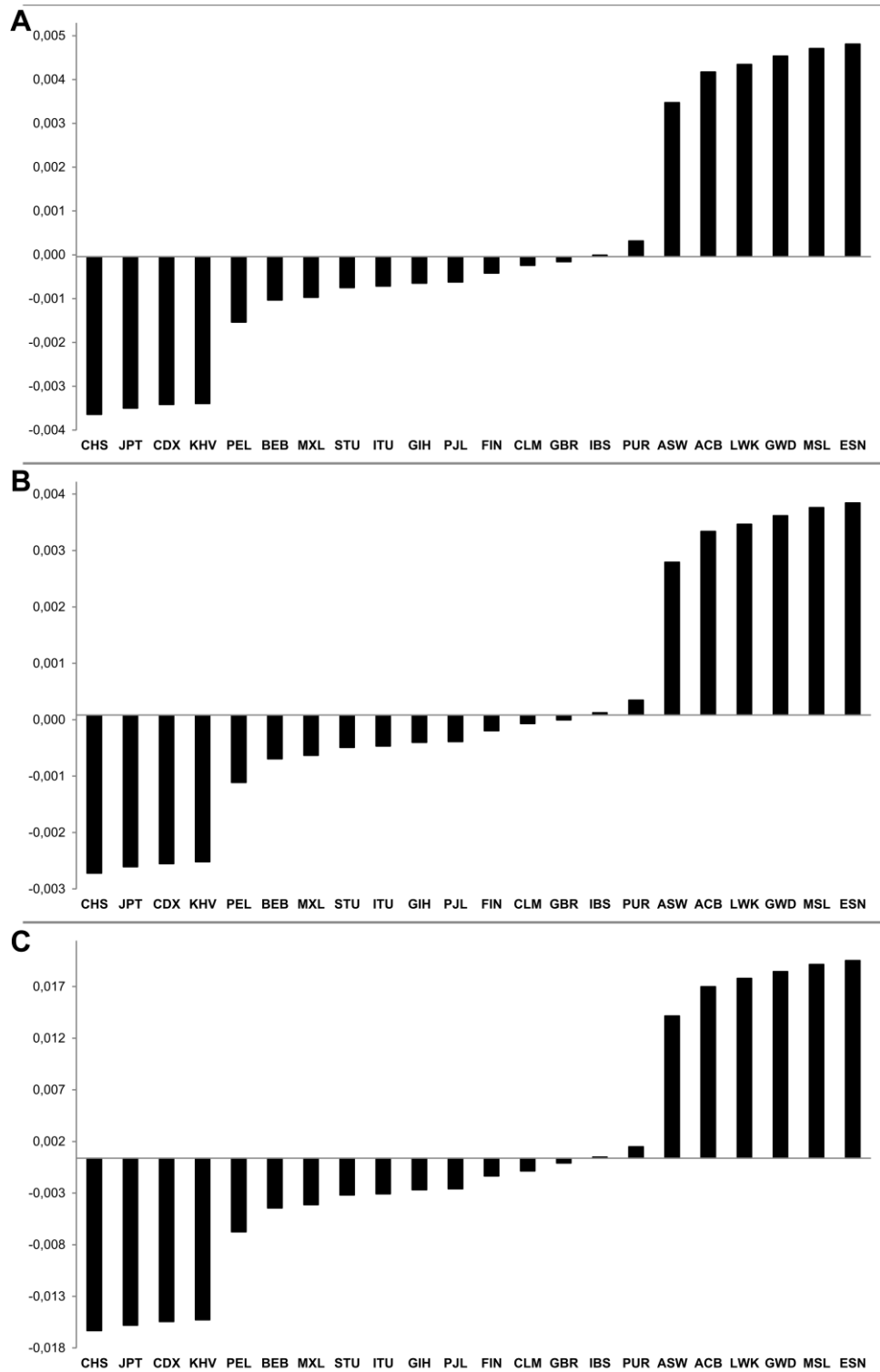
A) GENOME, B) EXOME and C) BEADCHIP, where X denotes the test population indicated on the x-axis.

Figure S2

Calculated Cross Validation error for different admixture models (K=3 to 10) for each dataset.
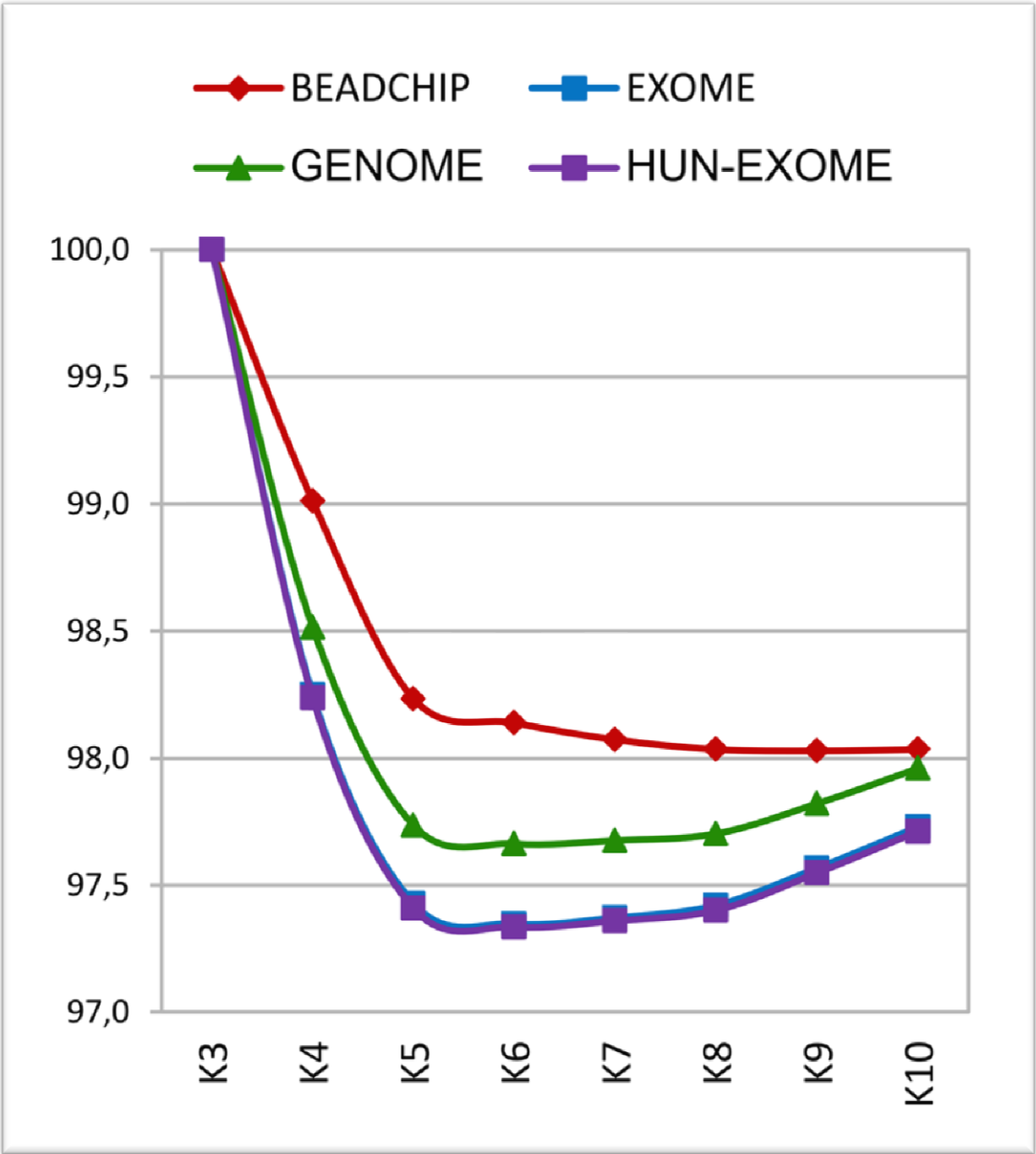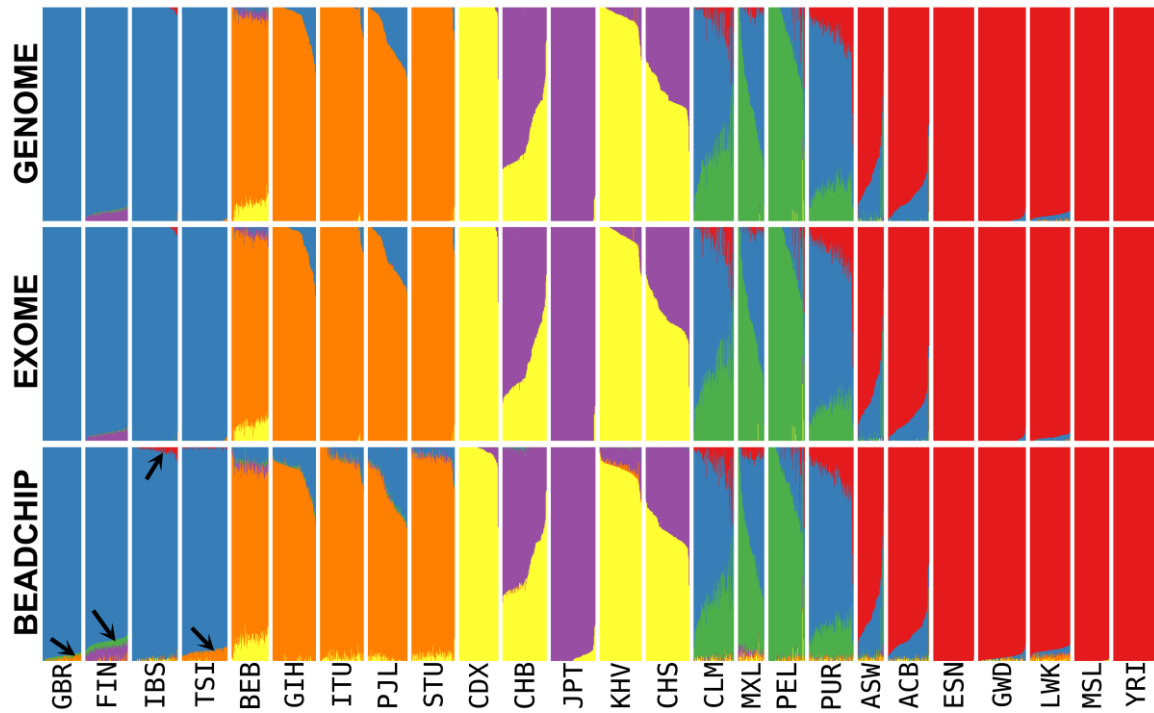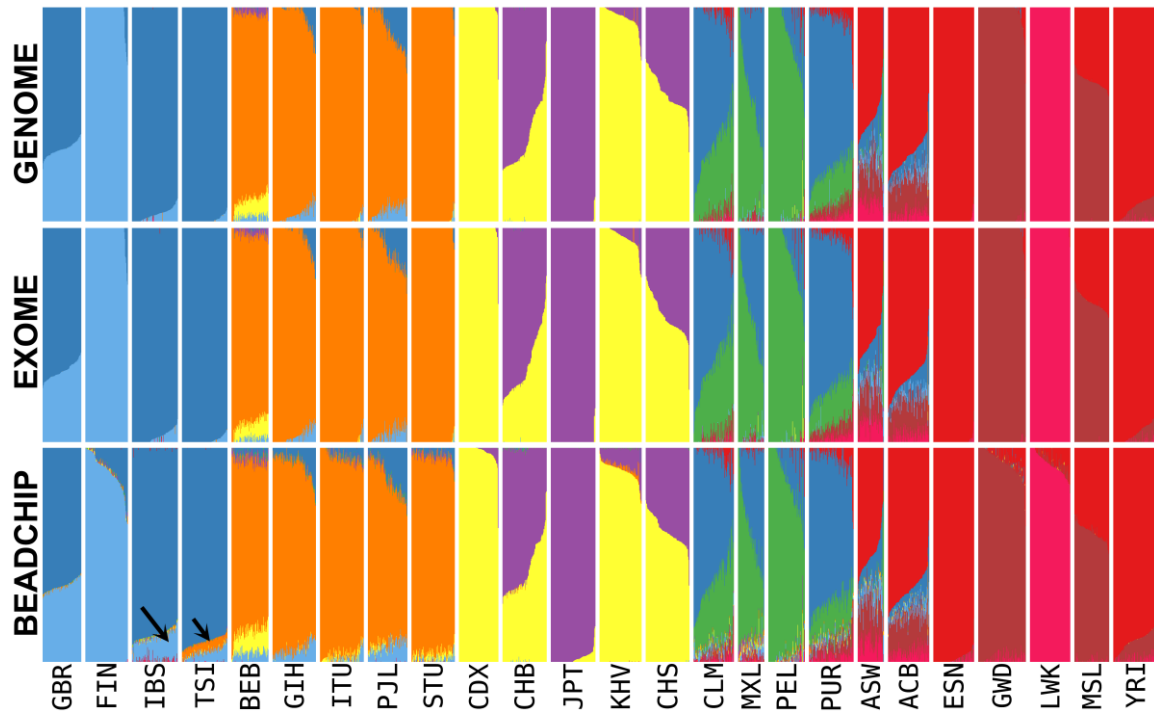
Figure S3

Admixture analysis of 25 populations using the GENOME, EXOME and BEADCHIP datasets with the

K=6 admixture model.

The arrows denote examples of minor genetic admixture components within the European super-
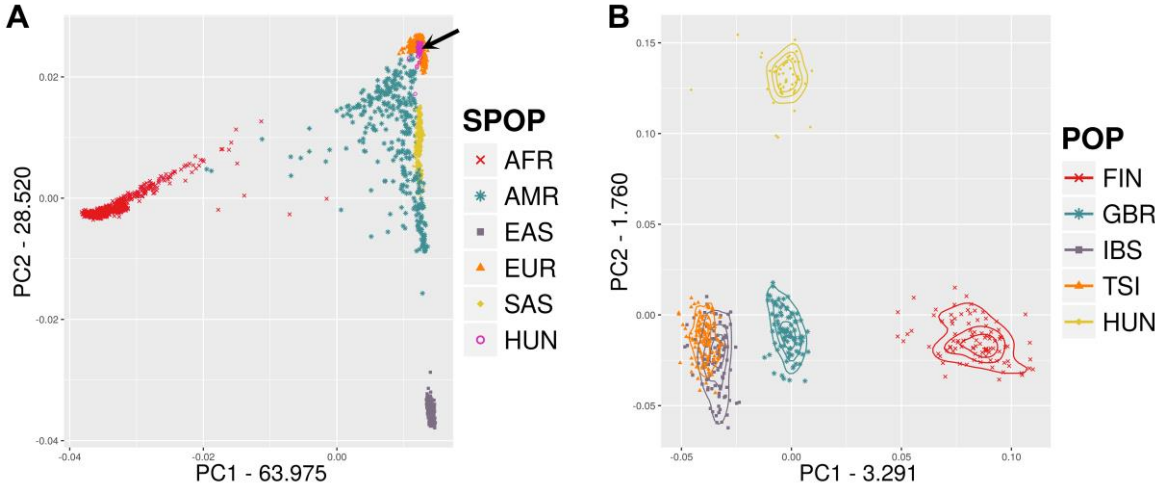
population.

Figure S4

Admixture analysis of 25 populations using the GENOME, EXOME and BEADCHIP datasets with the K=9 admixture model.

The arrows denote examples of minor genetic admix components within the European super-population.
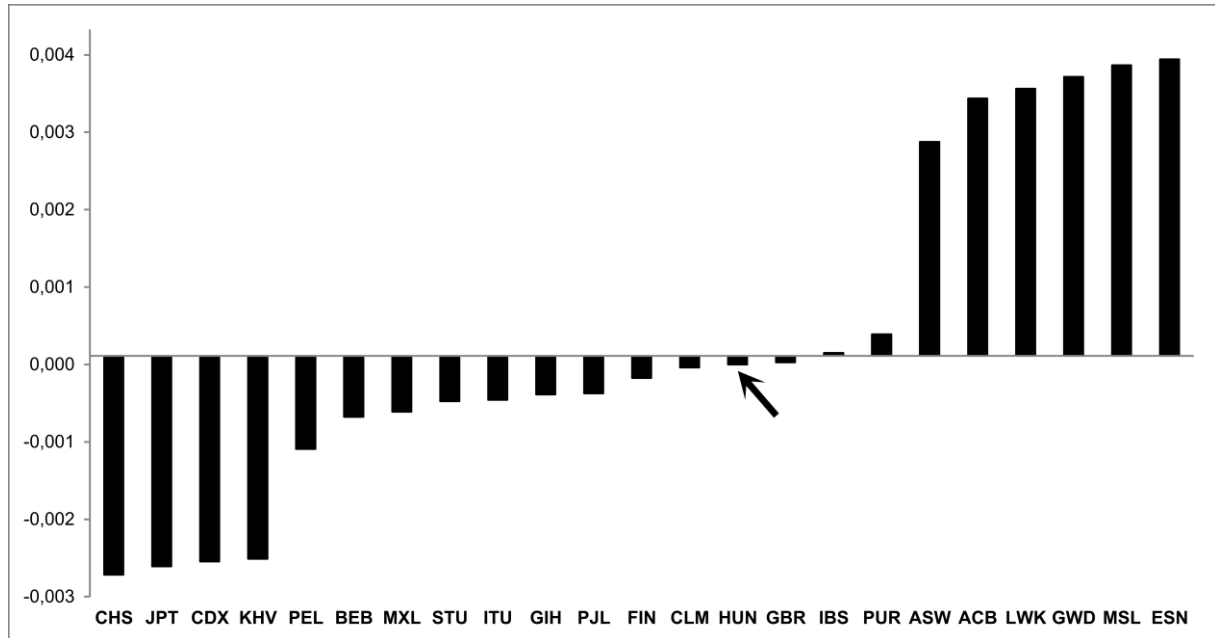
Figure S5

PCA of the HUN-EXOME dataset

 A) Hungarians (highlighted by the arrow) in relation to the super-population, **B)** Hungarians in

relation to European populations.

Figure S6

F4(TSI, X; CHB, YRI) values of the HUN-EXOME dataset, where X denotes the test population indicated on the x-axis.

The arrow shows the relative position of Hungarians, indicating that they have higher East Asian admixture component than British but lower than Finnish population.

Figure S7

Admixture analysis of 25 populations using the HUN-EXOME dataset for the **A)** K=6 and **B)** K=9 admixture models.
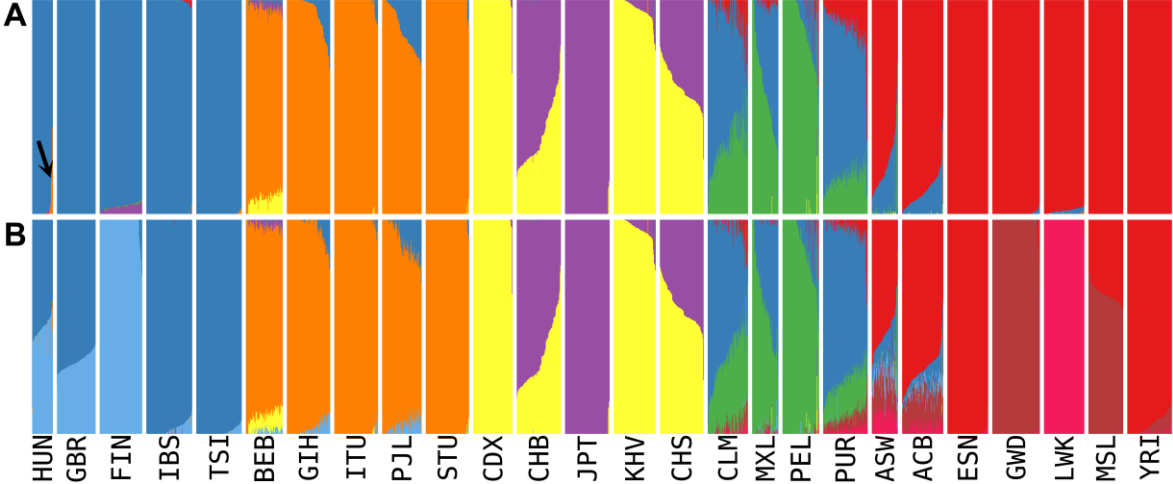
The arrow denotes Hungarian individuals with substantial South Asian admixture components.

Data

We have deposited the Supplementary Data to figshare:

https://figshare.com/s/e91794c7141a7eb16255