

1 **Selection signatures in worldwide Sheep populations -**
2 **Supplementary material**

3 Maria-Ines Fariello¹, Bertrand Servin¹, Gwenola Tosser-Klopp¹, Rachel Rupp¹, Carole Moreno¹, Magali
4 SanCristobal¹, Simon Boitard^{3,4,*}

5 **1 Génétique, Physiologie et Systèmes d'Élevage, INRA & ENVT & ENSAT,**
6 **Castanet-Tolosan, France**

7 **3 Génétique Animale et Biologie Intégrative, INRA & AgroParisTech, Jouy-en-Josas,**
8 **France**

9 **4 Origine, Structure et Evolution de la Biodiversité, Museum National d'Histoire**
10 **Naturelle & EPHE & CNRS, Paris, France**

11 *** E-mail: Corresponding sboitard@mnhn.fr**

Candidate Gene	Populations
BNC2	COM
KITLG	COM, EMZ
KIT	VBS
EDN3	MOG
MC1R	SUF, SAB, GAL
ASIP	MER
MITF	VBS, ERS, BOS

Table S1. List of genes potentially associated with coloration patterns found under selection signatures and the most likely candidate populations for selection.

group	K	μ	σ
AFR	5	1.97	0.53
ASI	25	7.78	0.92
CEU	10	3.41	0.62
ITA	15	2.83	0.44
NEU	40	4.09	0.45
SWA	10	3.28	0.55
SWE	25	3.32	0.39

Table S2. Parameters for the hapFLK genome scan. K : Number of haplotype clusters used in the LD model for each group, as determined by the fastPHASE cross-validation procedure. μ, σ mean and standard deviation of the normal distribution used to model the hapFLK neutral distribution

group	DF	\overline{FLK}
AFR	1	0.97
ASI	7	6.65
CEU	3	2.91
ITA	3	2.93
NEU	5	4.78
SWA	3	2.92
SWE	3	2.96

Table S3. Theoretical number of degrees of freedom (DF) and observed mean of the FLK statistic (\overline{FLK}) in each group of populations.

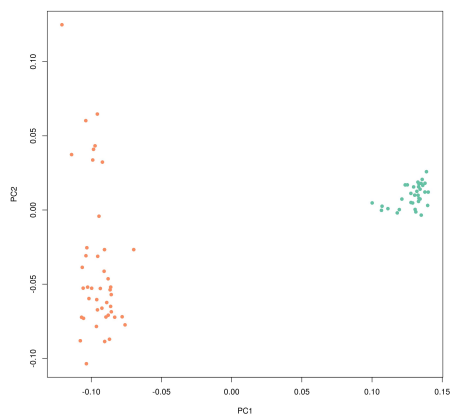


Figure S1. Projection of animals from the African group on the first 2 principal components.

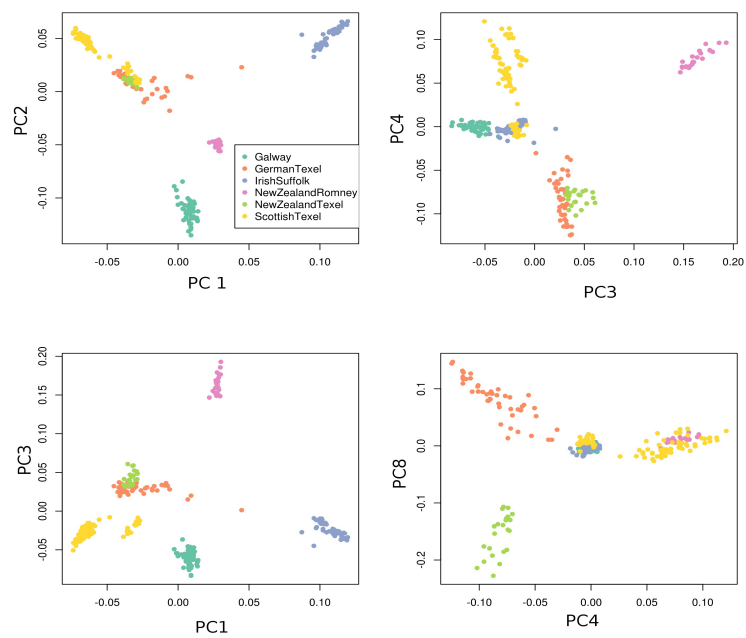


Figure S2. Projection of animals from the North European group on the first 8 principal components.

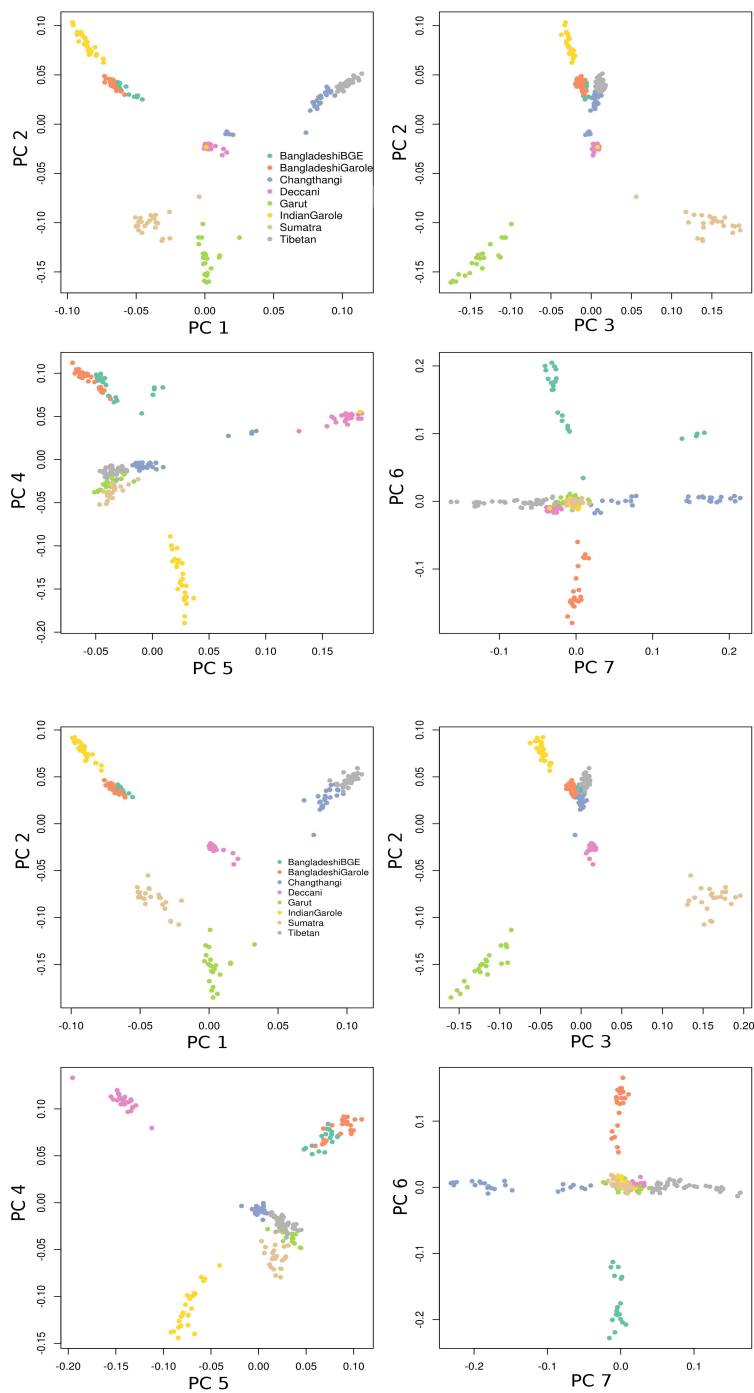


Figure S3. Projection of animals from the Asian group on the first 8 principal components, before (top) and after (bottom) correction. One Indian Garole animal located in the Deccani cluster, was attributed to this breed. Five Bangladeshi BGE animals clustering away from the rest of the breed were removed. Four Changthangi animals, which clustered away from the rest of the breed and appeared admixed with the Deccani breed, were removed. One outlier Sumatra animal was also removed.

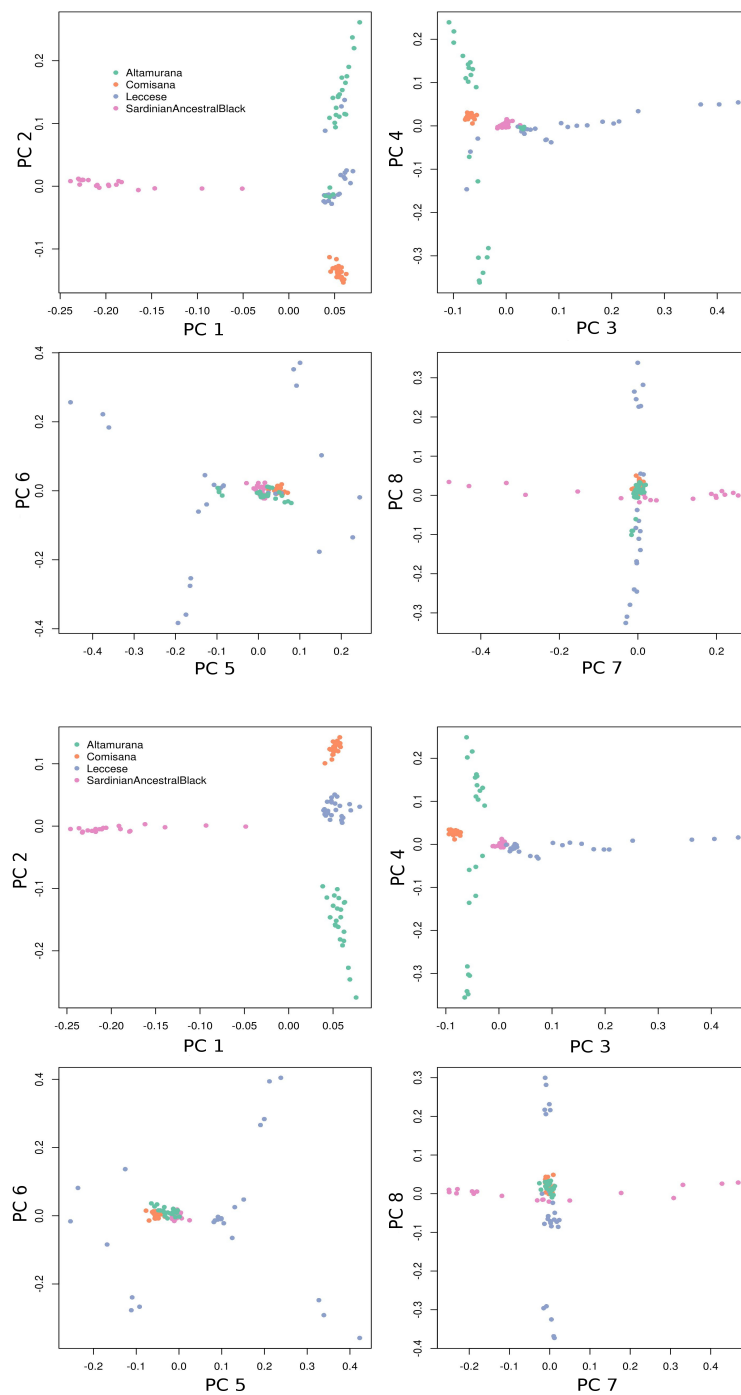


Figure S4. Projection of animals from the Italian group on the first 8 principal components, before (top) and after (bottom) correction. Four Altamurana animals located in the Lecceese cluster were attributed to the Lecceese breed. Similarly, three Lecceese animals located in the Altamurana cluster was attributed to the Altamurana cluster.

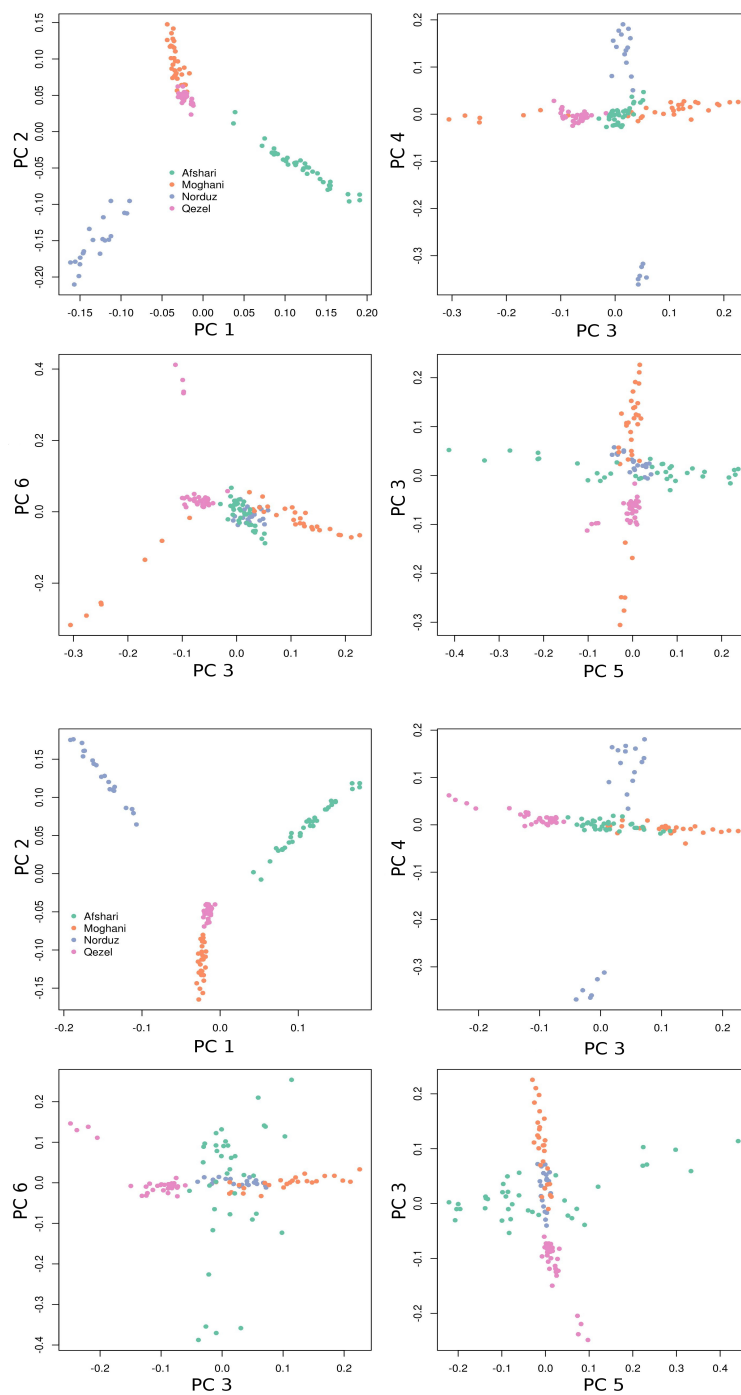


Figure S5. Projection of animals from the South West Asian group on the first 8 principal components, before (top) and after (bottom) correction. Seven Moghani animals clustering away from the rest of the breed, and possibly admixed with the Qezel breed, were removed.

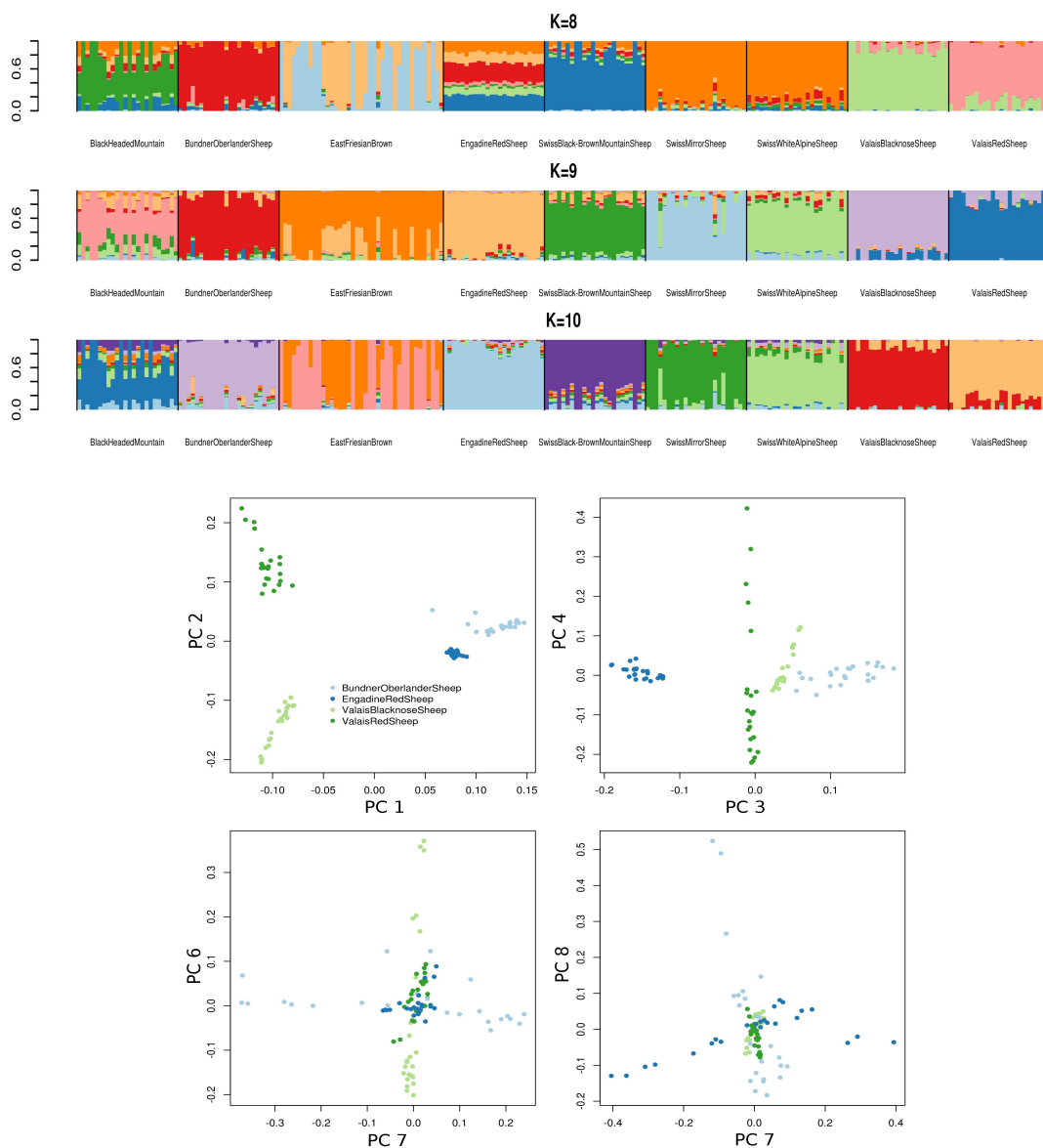


Figure S6. Admixture analysis for animals of the Central European group (top), and projection of the selected animals on the first 8 principal components (bottom). Three breeds (Swiss Mirror Sheep, Swiss Alpine White Sheep and Swiss Black-Brown Mountain Sheep) were reported as admixed breeds in [1] and were consequently removed, although they appeared relatively homogeneous in the admixture analysis. Two further breeds (East Friesian Brown and Black Headed Mountain) were removed based on the admixture analysis.

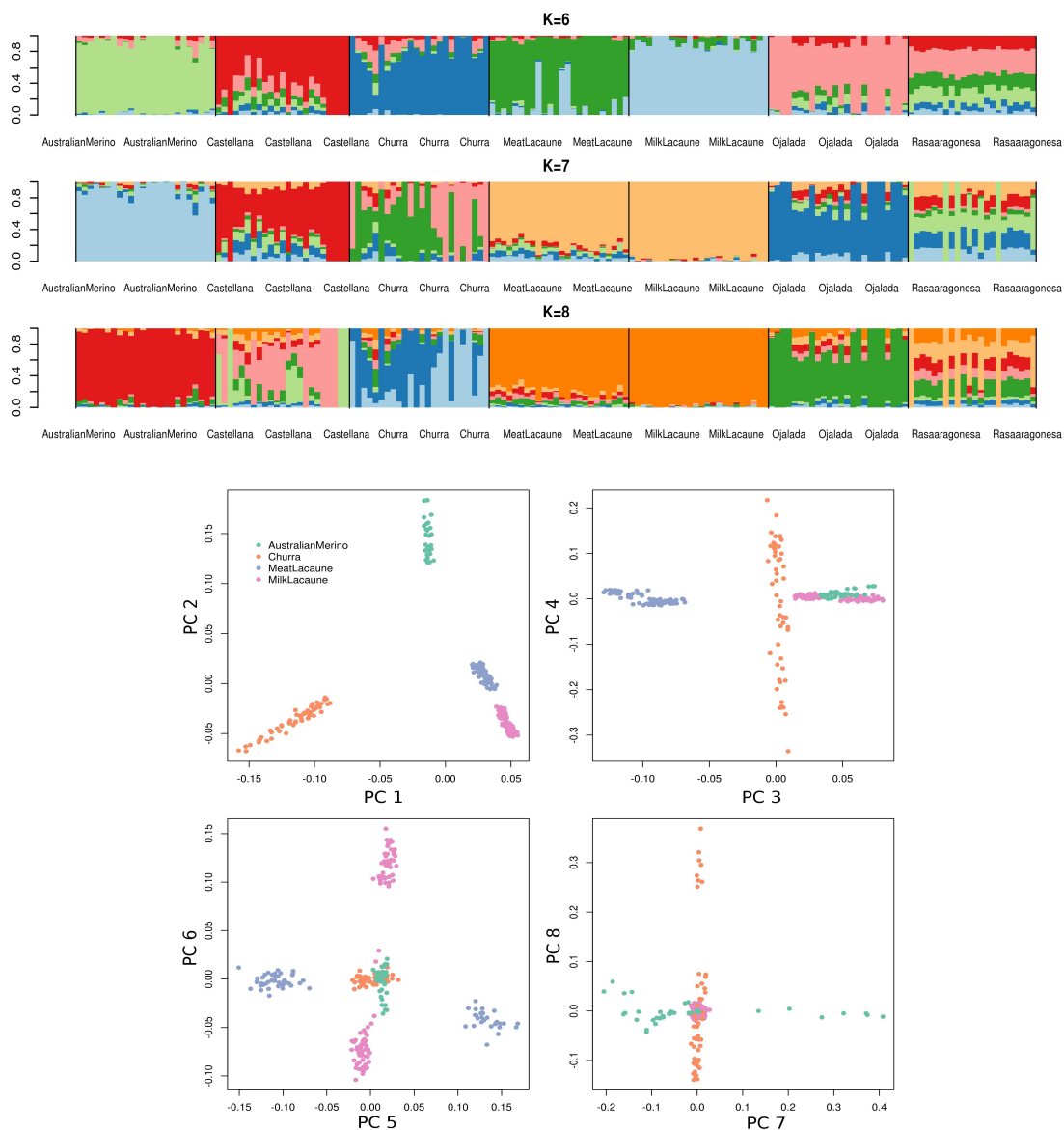


Figure S7. Admixture analysis for animals of the South West European group (top), and projection of the selected animals on the first 8 principal components (bottom). For the admixture analysis, a subset of 24 animals was sampled at random within each breed in order to obtain balanced sample sizes. Three breeds (Castellana, Ojalada and Rasaaragonesa) were removed based on the admixture analysis, as they were clearly admixed. Historical records concerning those 3 breeds are ambiguous, some of them reporting that they may result from a cross between Merino and Churra animals. The Churra breed also looked admixed in this analysis, but this was no longer the case after removing the three other problematic breeds.

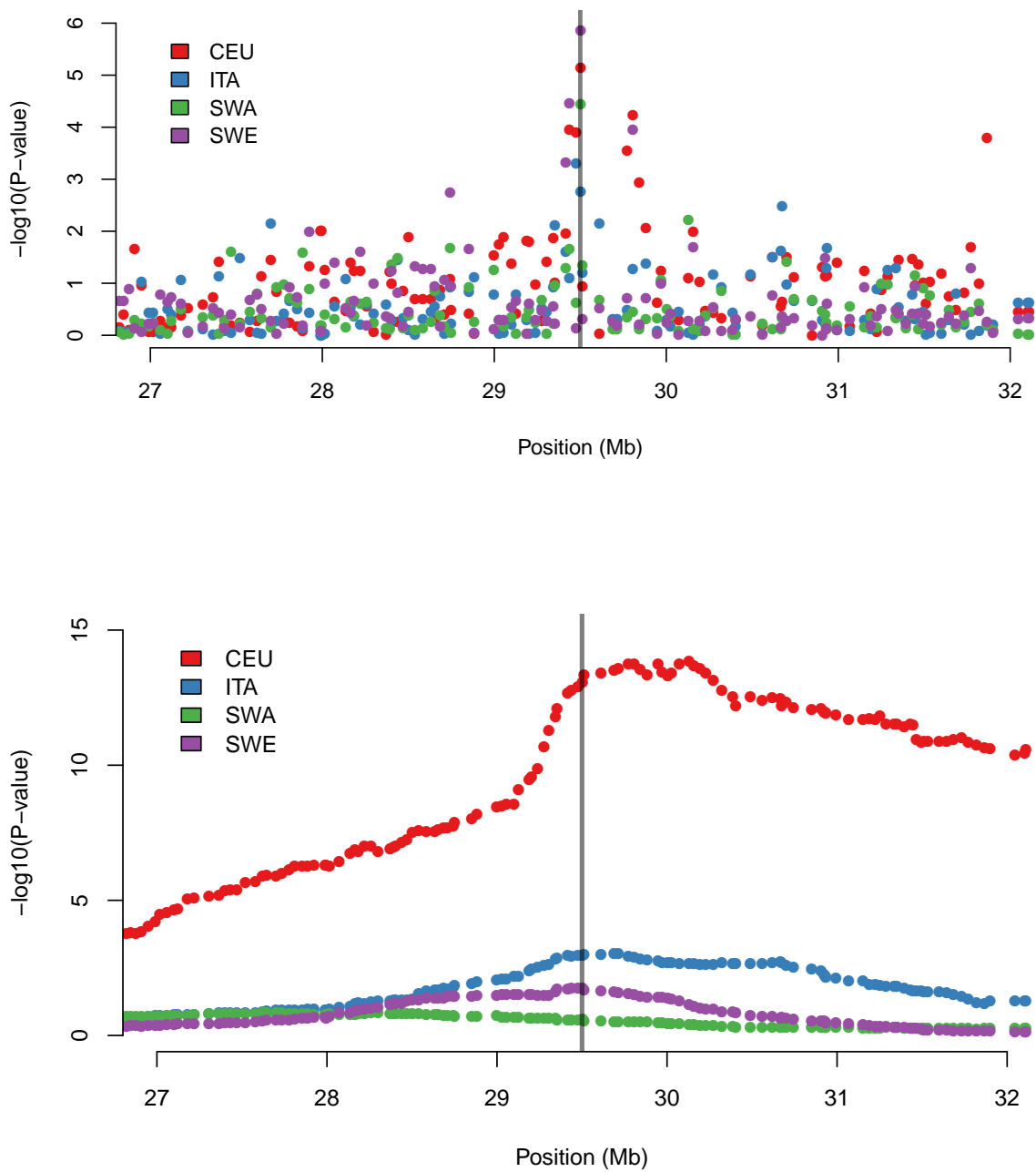


Figure S8. FLK (top) and hapFLK (bottom) p-values in the region of RXFP2. The position of RXFP2 is indicated by the vertical gray line.

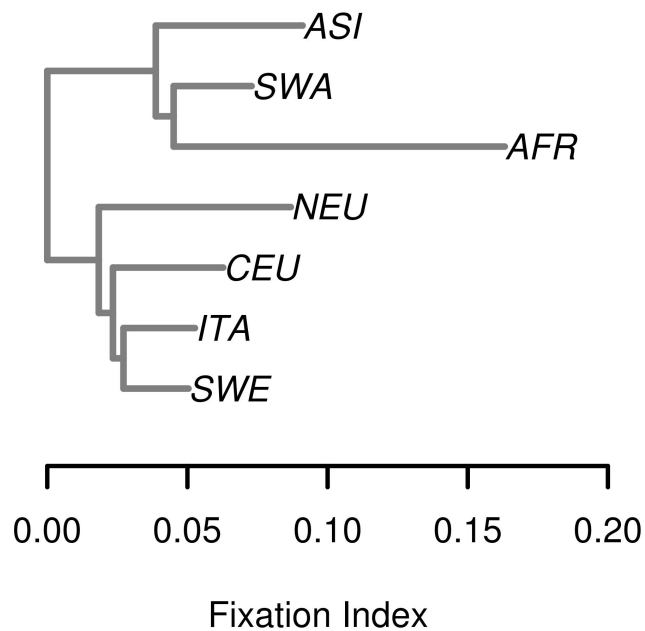


Figure S9. Phylogenetic tree of the ancestral populations of geographical groups.

12 References

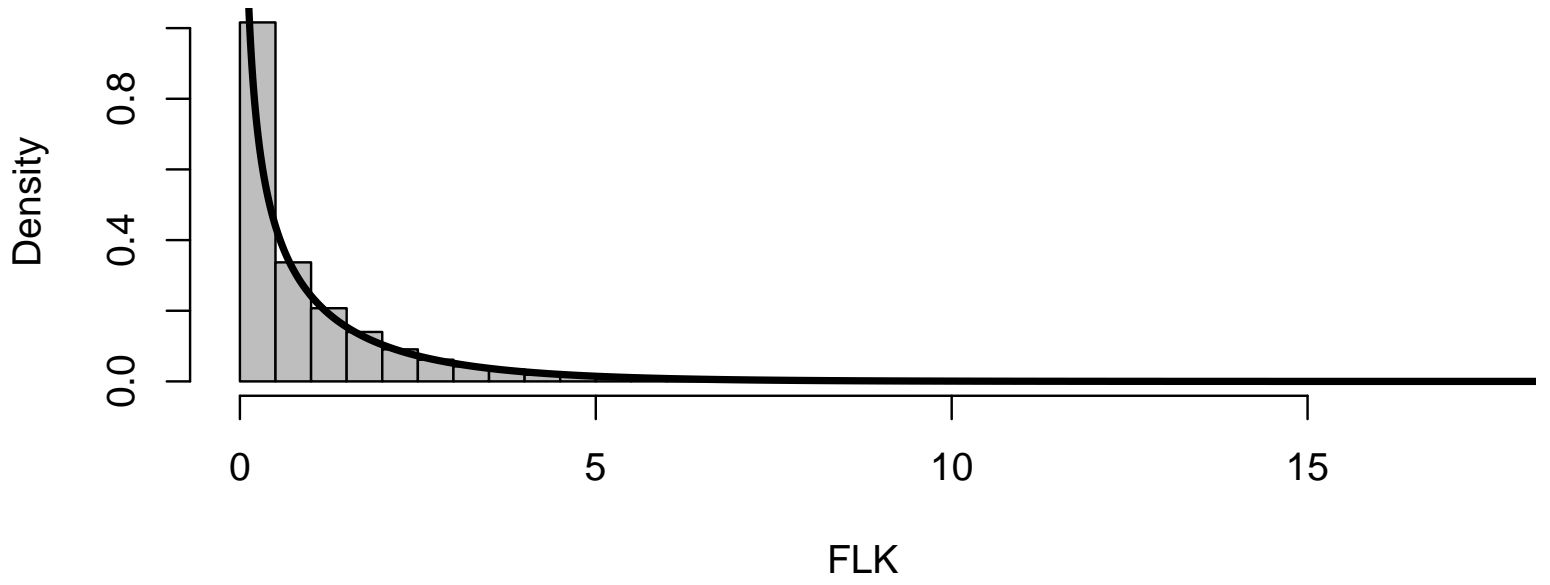
- 13 1. Kijas JW, Lenstra JA, Hayes B, Boitard S, Porto Neto LR, et al. (2012) Genome-wide analysis of
14 the world's sheep breeds reveals high levels of historic mixture and strong recent selection. PLoS
15 Biol 10: e1001258.

16 Text S1. FLK and hapFLK genome scans within groups of populations

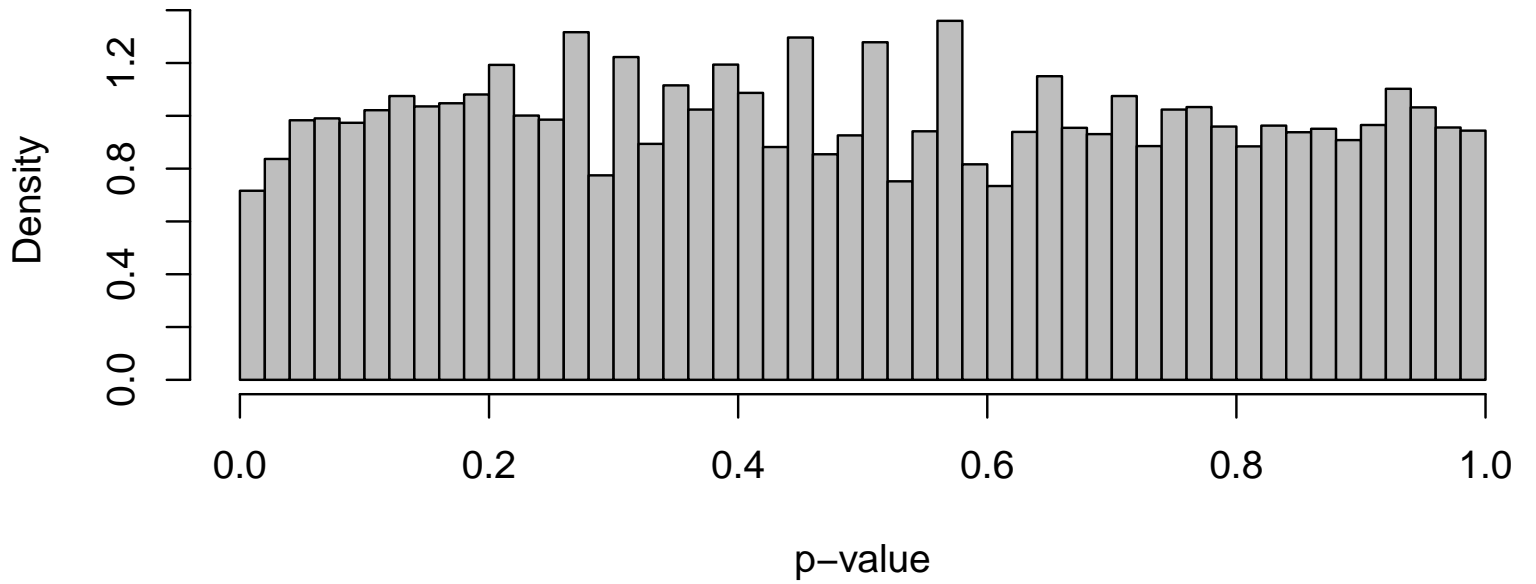
17 For each population group, we show, in this order

- 18 1. histogram of the observed FLK distribution, and corresponding theoretical χ^2 distribution
- 19 2. the FLK p-value distribution
- 20 3. histogram of the hapFLK distribution and corresponding estimated normal distribution (see details
21 in Methods)
- 22 4. the hapFLK p-value distribution
- 23 5. Manhattan plot of the FLK p-values
- 24 6. Manhattan plot of the hapFLK p-values

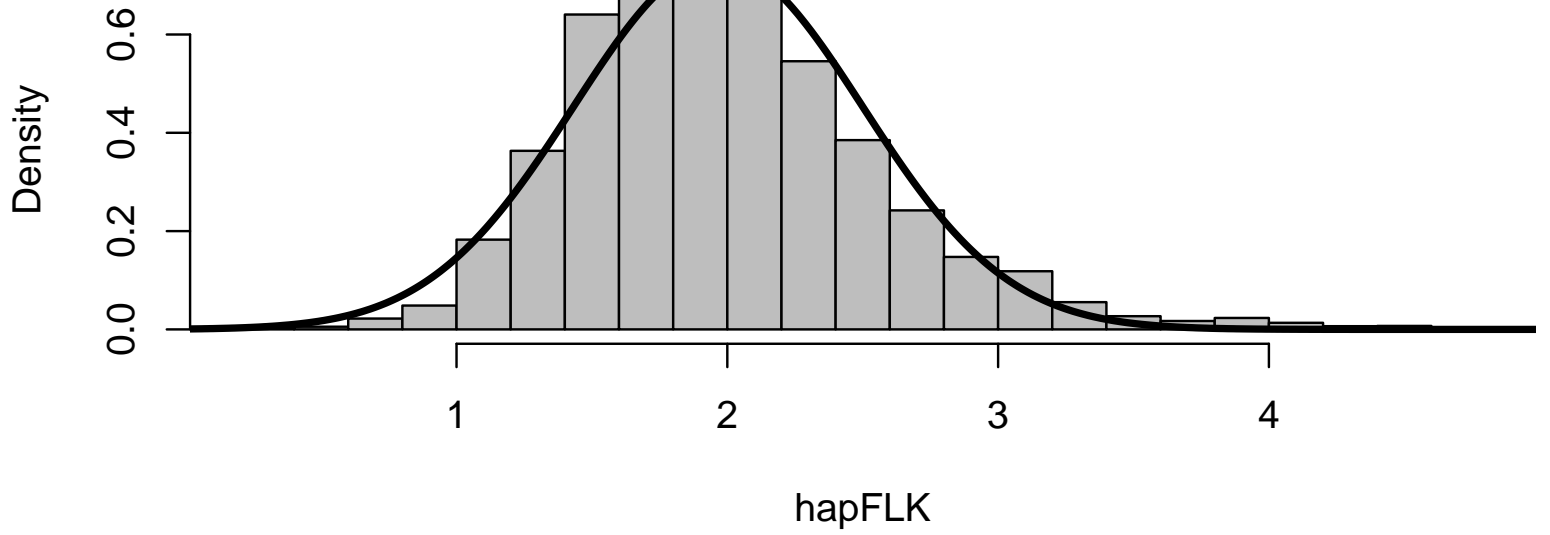
FLK distribution for group AFR



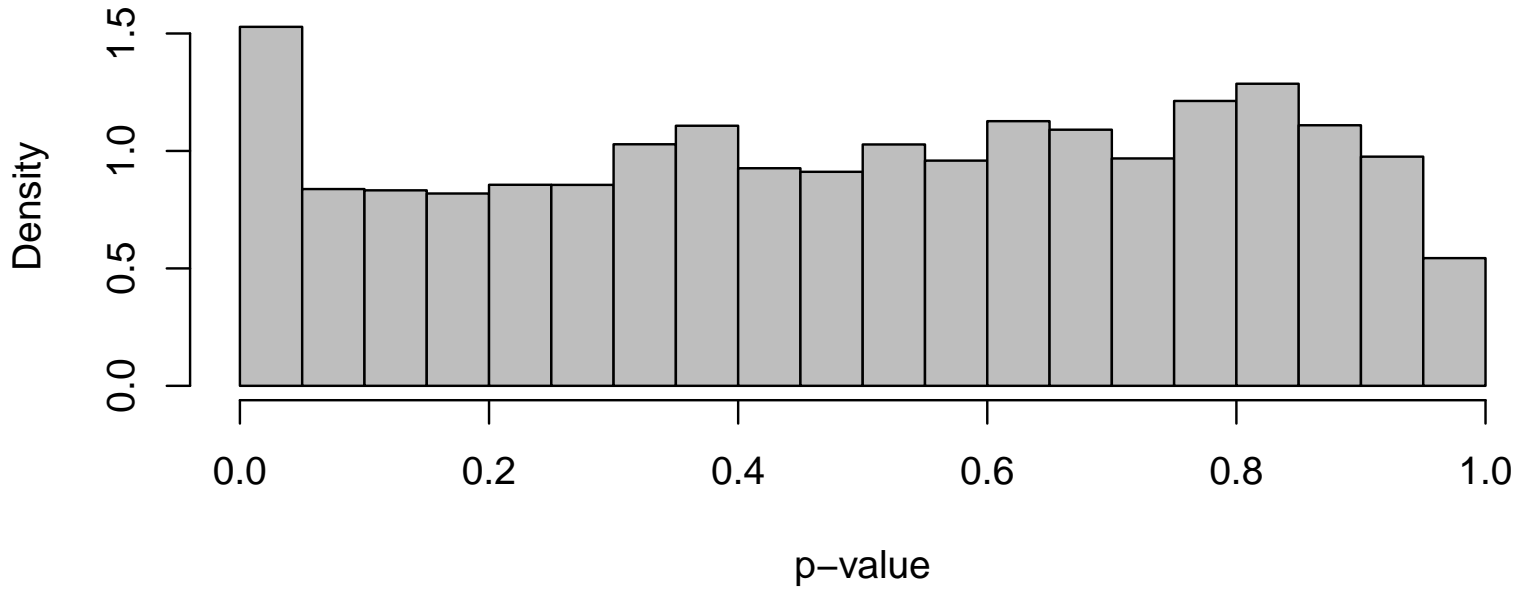
P-value distribution for group AFR



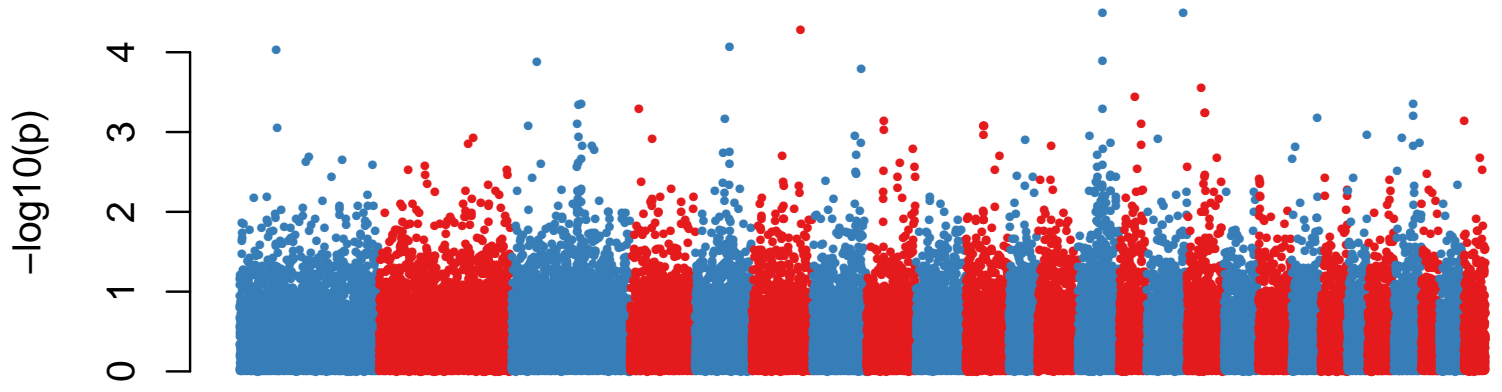
hapFLK distribution for group AFR



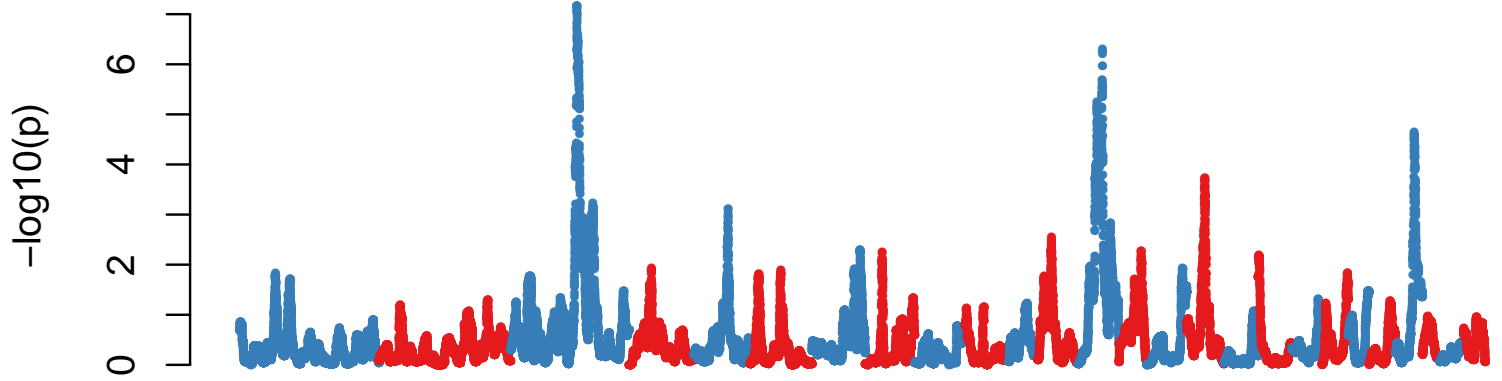
P-value distribution for group AFR



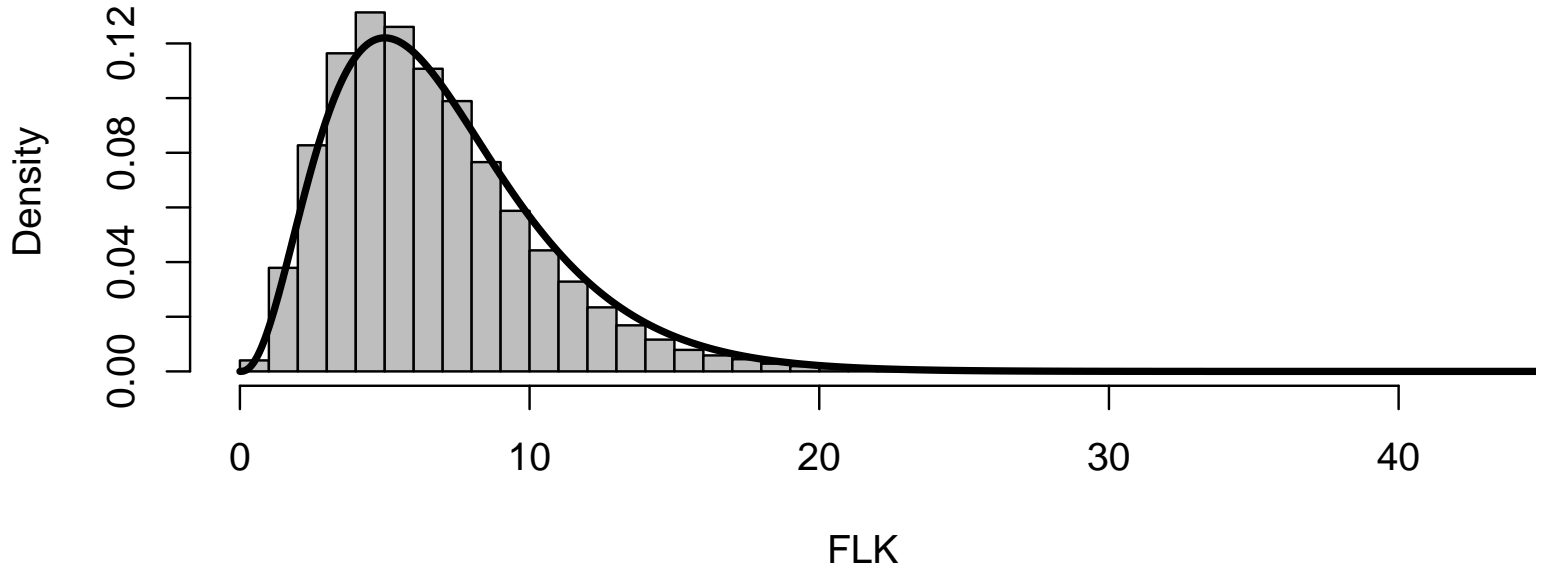
FLK genome scan for group AFR



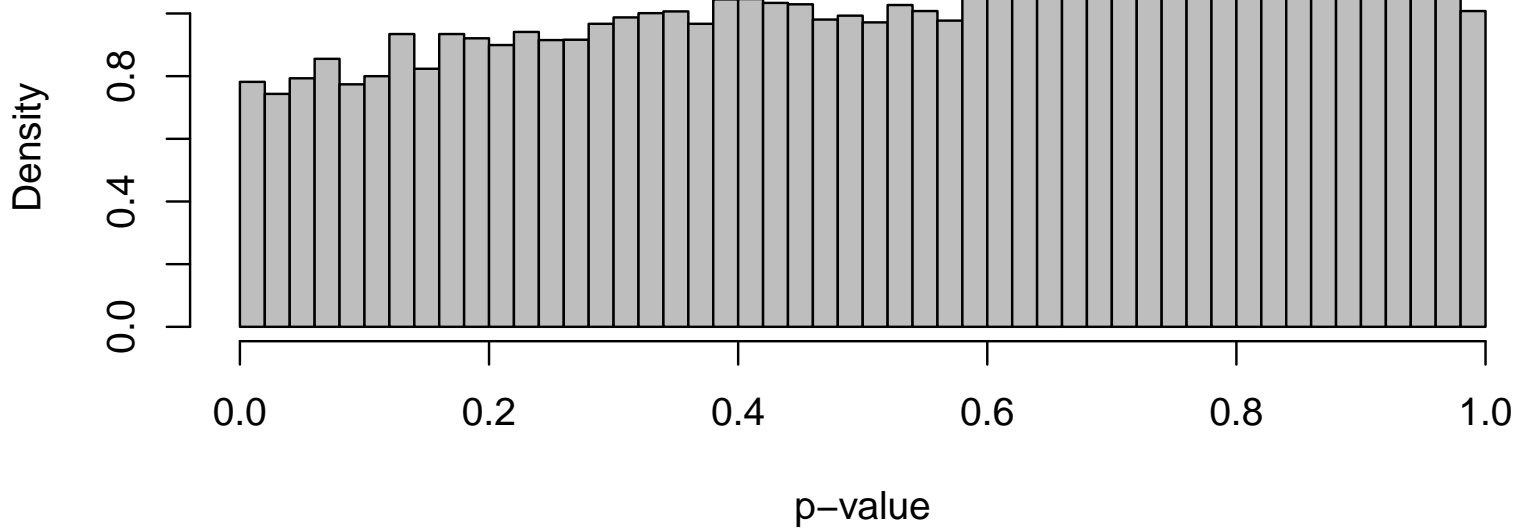
hapFLK genome scan for group AFR



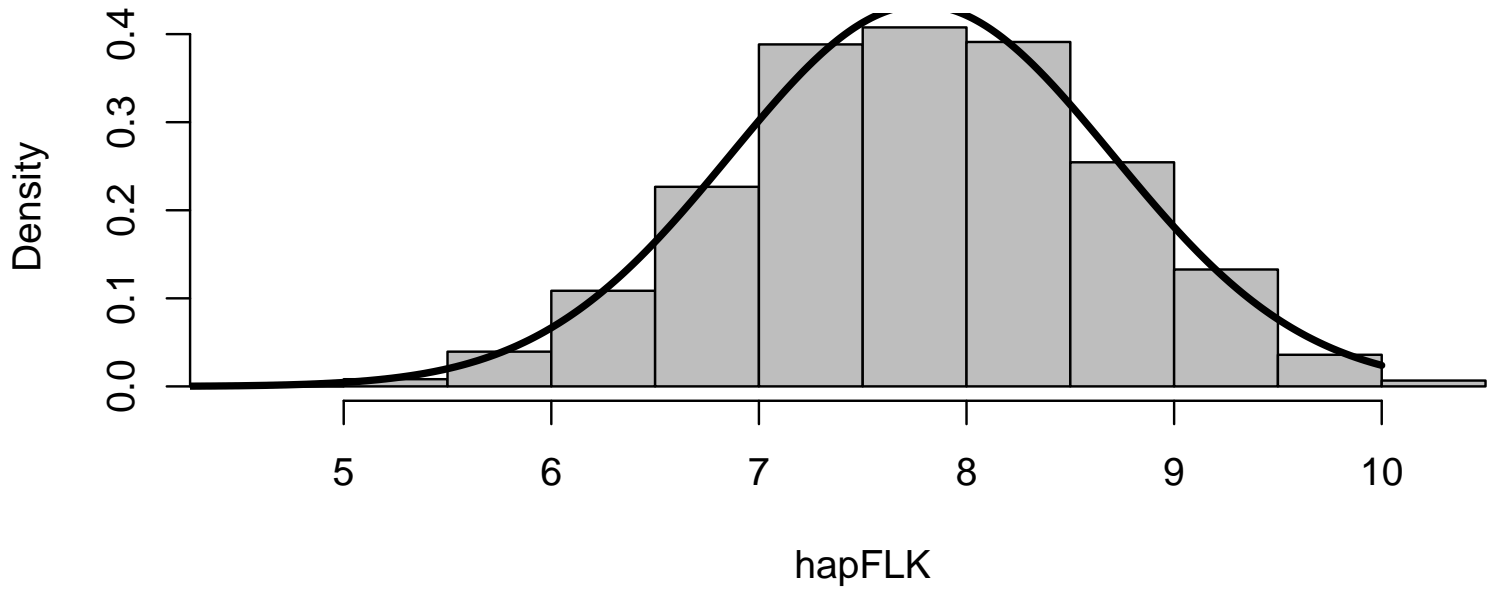
FLK distribution for group ASI



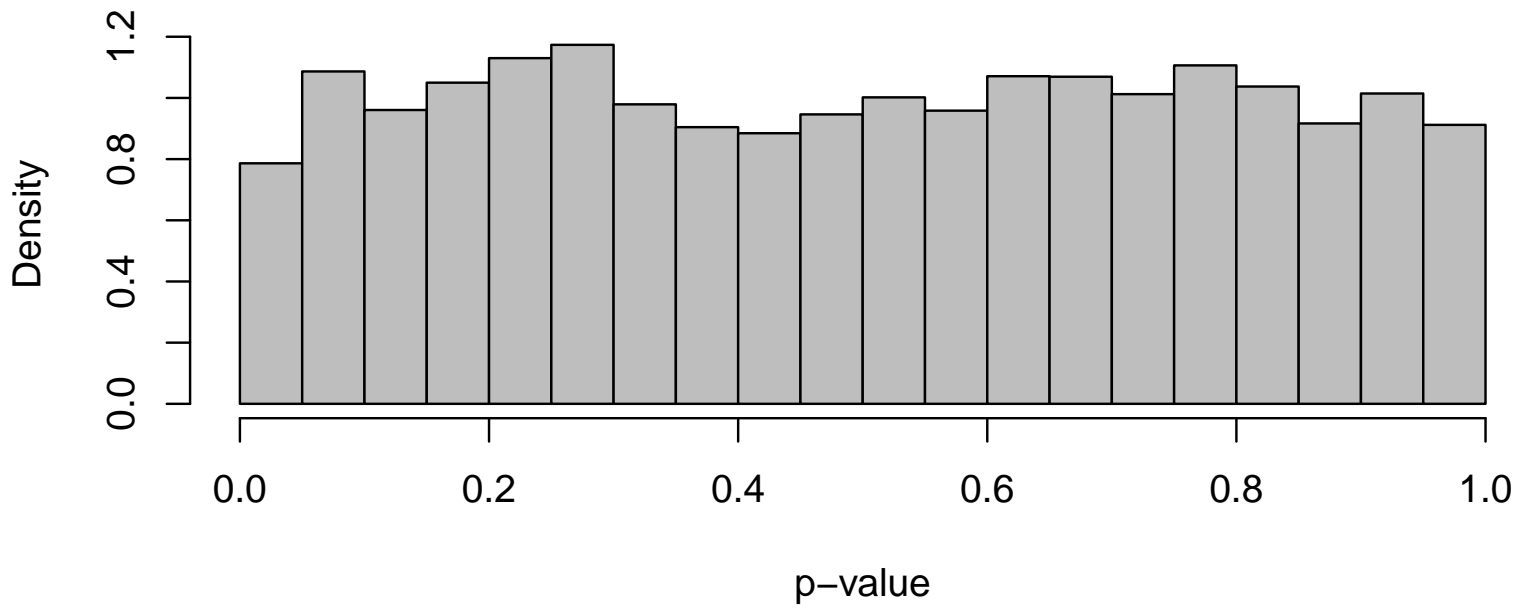
P-value distribution for group ASI



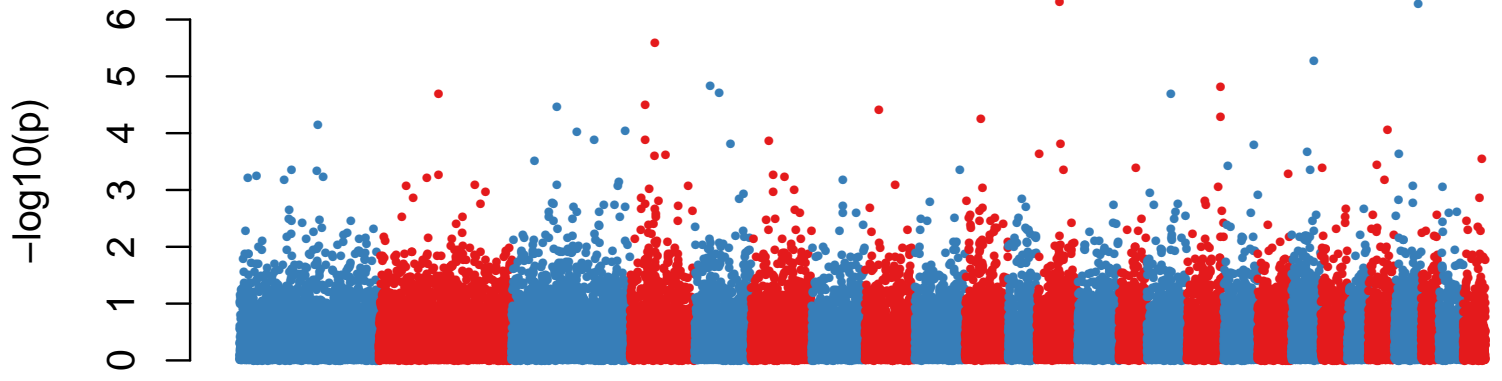
hapFLK distribution for group ASI



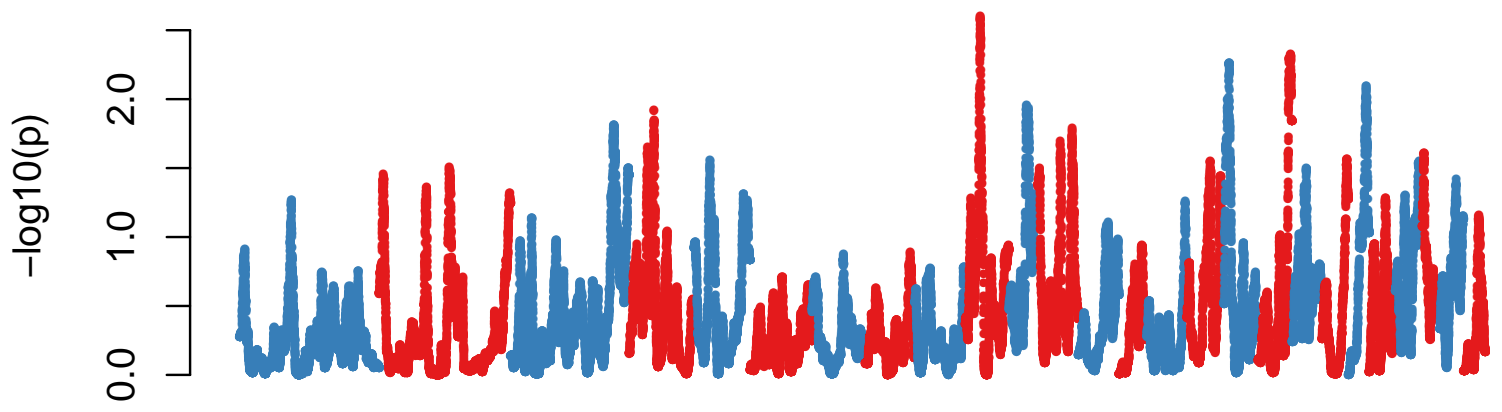
P-value distribution for group ASI



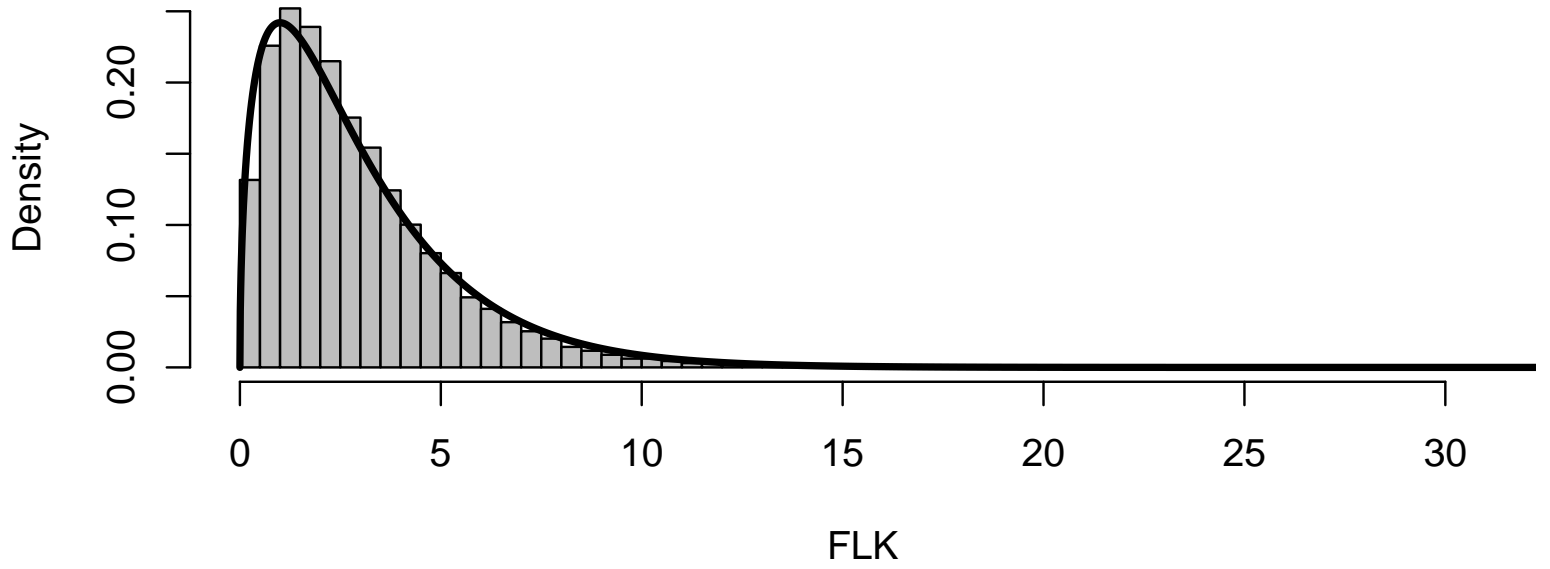
FLK genome scan for group ASI



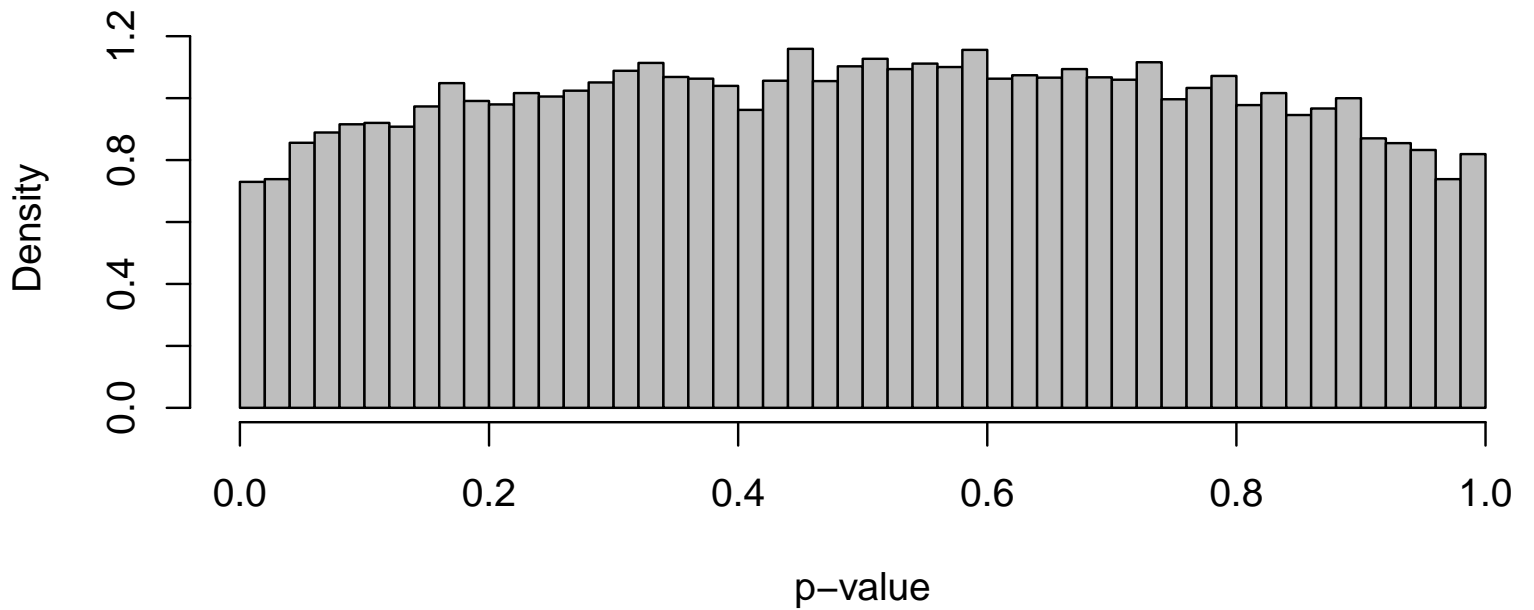
hapFLK genome scan for group ASI



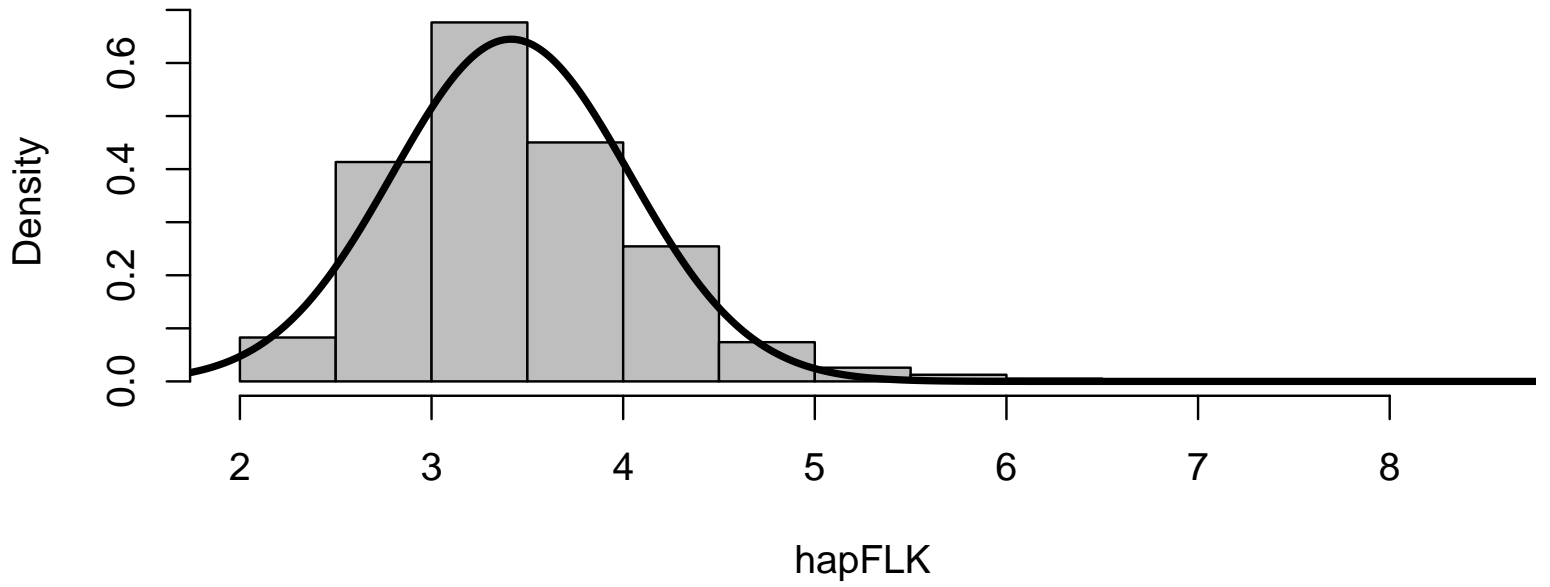
FLK distribution for group CEU



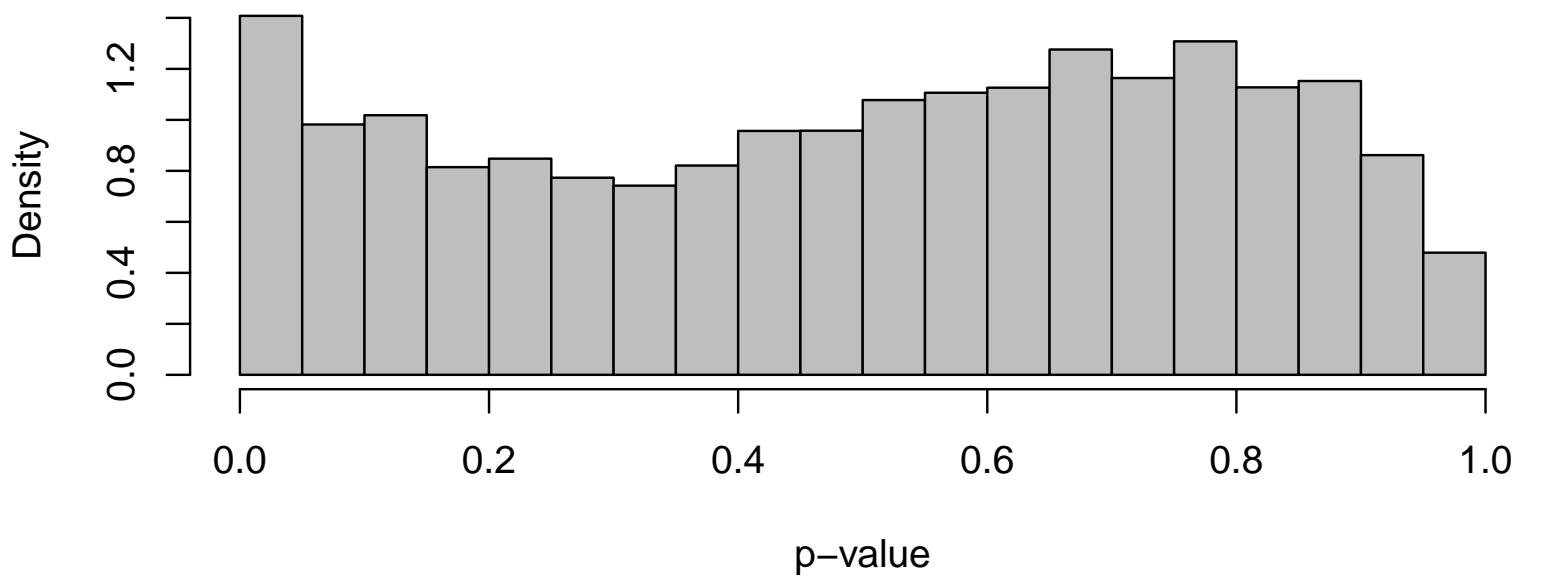
P-value distribution for group CEU



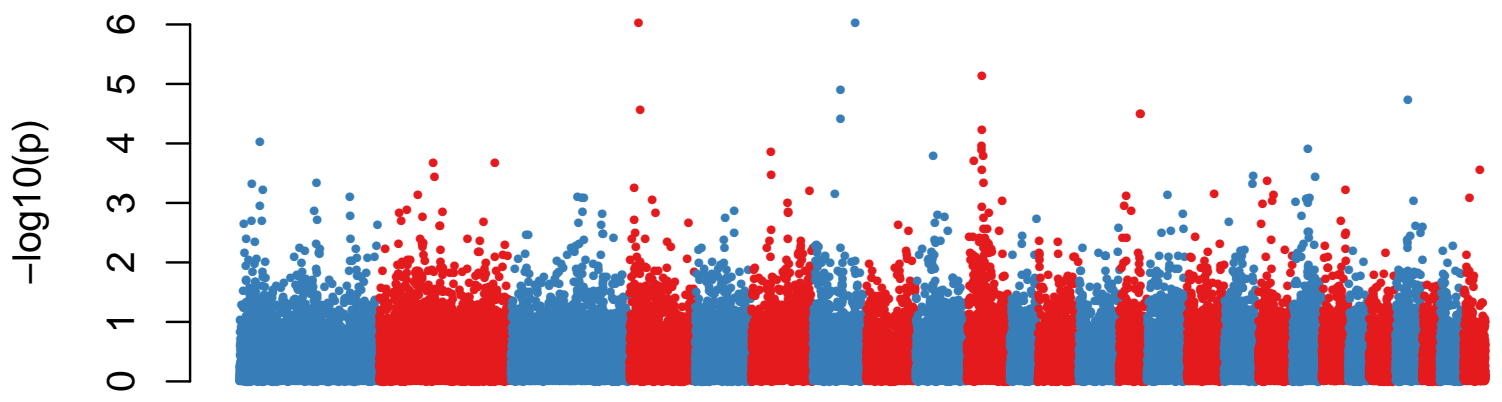
hapFLK distribution for group CEU



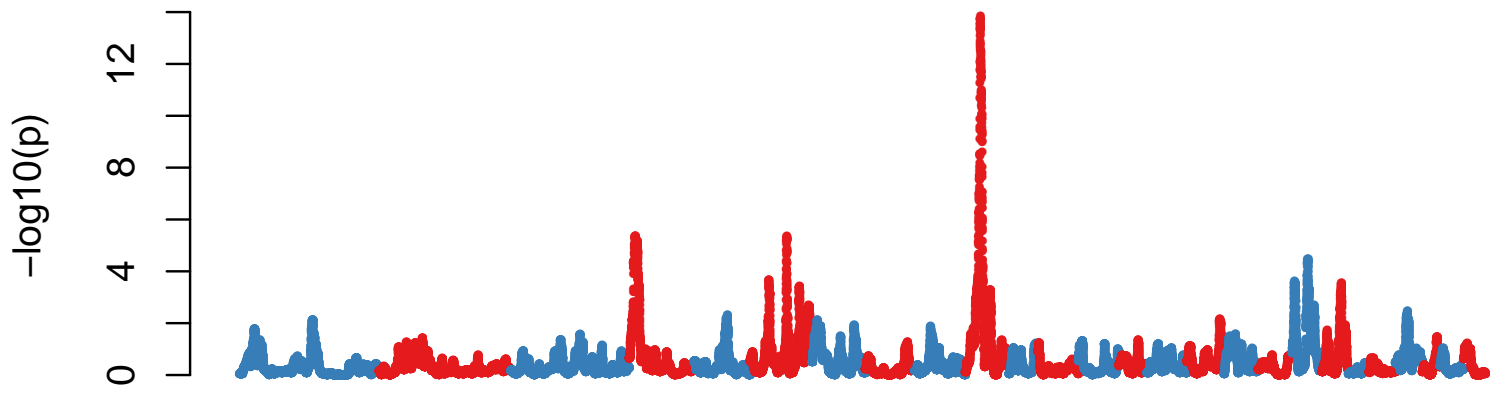
P-value distribution for group CEU



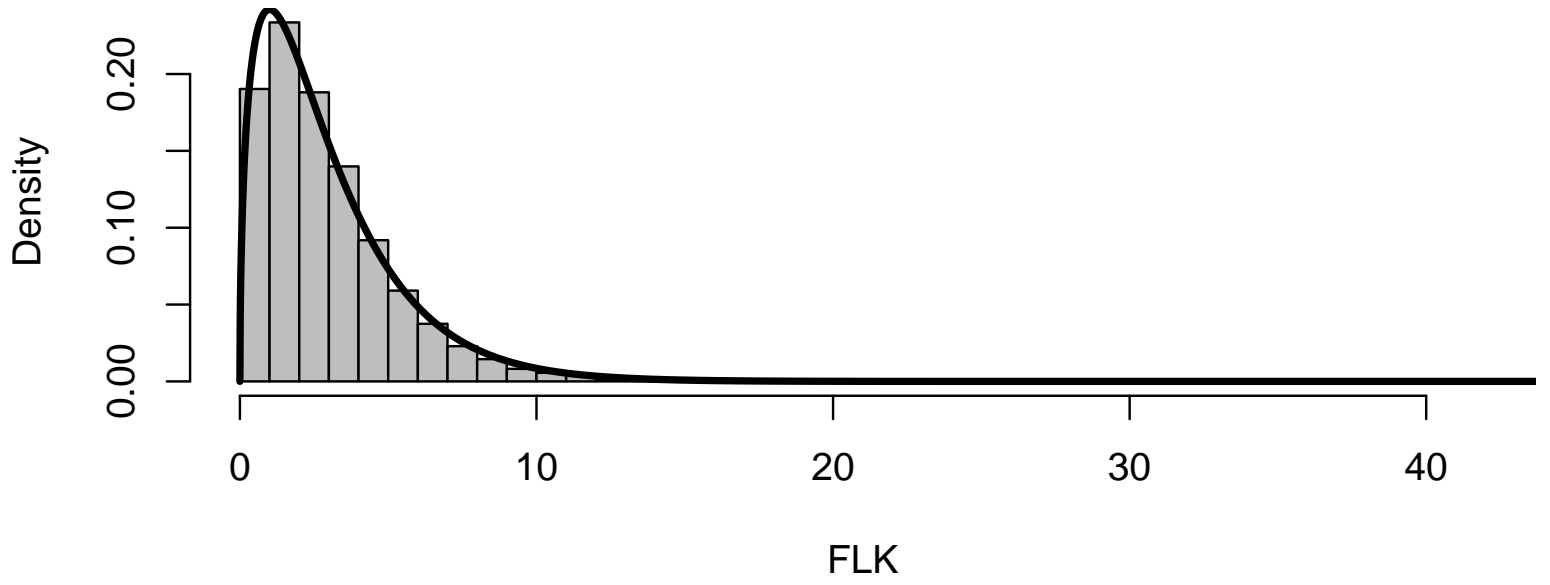
FLK genome scan for group CEU



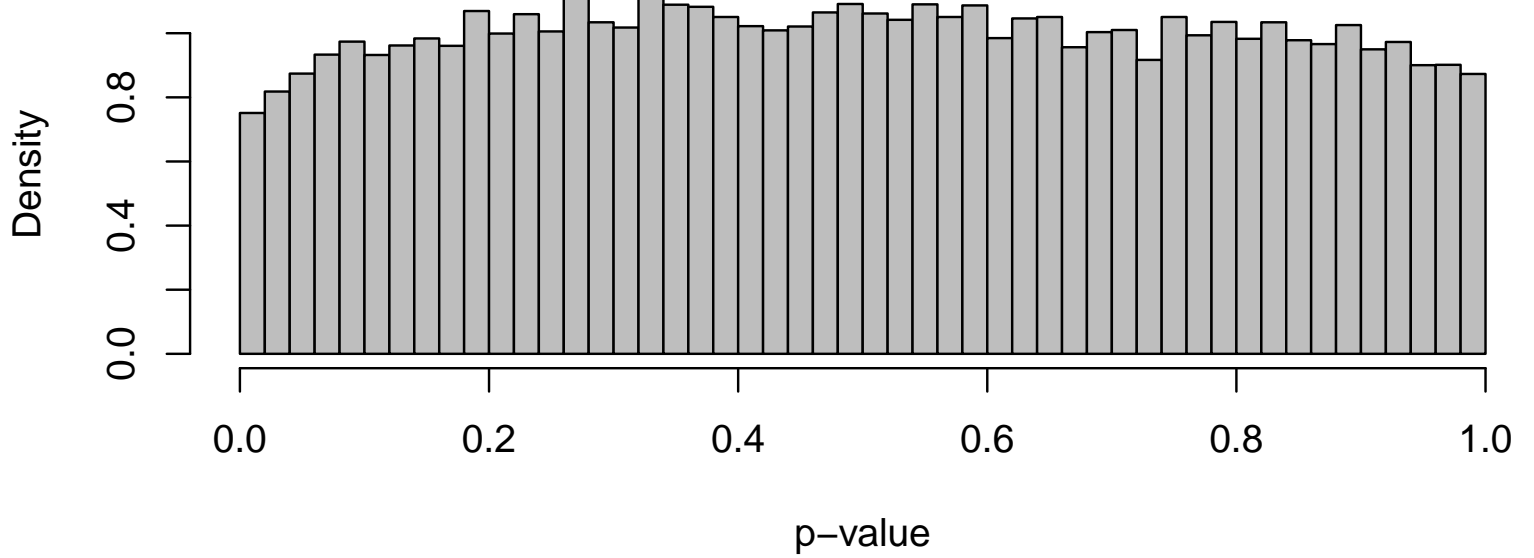
hapFLK genome scan for group CEU



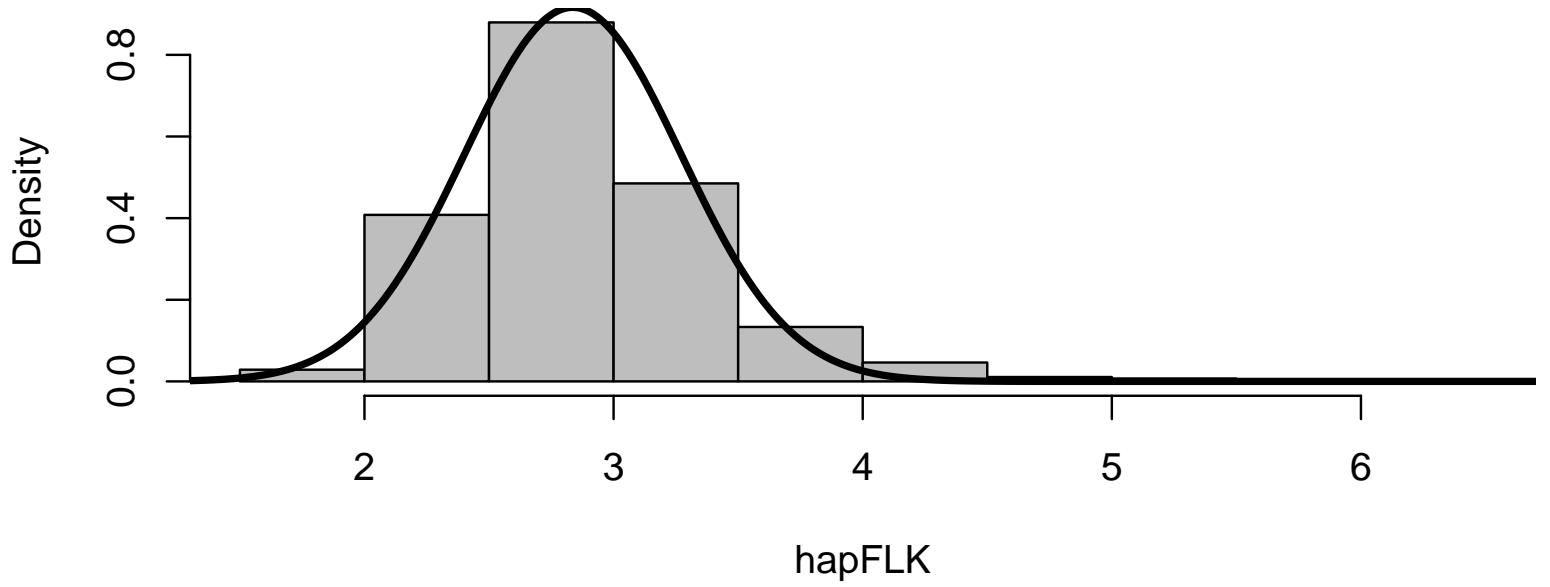
FLK distribution for group ITA



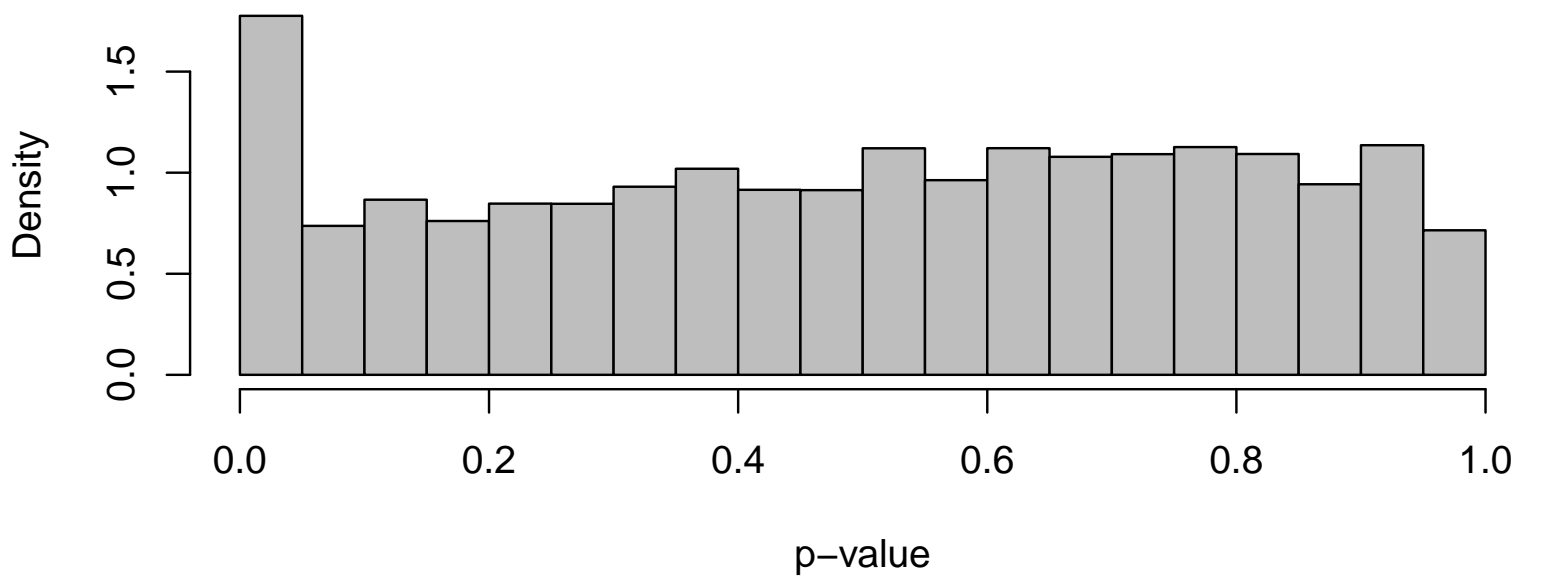
P-value distribution for group ITA



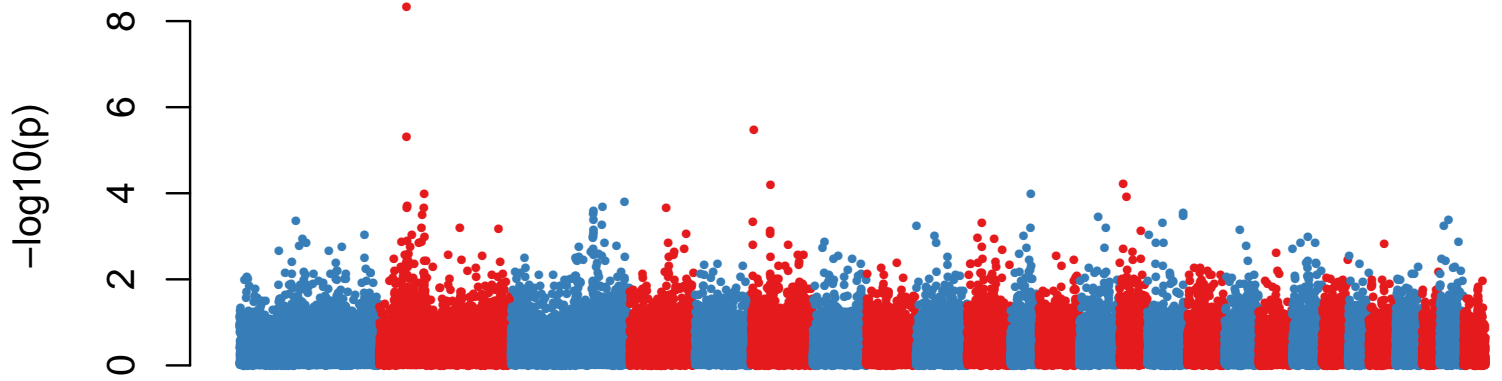
hapFLK distribution for group ITA



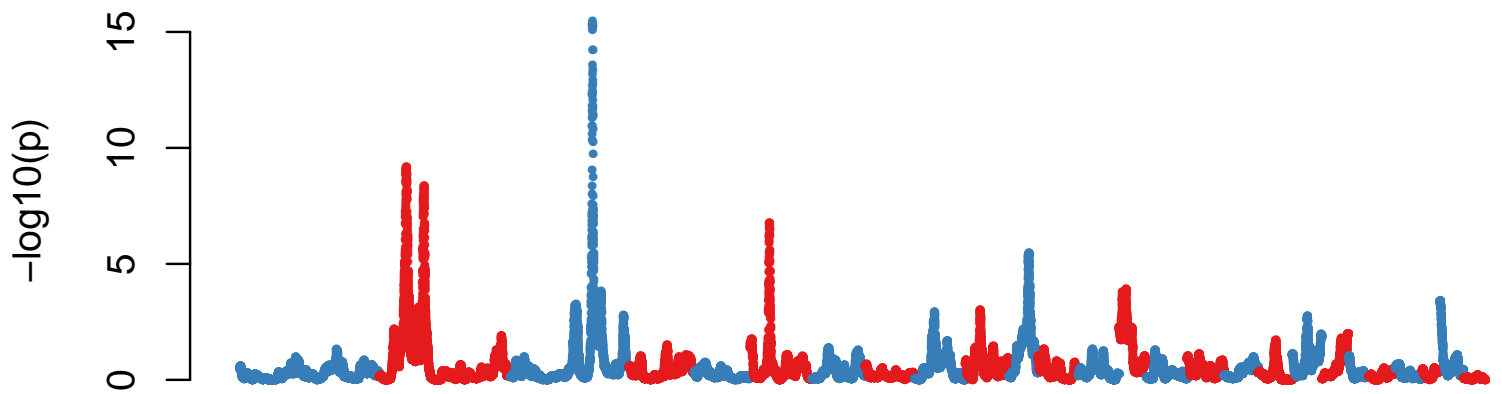
P-value distribution for group ITA



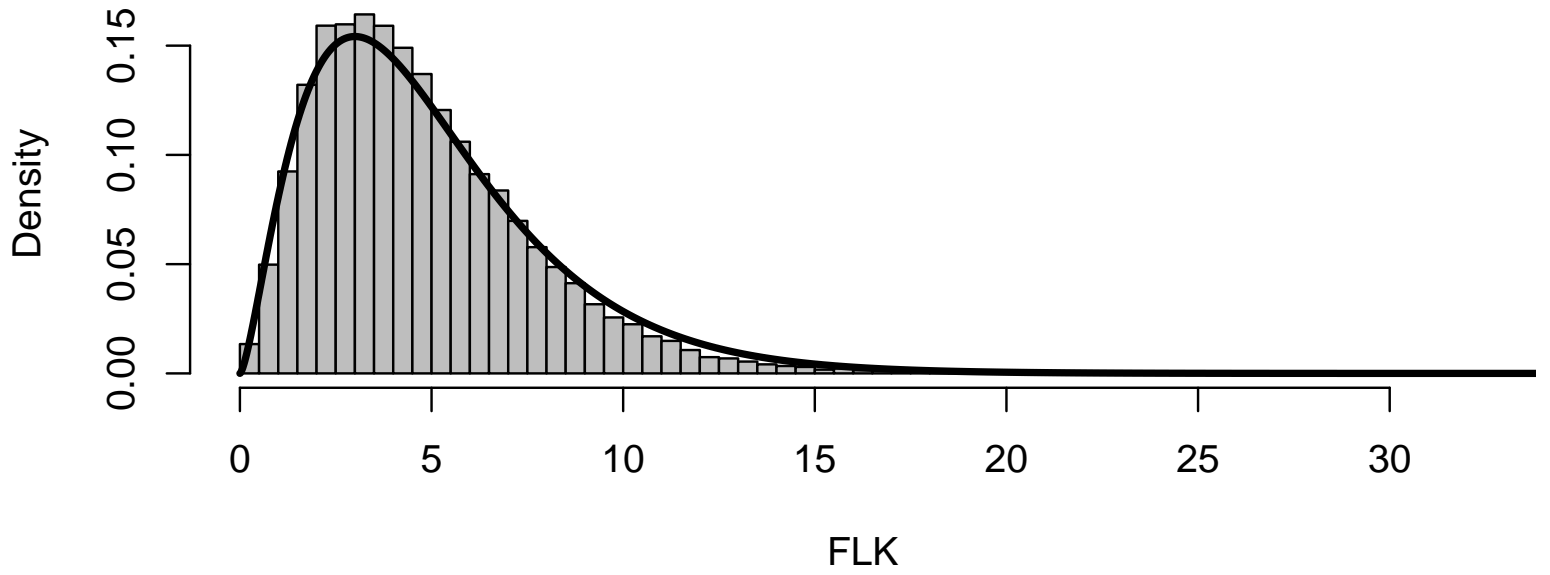
FLK genome scan for group ITA



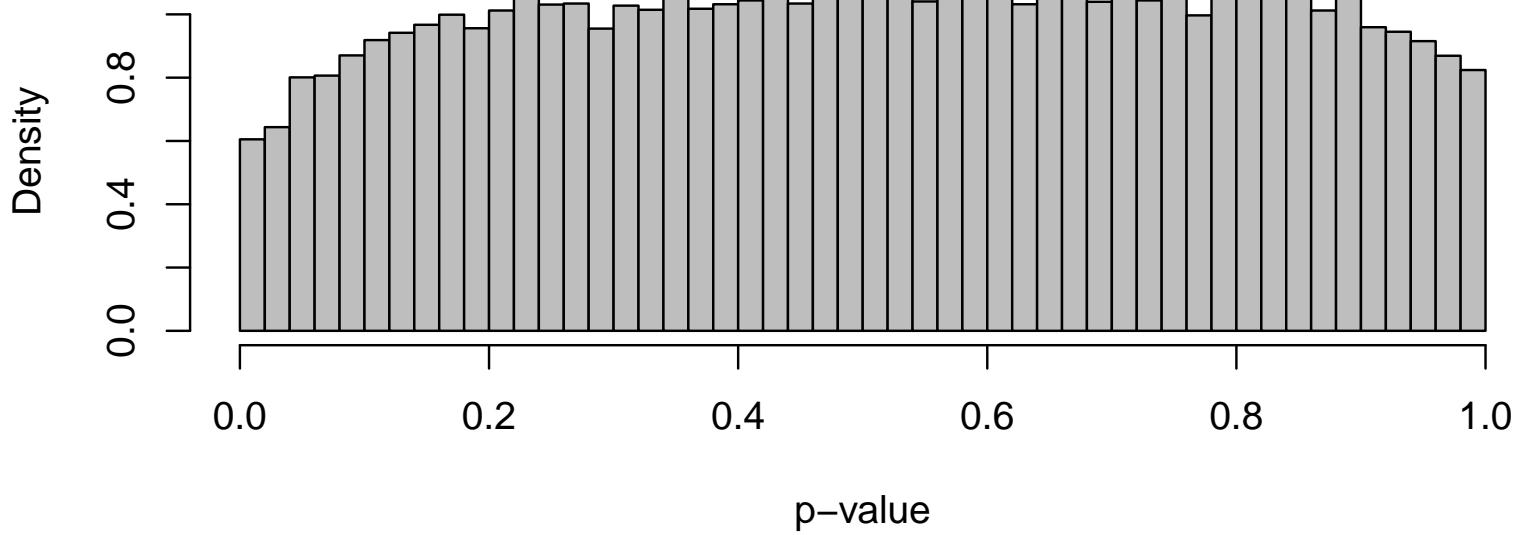
hapFLK genome scan for group ITA



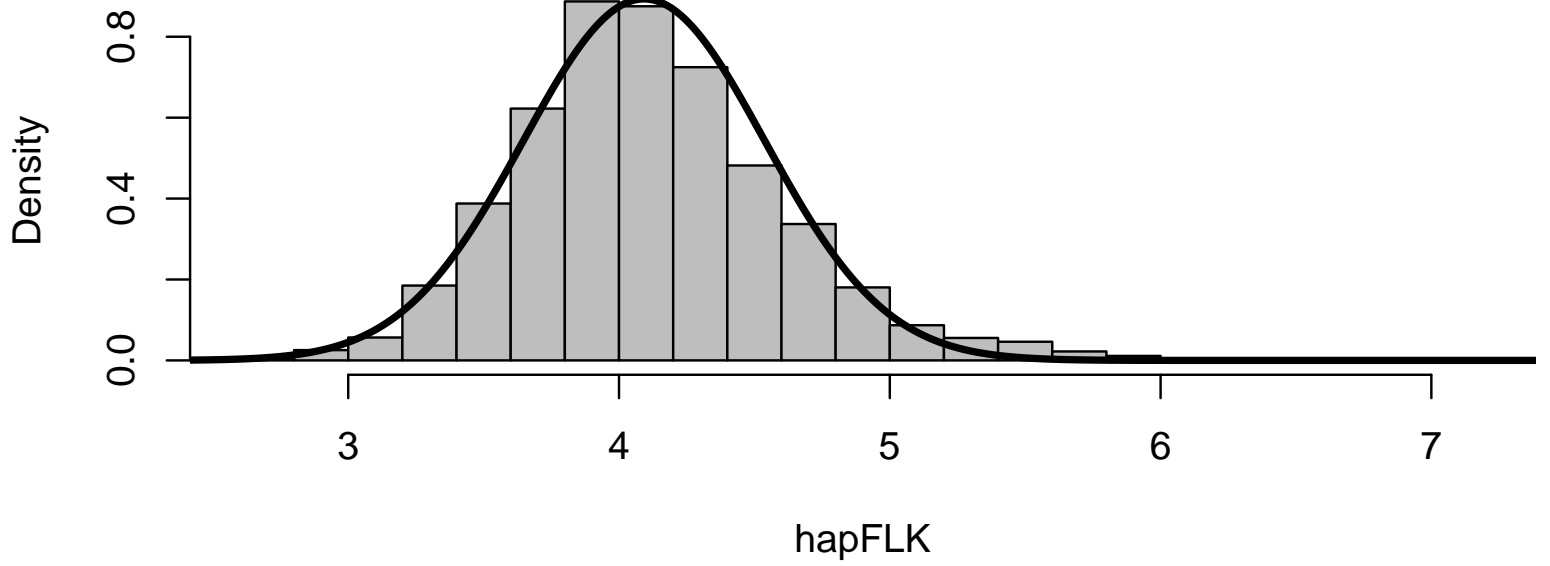
FLK distribution for group NEU



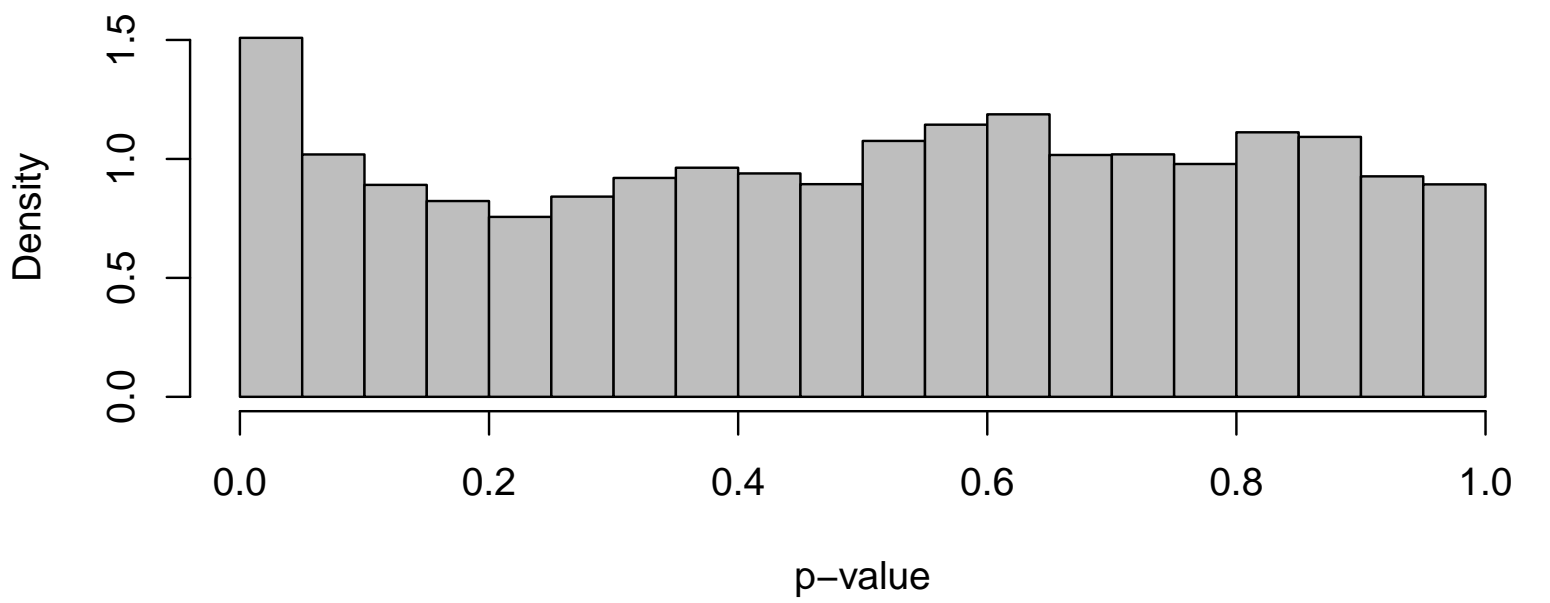
P-value distribution for group NEU



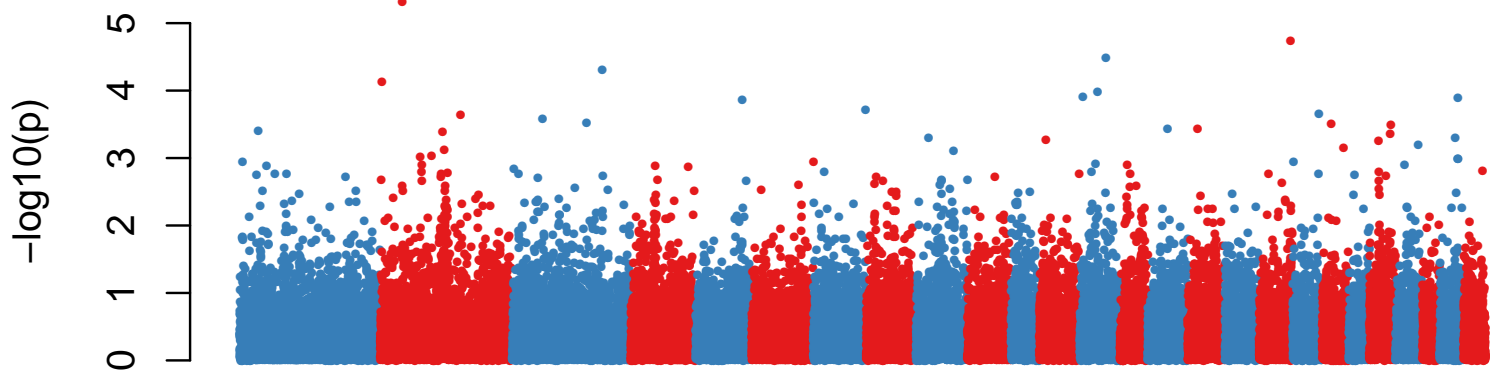
hapFLK distribution for group NEU



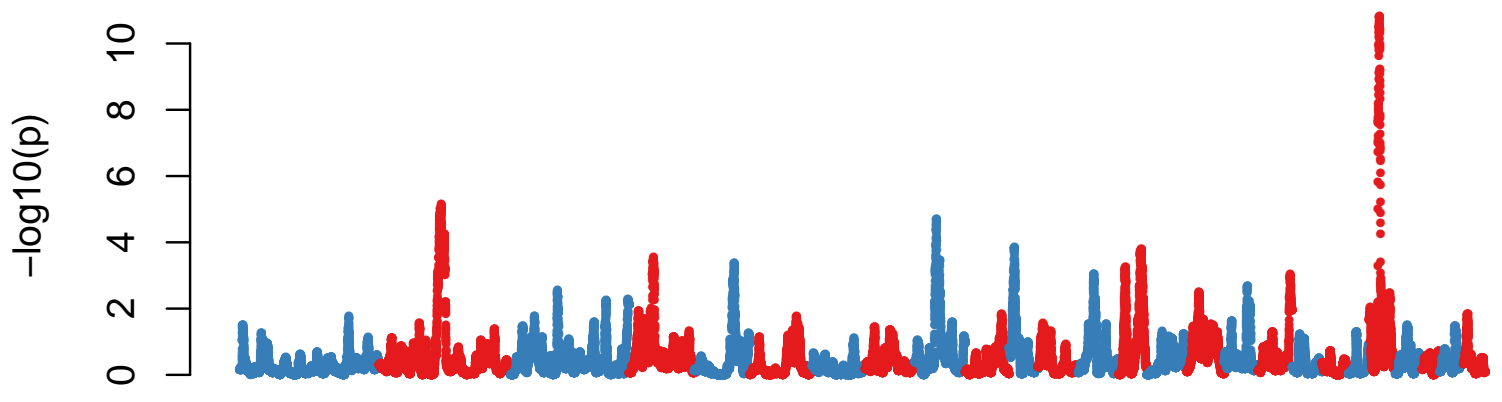
P-value distribution for group NEU



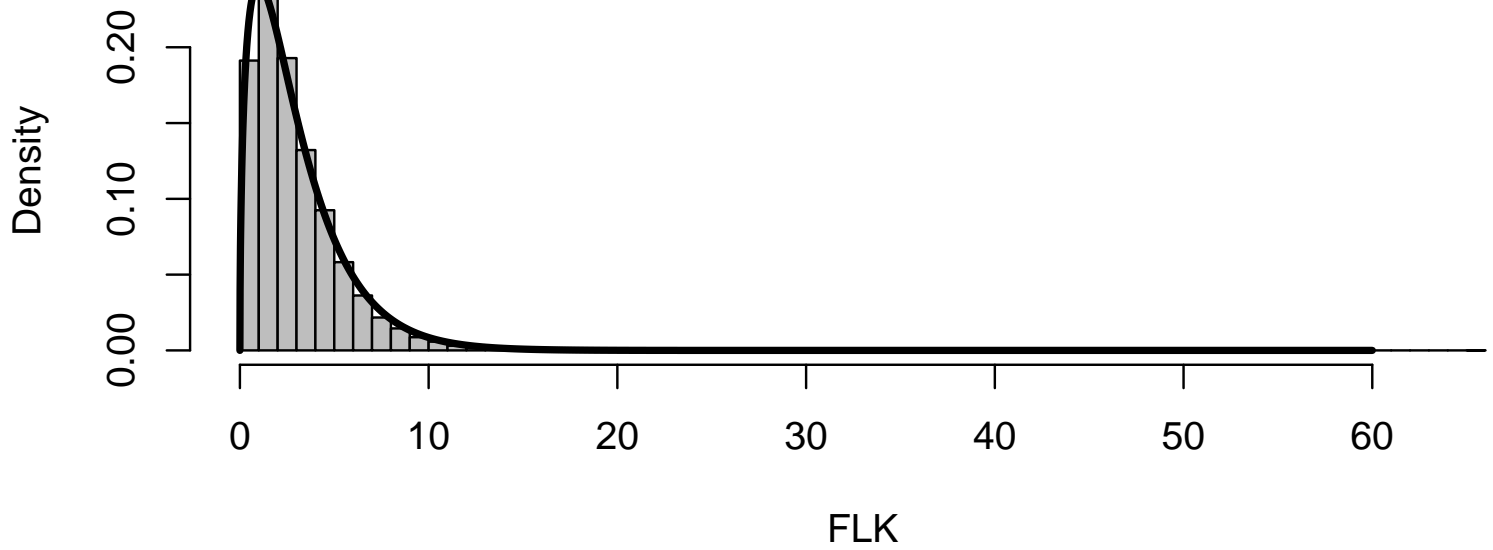
FLK genome scan for group NEU



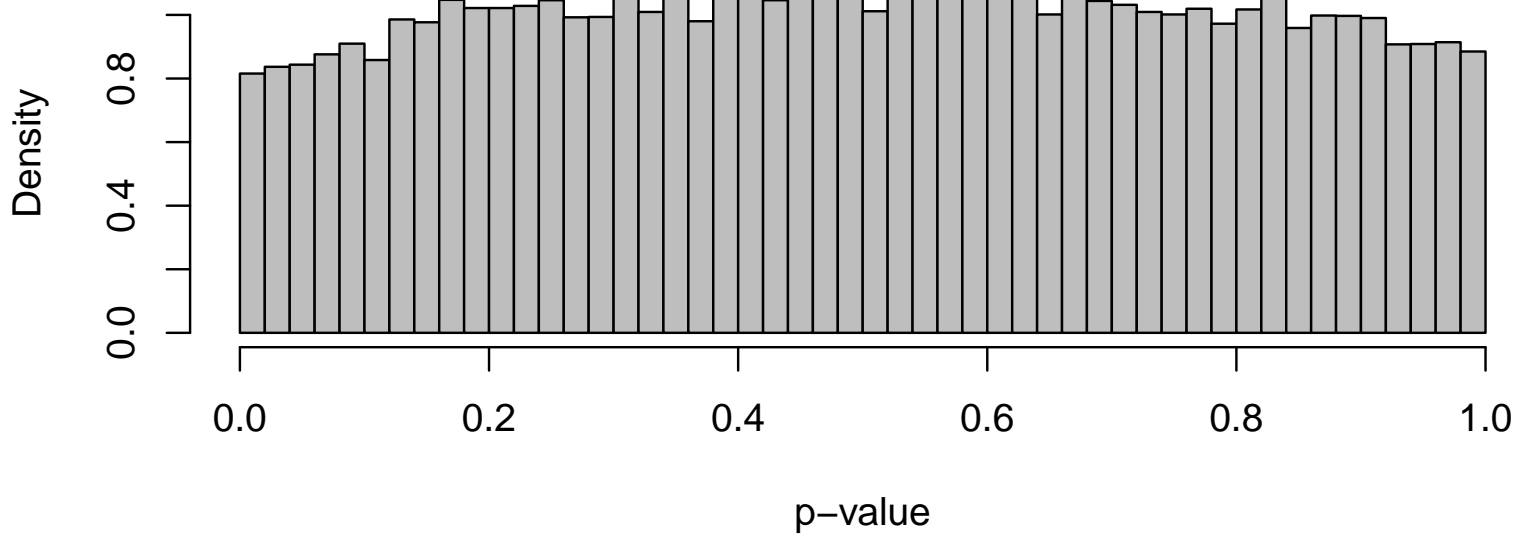
hapFLK genome scan for group NEU



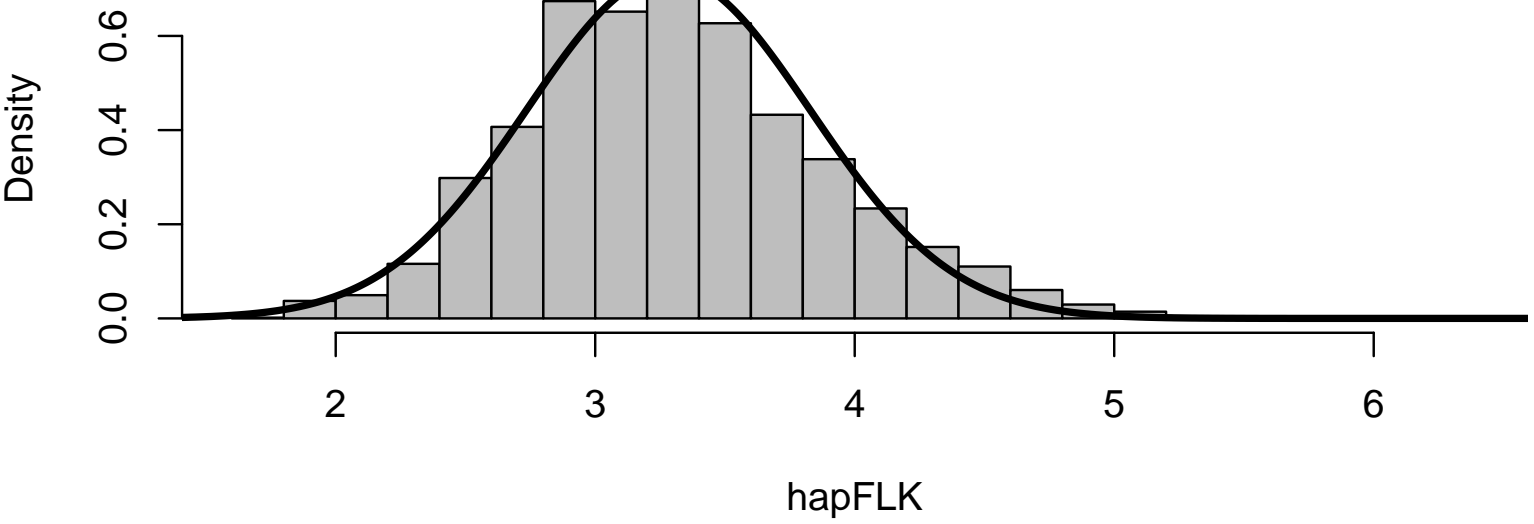
FLK distribution for group SWA



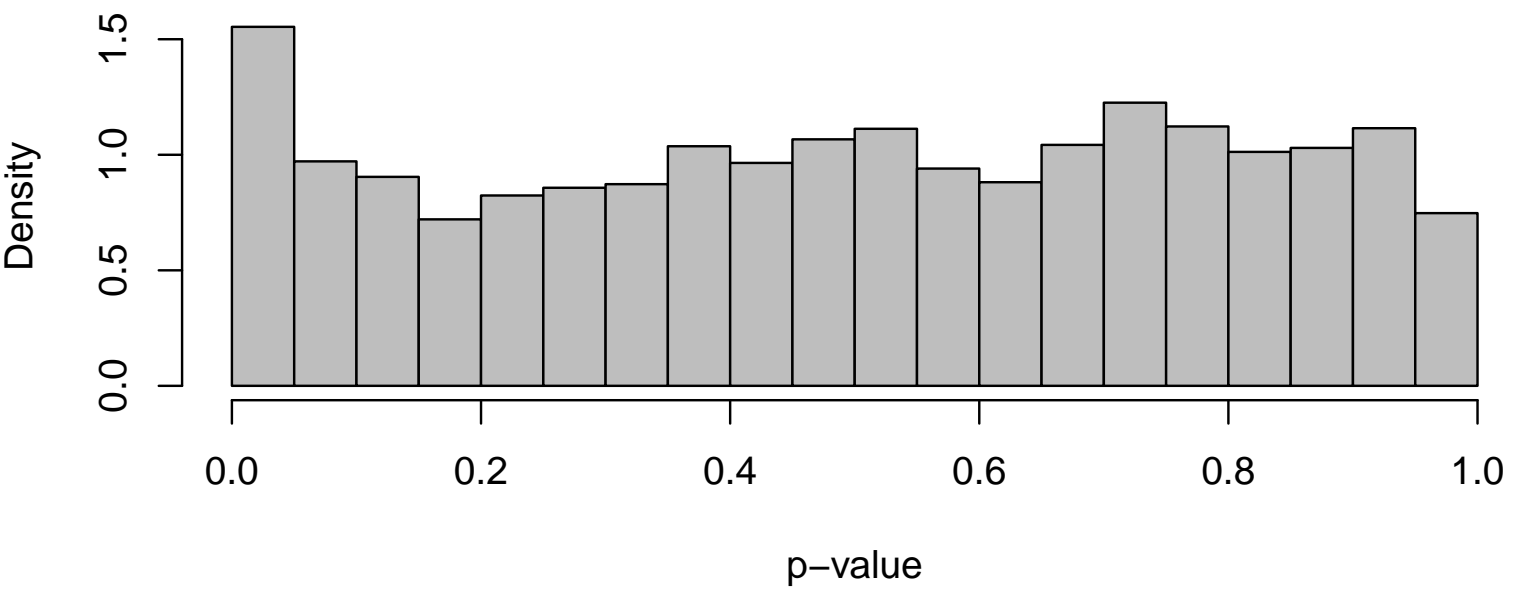
P-value distribution for group SWA



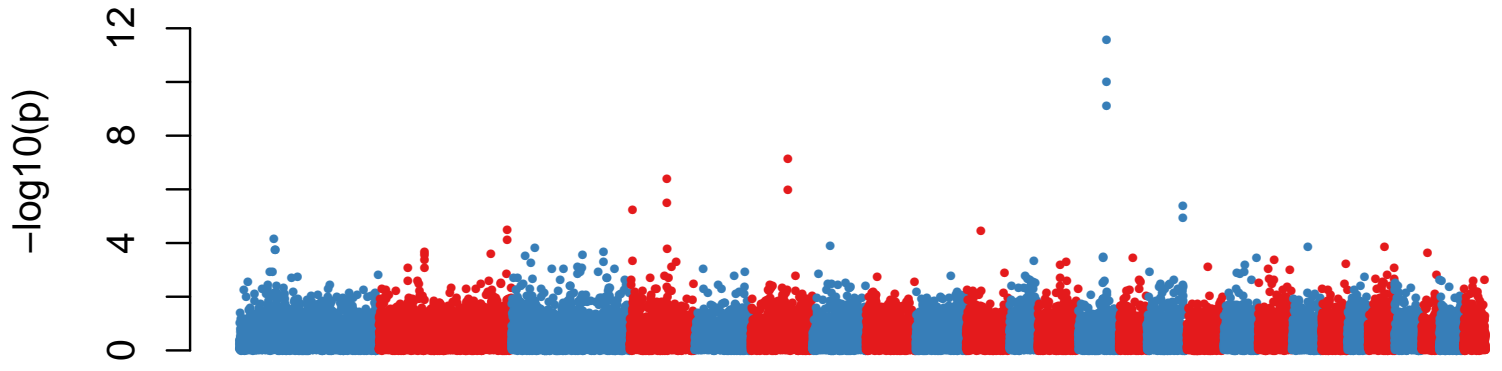
hapFLK distribution for group SWA



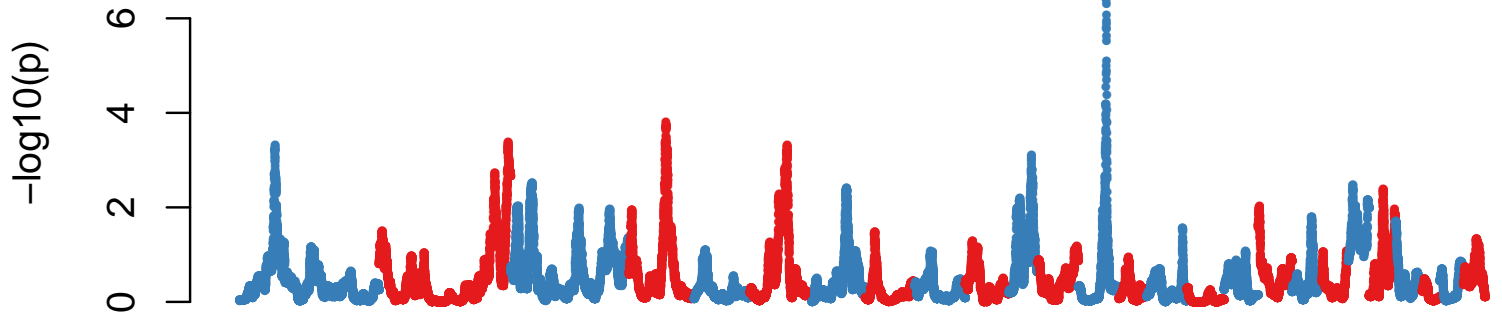
P-value distribution for group SWA



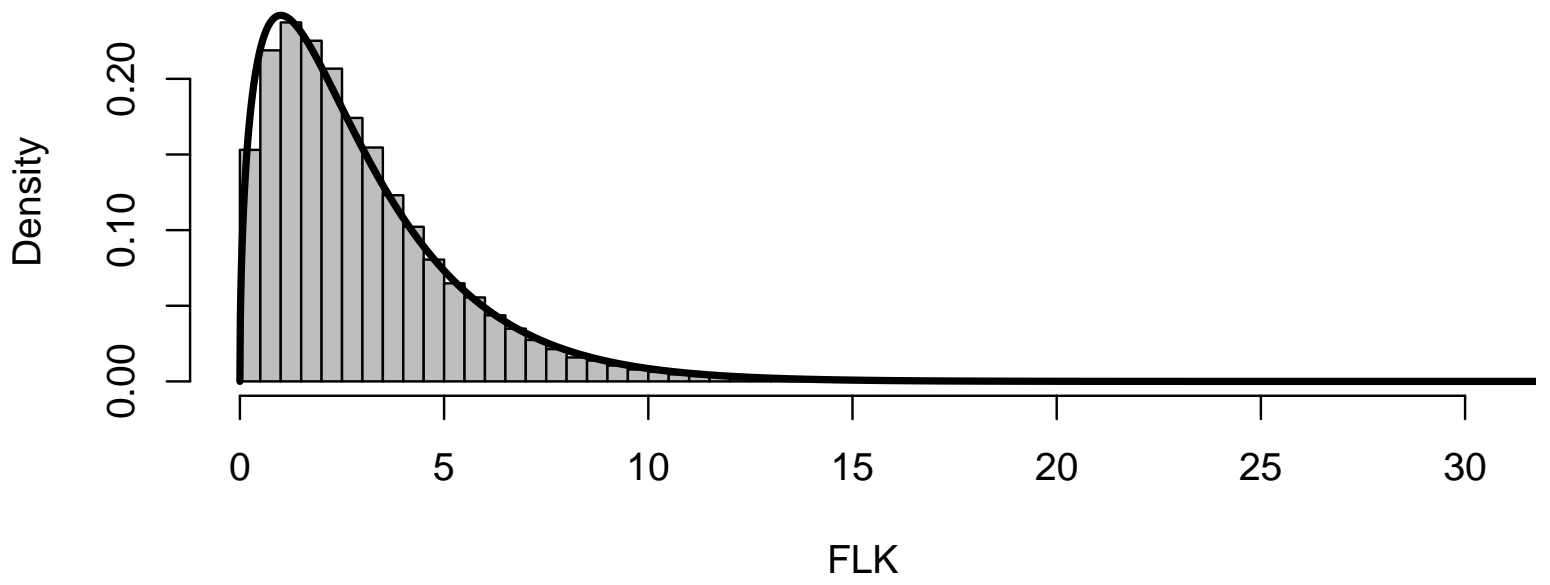
FLK genome scan for group SWA



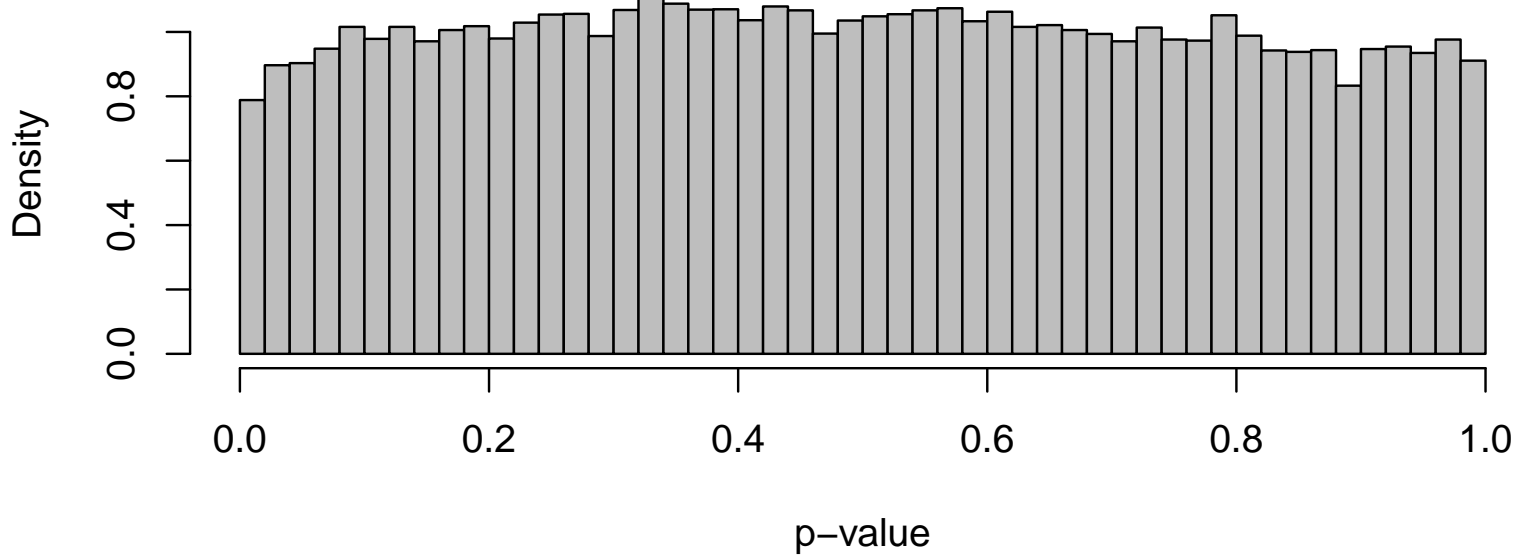
hapFLK genome scan for group SWA



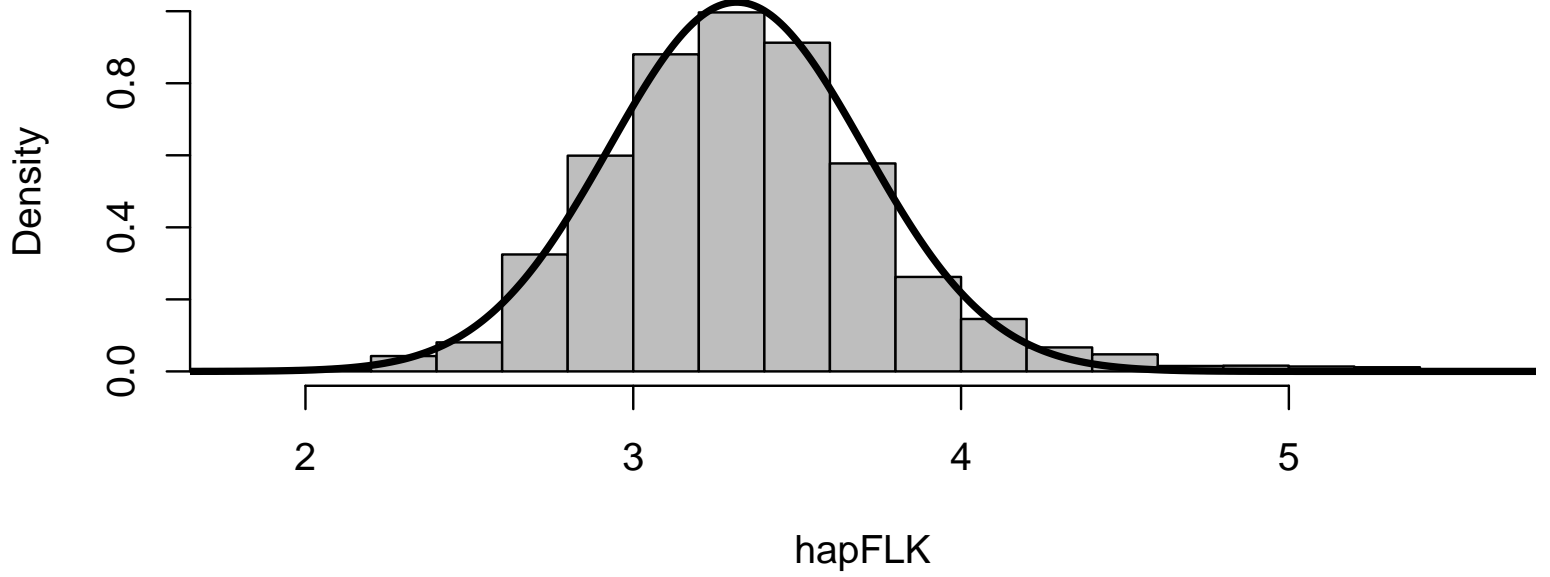
FLK distribution for group SWE



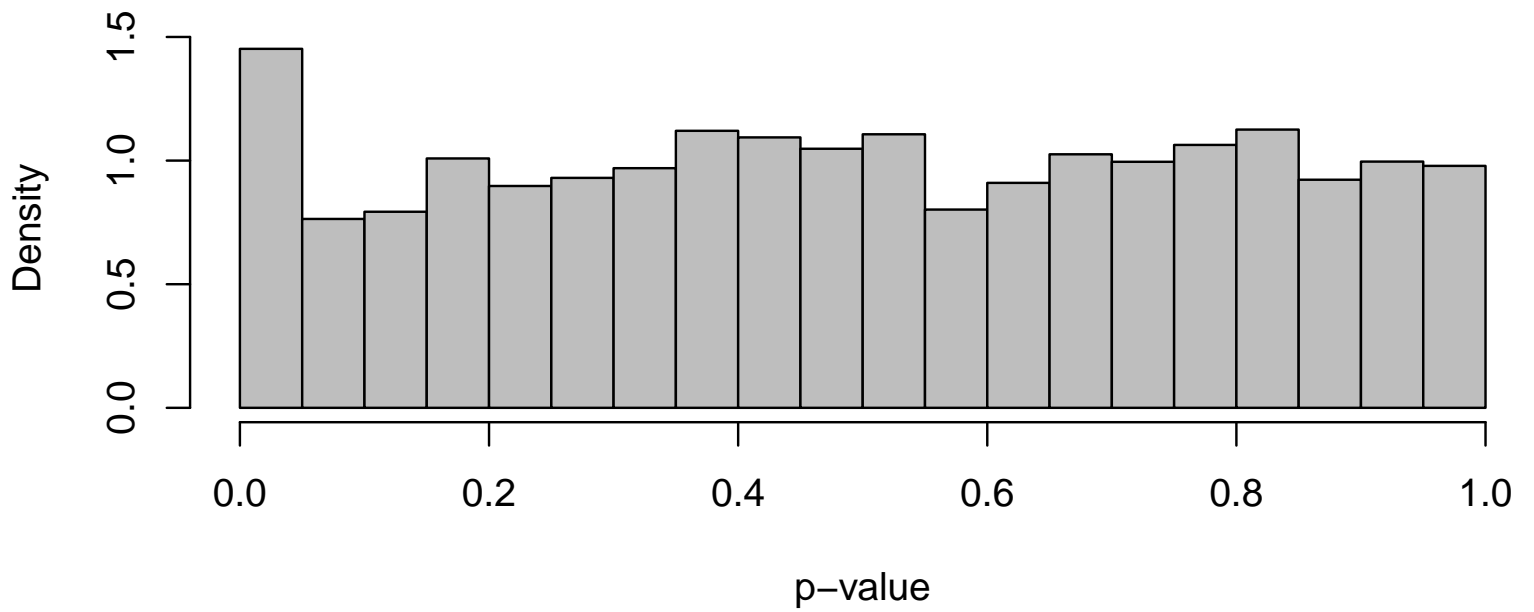
P-value distribution for group SWE



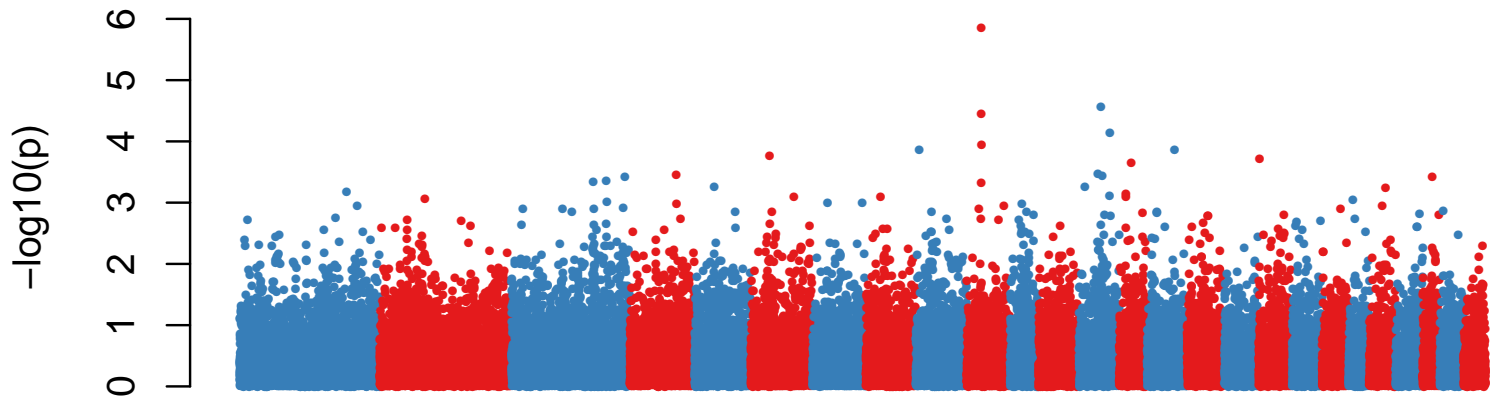
hapFLK distribution for group SWE



P-value distribution for group SWE



FLK genome scan for group SWE



hapFLK genome scan for group SWE

