Hunter-gatherer admixture facilitated natural selection in Neolithic European farmers

Tom Davy[1]*, Dan Ju[2], Iain Mathieson[2§], Pontus Skoglund[1*§]

[1]Ancient Genomics Laboratory, Francis Crick Institute, London, UK
[2]Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
*Correspondence to: tom.davy@crick.ac.uk, pontus.skoglund@crick.ac.uk
[§]co-senior authors

Abstract

Ancient DNA has revealed multiple episodes of admixture in human prehistory during geographic expansions associated with cultural innovations. One important example is the expansion of Neolithic agricultural groups out of the Near East into Europe, and their consequent admixture with Mesolithic hunter-gatherers. Ancient genomes from this period provide an opportunity to study the role of admixture in providing new genetic variation for selection to act upon, and also to identify genomic regions that resisted hunter-gatherer introgression and may thus contribute to agricultural adaptations. We used genome-wide DNA from 728 individuals spanning Mesolithic and Neolithic Europe to infer ancestry deviations in the genomes of admixed individuals, and to test for natural selection after admixture using a new method based on testing for deviations from a genome-wide null distribution. We find that the region around the pigmentation-associated gene *SLC24A5* shows the greatest overrepresentation of Neolithic ancestry in the genome ($|Z| = 3.45$). In contrast, we find the greatest overrepresentation of Mesolithic local ancestry across the key immunity locus that is the Major Histocompatibility Complex (MHC; $|Z| > 4$) which also shows allele frequency deviations indicative of a selective sweep following admixture ($p = 1 \times 10^{-29}$). This could reflect negative frequency dependent selection on MHC alleles common in Neolithic populations, or that Mesolithic alleles were positively selected for and facilitated adaptation by Neolithic populations to pathogens, new diets, or other environmental factors. Our results extend previous results that highlight immune function and pigmentation as targets of adaptation in more recent populations to selection processes in the Stone Age, and demonstrate that admixture facilitated selection by contributing new genetic variation.

Introduction

Despite the evidence from studies of ancient DNA that admixture among Holocene populations is ubiquitous, less is known about how admixture provided variation for natural selection to act upon during transitional periods. Given that Holocene admixture was often associated with dramatic migrations or changes in lifestyle, we might expect an important role for adaptive introgression. Perhaps the best-studied example of ancient admixture is in the Mesolithic-Neolithic transition in Europe. As

44 Early Neolithic groups expanded across Europe from Anatolia in the period 10,000 to
45 5,000 years ago[1–7], they admixed with local Mesolithic hunter-gatherers, and by the
46 later Middle Neolithic period derived 20-30% of their ancestry from these local
47 groups[1,2,8,9]. The admixed Neolithic ancestry thus found itself in a new demographic,
48 cultural, and geographic landscape. In addition to dramatic dietary changes, the
49 Neolithic transition included an increase in population density, which together with
50 the presence of domesticated animals has been hypothesised to lead to increased
51 infectious disease load[10].
52
53 Despite the potential of introgression as a source of adaptive genetic variation,
54 examples of adaptive admixture in humans are relatively rare. Examples include
55 American populations experiencing selection following admixture with European and
56 African populations at immune loci[11]. Within Africa, the *DARC* 'Duffy-null' allele has
57 seen at least two adaptive admixture events, spreading from the mainland to both
58 Madagascar and the Capo Verde islands[12,13], alongside signals at other immune
59 genes[14,15]. There are also several examples of adaptive alleles inherited from
60 archaic humans such as a *WARS2-TBX15* haplotype in Greenlandic Inuit[16] and an
61 allele of *EPAS1* in present-day Tibetans[17] both introgressed from Denisovans.
62 Studies of adaptive admixture among present-day populations are in general
63 confined to recent admixture in the past ten or so generations, which is too short to
64 detect anything but the strongest selection. On the other hand, signals of adaptive
65 archaic introgression from events thousands of generations ago may have decayed
66 with time and be difficult to detect.

67 Previous studies of natural selection in the European Neolithic have either compared
68 allele frequencies or haplotype structure with other ancient and modern
69 populations[2,18–21]. However, no study to date has attempted to assign signals of
70 adaptive admixture to a particular ancestry close to the time of admixture. The
71 admixture of Early Neolithic and hunter-gatherer ancestry represents an opportunity
72 to study adaptive admixture, similar to recently admixed present-day populations, but
73 over a much longer timescale; around a hundred generations (**Figure 1A**). A recent
74 study identified two optimal approaches to detect adaptive admixture, based on
75 allele frequencies and local ancestry, respectively[22]. Here, we adapt these two
76 approaches to ancient populations with a new framework to obtain p-values from
77 genome-wide null distributions, investigating adaptive admixture in 728 Mesolithic,
78 Early and Middle Neolithic individuals.

79 **Results and Discussion**
80 We used clustering analyses to assign individuals with genome-wide ancient DNA
81 data from Europe and Anatolia in the past ~15,000 years to one of three groups: HG
82 for 135 Mesolithic and Upper Palaeolithic hunter-gatherer ancestry individuals, NEO
83 for 56 early Neolithic individuals from Anatolia and the Balkans without evidence for
84 substantial hunter-gatherer admixture, and MNEO for 537 later Middle Neolithic
85 individuals with substantial HG admixture. In total, our analysis contains 728 ancient

2

86   individuals, spanning 7,500 years across the European continent (**Figure 1A-B;**
87   **Supplementary Figure 1; Supplementary Figure 2**).

88

89   We first used an approach to find natural selection that is admixture-unaware,
90   searching for increased differentiation as the squared allele frequency difference
91   between populations, the f_2 statistic[23]. We computed the average for each SNP and
92   25 SNPs flanking SNPs on each side (*i.e.* in 51 bp sliding windows), and obtain p-
93   values by fitting a gamma distribution to a null sample of 532 approximately
94   independent loci, separated by at least 5 million basepairs (mb). We observe no
95   statistically significant outliers in the HG-NEO or HG-MNEO contrasts
96   (**Supplementary Figure 3**), but in the NEO-MNEO contrast we observe a highly
97   significant excess differentiation across the Major Histocompatibility Complex (MHC)
98   region on chromosome 6, centred upon *HLA-DQB1* (p = 1×10^{-21}) (**Supplementary**
99   **Figure 3**). This suggests natural selection at the MHC in the period covered by the
100   data from NEO and MNEO groups, although this analysis does not address whether
101   that might be due to adaptive HG admixture.

102

103   To search for adaptive admixture, we applied a statistic based on previous work[2]
104   termed F_{adm}[22], which tests for deviations from the expected allele frequencies given
105   the genome-wide average mixture proportions of the contributing ancestries. We
106   again observe an excess signal across the MHC, centred upon *HLA-DQB1* (p =
107   1×10^{-29}) (**Figure 1A**). To confirm that our findings were not driven by ascertainment
108   bias in the 1.2M SNP panel[24], we analysed the MHC region in 156 whole-genome
109   shotgun sequences from 67 Mesolithic (HG), 27 Early Neolithic (NEO), and 62
110   Middle Neolithic Admixed (MNEO) individuals (Extended Methods), and observe a
111   concordant peak stretching across the Class III MHC region (**Figure 1E**).

112

113   We next sought to quantify the direction of admixture by searching for deviations in
114   local ancestry across the genome (Local Ancestry Deviation; LAD, ref. [22]). We
115   inferred local ancestry in 537 admixed Middle Neolithic individuals with genome-wide
116   SNP data using *ancestryHMM*[25], which estimates local ancestry in low-coverage
117   genomic data using allele frequencies from two populations. We computed standard
118   errors and Z-scores for local ancestry deviation (LAD) using an approximately
119   independent subsample of the genome-wide distribution consisting of 555 sites
120   separated by at least 5Mb (**Methods**).

121

122   The greatest excess of Neolithic ancestry centred on *SLC24A5* (**Figure 2C;**
123   **Supplementary Figure 4; Supplementary Figure 5**), with a peak of +18.6%
124   (|Z|=3.45). The derived *SLC24A5* allele, which is carried on the Neolithic ancestry
125   background, is one of the two alleles which contributes most to light skin
126   pigmentation in present-day European populations[26]. It has been previously shown
127   to have been at relatively high frequency in the Neolithic and absent in the Mesolithic
128   hunter-gatherers[2], and our results show that the selection removed hunter-gatherer
129   ancestry at this locus in later admixed Neolithic groups.

130

131   Meanwhile, the lowest amount of Neolithic ancestry is found at the Major

132   Histocompatibility Complex (MHC) region on chromosome 6. Within this locus, the

133   region of highest Mesolithic ancestry is centred on *HLA-E*, with a peak excess of

134   +23.2% (|Z|=4.34). This region of elevated Mesolithic ancestry continues as a

135   contiguous region which extends across the Class II region of the MHC, with an

136   average of ancestry across the entire MHC (between hg19 positions 28,477,797-

137   33,448,354 of chromosome 6) of +9.16% (|Z|=1.82), and a secondary peak centred

138   upon the class II region of +17.26% (|Z|=3.23) (**Figure 1E**).

139

140   In the F_{adm} analysis, we also find a number of genome-significant peaks outside of

141   the MHC, though these are an order of magnitude below the MHC p-value

142   (**Supplementary Figure 6**). *TYRP1* has been previously reported to have

143   experienced selection in European populations[27,28]. *TLR2* is an important gene in

144   immunity, and may also be a target of archaic adaptive admixture[29]. *MUC19* also

145   plays a role in immunity, and was found to be under selection in Central Mexican

146   indigenious populations[30]. *OPRM1* is documented to contribute towards skin

147   pigmentation in Native American and East Asian populations[31,32], while a *HRNR*

148   haplotype linked to atopic dermatitis may have seen a selective sweep in East Asia

149   alone[33] *ROR1* is adjacent to *PGM1* which may have undergone selection in Mexican

150   populations[34]. Similarly, the LAD analysis finds some genes with previous reports of

151   selection; *HECW2* has been identified to be under selection in Khomani San[35], a

152   southern African group with recent hunter-gatherer lifestyle, while *PKP2* was

153   similarly seen to be under selection in the Yoruba in west Africa in the same

154   study.

155

156   It is also possible that adaptive admixture acted on multiple variants with small effect,

157   spread across the genome. To test for evidence of such polygenic selection, we

158   computed the Pearson correlation between the local ancestry deviation and the

159   effect size for 38 traits in the UK BioBank, using genome-wide significant SNPs

160   thinned to be approximately independent[36] (**Figure 3**). We see significant evidence

161   of correlation between trait scores and LAD in Skin Colour (p = 3e^{-4}), consistent with

162   the adaptive admixture around *SLC24A5*. Indeed, this signal is solely driven by two

163   loci, with a *HERC2* variant with a skew towards the Mesolithic (Z=1.7) also

164   contributing to a lighter level of skin pigmentation alongside *SLC24A5*. Without these

165   two loci, there is no significant evidence of polygenic selection (P = 0.58). We also

166   observe a weaker but significant correlation for hip size (**Figure 3, Supplementary**

167   **Figure 7**).

168

169   The Neolithic transition brought about drastic changes in demography, culture and

170   diet, as well exposure to novel pathogens and increased potential of zoonotic

171   disease. In admixed middle Neolithic individuals, we found excess Neolithic farmer

172   ancestry at the pigmentation locus *SLC24A5* and excess Mesolithic hunter-gatherer

173   ancestry at the MHC immunity locus. Previous studies also found evidence of natural

174   selection at *SLC24A5* in European populations[26,27] and showed that the allele was
175   introduced into Europe in the Neolithic[2,37,38] but our study now further demonstrates
176   that this resulted in a removal of hunter-gatherer ancestry across the wider locus. In
177   a similar but opposite process, the MHC locus has previously been demonstrated to
178   have undergone selection in the ancestry of present-day Europe[239] and specifically
179   in Neolithic Europe[18]. Here, we obtain further robust results for selection at the MHC
180   locus corrected for multiple testing, and demonstrate that this process specifically
181   increased hunter-gatherer ancestry at the locus.

182

183   In contrast to *SLC24A5*, the second high-effect pigmentation variant, in *HERC*,
184   displays an excess of Mesolithic ancestry (+17.23%, $|Z| = \sim3.11$). Together with the
185   third high-effect pigmentation variant at *SLC45A2*, which arrived in Europe via later
186   expansions from the steppe, selection on pigmentation in Europe thus targeted
187   variants from each of the three major ancestral populations[9]. This highlights the
188   prominent role of admixture in the evolution of skin pigmentation in Western Eurasia.
189   That this signal is not found in the allele-frequency based analysis with F_{adm} can
190   likely be attributed to the small absolute change in allele frequency between our
191   Neolithic populations, confirming recent demonstrations that local ancestry can in
192   some cases be more powerful than allele frequency analysis for detecting selection
193   in admixed popuations[22].

194

195   Evidence of selection on Mesolithic ancestry across the MHC locus, highlights its
196   role in facilitating adaptation in immunity during the Neolithic transition in Europe.
197   One hypothesis is that this reflects the fact that Neolithic populations were expanding
198   into environments containing pathogens to which Mesolithic populations had already
199   adapted. This is contrary to the idea that the pathogen load in Neolithic populations
200   was solely driven by increased population density and proximity to zoonotic vectors
201   via animal husbandry. While examples of putative adaptive admixture involving the
202   MHC have been previously described[11,15,40], a clear link between the alleles under
203   selection within this region and a specific pathogen has not been previously
204   observed.

205

206   Another possibility is that this adaptation reflects negative frequency-dependent
207   selection[41], for example because pathogens adapt to the host population's immunity
208   genes, with a bias towards more common variants. Rare alleles would thus confer a
209   fitness advantage, until they become common enough that other, now-rarer, alleles
210   have higher fitness. Under this model, HLA alleles unseen by a given pathogen will
211   attain a higher fitness initially following admixture owing to their novelty for the
212   pathogen[42]. Thus, the selection on the minor Mesolithic ancestry could simply reflect
213   admixture introducing rarer variants into Neolithic populations. Future studies,
214   including whole-genome shotgun data in tandem with improved functional
215   annotation, may shed further light on this adaptive process.

216

217

230 FIGURES
231

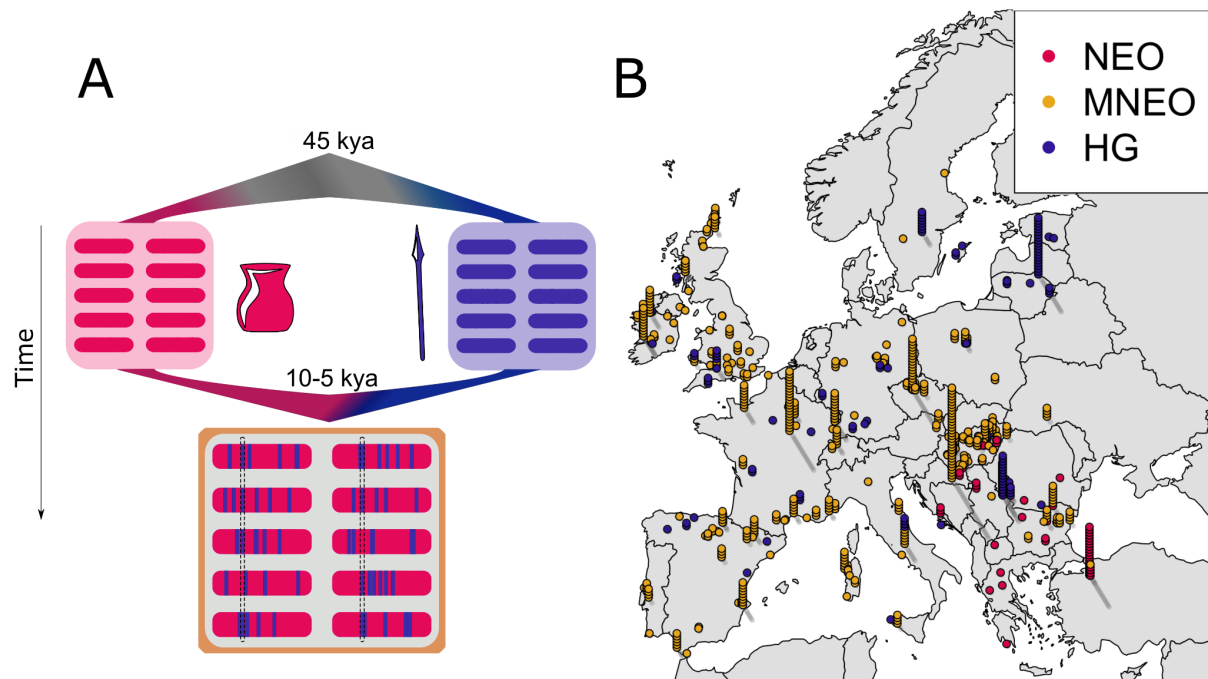A

45 kya

Time

10-5 kya

B

NEO
MNEO
HG

232
233
234

Figure 1. Admixture model and geographic distribution of Neolithic and Mesolithic individuals with genome-wide ancient DNA. A) An illustration of the genetic history of the Neolithic-Mesolithic transition in Western Eurasia. **B)** 'Casino-plot' of individuals included for analyses, coloured by the ancestry group for which those individuals were used in this paper. For sites with multiple samples, we stack those individuals above the reported coordinates.
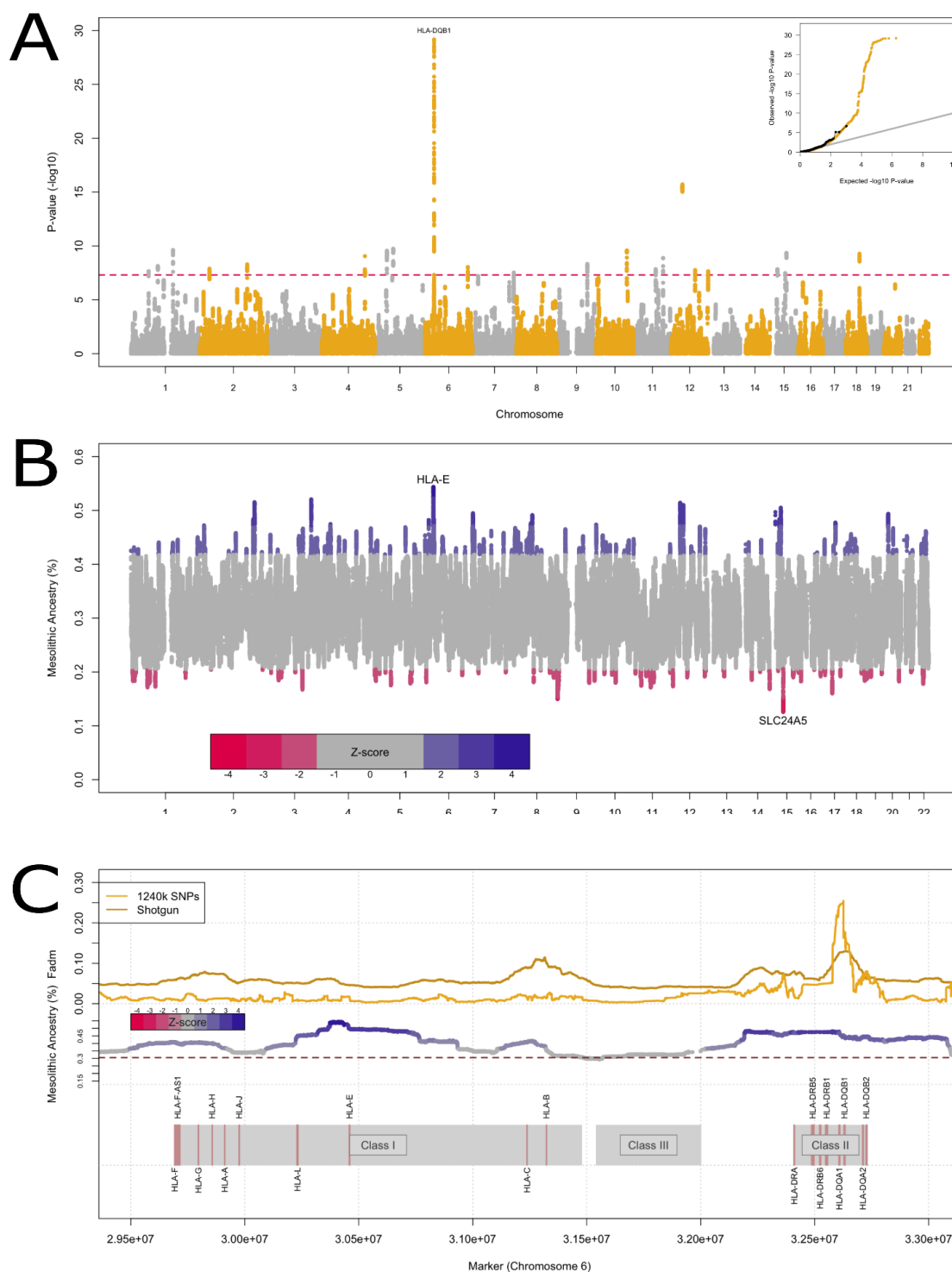
241

7

Figure 2. Genome-wide significant signals of adaptive admixture. A) Manhattan plot of p-values from the F_{adm} scan across the genome for deviations from expected admixed allele frequencies. Inset, quantile-quantile plot of expected and observed p-values. **B)** Local Ancestry Deviations (LAD) in the Middle-Neolithic across the genome, with top peaks of each ancestry labelled. **C)** Zoomed-in region of the MHC (chromosome 6), with statistics derived from 1240k and whole-genome shotgun data across the MHC regions I,II & III on chromosome 6.
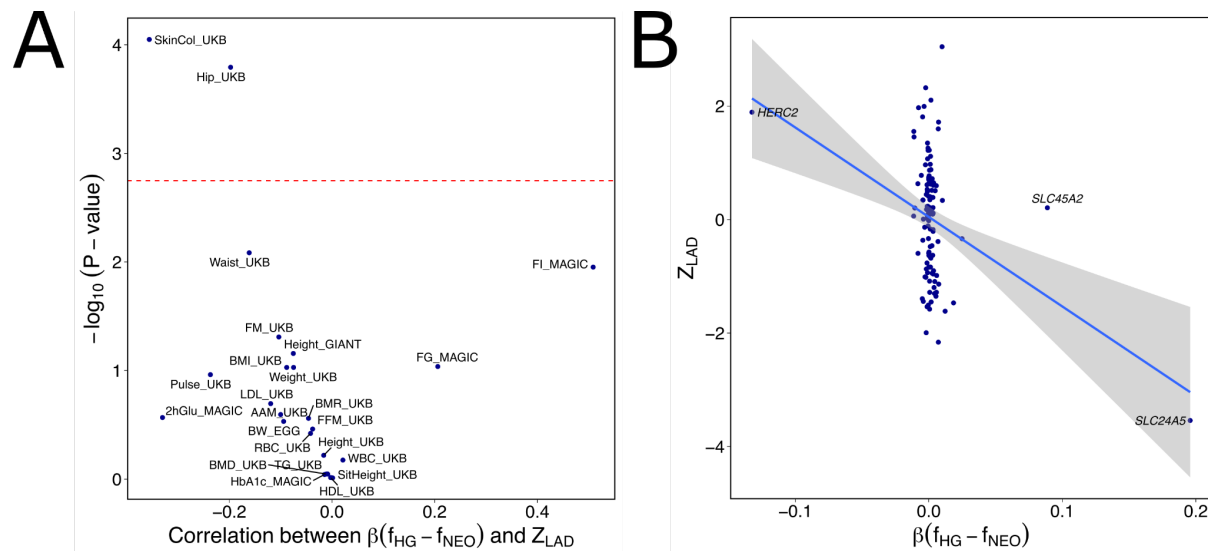
Figure 3. Test for polygenic adaptive admixture. A) Pearson correlation of polygenic traits against local ancestry. B) Correlation of LAD Z-scores with Skin Colour SNP effect size weighted by the signed allele frequency difference between the two source populations.

Experimental Procedures

Data preparation

We first used clustering approaches on a large set of previously published[1–9,37,38,43–66] individuals to identify those individuals exclusively or almost exclusively HG (Mesolithic, Western or Siberian Hunter-Gatherers [WHG, SHG]) or NEO ancestry (Early Neolithic, which we identify using Neolithic Anatolia as a baseline). We downloaded v50 of 10,391 individuals (3,589 ancient) from the Reich lab https://reich.hms.harvard.edu/downloadable-genotypes-present-day-and-ancient-dna-data-compiled-published-papers, accessed July 19th 2022 for the 1240K set of SNPs. We ran a first set of clustering with ADMIXTURE from K=3 to K=11, conditioned on European individuals dated to >5kya excluding siblings, parents, and duplicates. We then further filtered on age, setting a max age for HG, NEO and MNEO of 12kya, 8.5kya and 8kya respectively, retaining 728 individuals. Using PLINK, we then make a subset of files containing only individuals fitting into either Mesolithic (EHG+WHG) or admixed Neolithic ancestry, alongside a geographically and temporally conservative Neolithic cluster.

Detecting selective sweeps

We filtered the full set of SNPs by conditioning on observing at least 10 non-missing pseudohaploid genotypes in earch population. We utilised the F_{adm} statistic , i.e. (F_{MNEO} - $E(F_{MNEO}))^2$ / (1 - $E(F_{MNEO})^2$), where $E(F_{MNEO})$, where F is the frequency of a given SNP and $E(F_{MNEO})$ is an expectation of allele frequency derived by the contributing ancestry proportions weighted by allele frequency ($F_{MESO}Anc_{MESO}$ + $F_{NEO}Anc_{NEO}$). In all cases, the given statistic is calculated on a per-SNP basis before sliding-window is applied across the genome with a step size of 1. We derive the f_2 analyses in an identical manner. We annotate genes using the gencode v39liftover37 annotation files.

For each analysis, we then draw a null distribution from this sliding-window whole-genome distribution, ensuring that each sliding-window datum that contributes to the null distribution is sampled with SNPs that are, as a group, no less than 5Mb away from the previous sample. A gamma distribution is then fitted to this null distribution using the R package *fitdistrplus v1.1-3* [67] with the flags ""gamma", method = "mle", keepdata = T". We then derive p-values for the genome-wide distribution from this null-fitted gamma distribution..

Confirmation of HLA signal in whole-genome shotgun data

To confirm the presence of the HLA signal in the whole-genome shotgun data, we collected a set of 81 individuals used in the original analyses which had whole-genome shotgun genomes, and added 74 unpublished genomes available pre-publication under the Ft. Lauderdale principles (https://reich.hms.harvard.edu/ancient-genome-diversity-project). We thus only use this data for local analyses of the MHC region. Exclusion of the pre-publication data results in too few NEO source genomes for reliable inference, and so we do not report genome-wide p-values. We again partition out our target populations using appropriate populations within this dataset. We filter the region initially on sites with more than 3 allele counts seen within the whole dataset, and a minimum of 30 sites seen in the Neolithic population, resulting in 11,086,998 sites after filtering. We apply a sliding window of 593 snps on the mixed-weight statistic in a similar way to the 1240k analysis, using the same weighting for ancestry contribution to SNP frequency expectation. We choose this window

306 size to approximate the same mean genetic distance per window as in the 1240k analysis
307 with 51 snp windows.
308

Detection of Local Ancestry Outliers
310 To analyse biases in local ancestry, we progress the same panel of 728 individuals
311 (Supplementary Table 1) to analysis with *AncestryHMM* [25] . We derive frequency counts for
312 both the contributing ancestries (HG, NEO) and similarly obtain this for each individual
313 MNEO sample via plink. We then run Ancestry HMM with a prior of 30% HG and 70%
314 Neolithic ancestry, and an N_e of 10,000.
315

Detection for biased inheritance of polygenic trait alleles
317 As in ref . [36], we start with 28 quantitative traits of interest, subsetting SNPs $P < 1e^{-8}$
318 overlapping the 1240k SNP array. We then iteratively prune through subsetting the smallest
319 p-value and removing all other associations within 250kb. Departing from Mathieson &
320 Terhost, we weight these pruned values by the difference in allele frequency between HG
321 and NEO and sum these across the genome for each trait. and look for correlation to the
322 LAD derived in the local ancestry analysis via a Pearson correlation test.
323

References
325

326 1. Haak, W., Lazaridis, I., Patterson, N., Rohland, N., Mallick, S., Llamas, B., Brandt, G.,
327   Nordenfelt, S., Harney, E., Stewardson, K., et al. (2015). Massive migration from the
328   steppe was a source for Indo-European languages in Europe. Nature *522*, 207.

329 2. Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S.A.,
330   Harney, E., Stewardson, K., Fernandes, D., Novak, M., et al. (2015). Genome-wide
331   patterns of selection in 230 ancient Eurasians. Nature *528*, 499–503.

332 3. Hofmanová, Z., Kreutzer, S., Hellenthal, G., Sell, C., Diekmann, Y., Díez-del-Molino, D.,
333   van Dorp, L., López, S., Kousathanas, A., Link, V., et al. (2016). Early farmers from
334   across Europe directly descended from Neolithic Aegeans. Proceedings of the National
335   Academy of Sciences *113*, 6886–6891.

336 4. Cassidy, L.M., Martiniano, R., Murphy, E.M., Teasdale, M.D., Mallory, J., Hartwell, B.,
337   and Bradley, D.G. (2016). Neolithic and Bronze Age migration to Ireland and
338   establishment of the insular Atlantic genome. Proceedings of the National Academy of
339   Sciences *113*, 368–373.

340 5. Kılınç, G.M., Koptekin, D., Atakuman, Ç., Sümer, A.P., Dönertaş, H.M., Yaka, R., Bilgin,
341   C.C., Büyükkarakaya, A.M., Baird, D., Altınışık, E., et al. (2017). Archaeogenomic
342   analysis of the first steps of Neolithization in Anatolia and the Aegean. Proc. Biol. Sci.
343   *284*.

344 6. Omrak, A., Günther, T., Valdiosera, C., Svensson, E.M., Malmström, H., Kiesewetter,
345   H., Aylward, W., Storå, J., Jakobsson, M., and Götherström, A. (2015). Genomic
346   Evidence Establishes Anatolia as the Source of the European Neolithic Gene Pool.
347   Curr. Biol. *26*, 270–275.

348 7. Kılınç, G.M., Omrak, A., Özer, F., Günther, T., Büyükkarakaya, A.M., Bıçakçı, E., Baird,
349   D., Dönertaş, H.M., Ghalichi, A., Yaka, R., et al. (2016). The Demographic Development
350   of the First Farmers in Anatolia. Curr. Biol. *26*, 2659–2666.

351 8. Lipson, M., Szécsényi-Nagy, A., Mallick, S., Pósa, A., Stégmár, B., Keerl, V., Rohland,
352   N., Stewardson, K., Ferry, M., Michel, M., et al. (2017). Parallel palaeogenomic
353   transects reveal complex genetic history of early European farmers. Nature *551*, 368–
354   372.

355 9. Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., Sudmant,
356   P.H., Schraiber, J.G., Castellano, S., Lipson, M., et al. (2014). Ancient human genomes
357   suggest three ancestral populations for present-day Europeans. Nature *513*, 409–413.

358 10. Barrett, R., Kuzawa, C.W., McDade, T., and Armelagos, G.J. (1998). EMERGING AND
359   RE-EMERGING INFECTIOUS DISEASES: The Third Epidemiologic Transition. Annual
360   Review of Anthropology *27*, 247–271.

361 11. Mendoza-Revilla, J., Chacón-Duque, J.C., Fuentes-Guajardo, M., Ormond, L., Wang,
362   K., Hurtado, M., Villegas, V., Granja, V., Acuña-Alonzo, V., Jaramillo, C., et al. (2021).
363   Disentangling signatures of selection before and after European colonization in Latin
364   Americans. bioRxiv, 2021.11.15.467418.

365 12. Hamid, I., Korunes, K.L., Beleza, S., and Goldberg, A. (2021). Rapid adaptation to
366   malaria facilitated by admixture in the human population of Cabo Verde. Elife *10*.

367 13. Hodgson, J.A., Pickrell, J.K., Pearson, L.N., Quillen, E.E., Prista, A., Rocha, J.,
368   Soodyall, H., Shriver, M.D., and Perry, G.H. (2014). Natural selection for the Duffy-null
369   allele in the recently admixed people of Madagascar. Proc. Biol. Sci. *281*, 20140930.

370 14. Lopez, M., Choin, J., Sikora, M., Siddle, K., Harmant, C., Costa, H.A., Silvert, M.,
371   Mouguiama-Daouda, P., Hombert, J.-M., Froment, A., et al. (2019). Genomic Evidence
372   for Local Adaptation of Hunter-Gatherers to the African Rainforest. Curr. Biol. *29*, 2926–
373   2935.e4.

374 15. Patin, E., Lopez, M., Grollemund, R., Verdu, P., Harmant, C., Quach, H., Laval, G.,
375   Perry, G.H., Barreiro, L.B., Froment, A., et al. (2017). Dispersals and genetic adaptation
376   of Bantu-speaking populations in Africa and North America. Science *356*, 543–546.

377 16. Racimo, F., Gokhman, D., Fumagalli, M., Ko, A., Hansen, T., Moltke, I., Albrechtsen, A.,
378   Carmel, L., Huerta-Sánchez, E., and Nielsen, R. (2017). Archaic Adaptive Introgression
379   in TBX15/WARS2. Mol. Biol. Evol. *34*, 509–524.

380 17. Huerta-Sánchez, E., Jin, X., Asan, Bianba, Z., Peter, B.M., Vinckenbosch, N., Liang, Y.,
381   Yi, X., He, M., Somel, M., et al. (2014). Altitude adaptation in Tibetans caused by
382   introgression of Denisovan-like DNA. Nature *512*, 194–197.

383 18. Childebayeva, A., Rohrlach, A.B., Barquera, R., Rivollat, M., Aron, F., Szolek, A.,
384   Kohlbacher, O., Nicklisch, N., Alt, K.W., Gronenborn, D., et al. (2022). Population
385   Genetics and Signatures of Selection in Early Neolithic European Farmers. Molecular
386   Biology and Evolution *39*.

387 19. Ju, D., and Mathieson, I. (2021). The evolution of skin pigmentation-associated variation
388   in West Eurasia. Proc. Natl. Acad. Sci. U. S. A. *118*.

389 20. Allentoft, M.E., Sikora, M., Refoyo-Martínez, A., Irving-Pease, E.K., Fischer, A., Barrie,
390   W., Ingason, A., Stenderup, J., Sjögren, K.-G., Pearson, A., et al. (2022). Population
391   Genomics of Stone Age Eurasia. bioRxiv, 2022.05.04.490594.

392 21. Kerner, G., Neehus, A.-L., Abel, L., Casanova, J.-L., Patin, E., Laval, G., and Quintana-
393   Murci, L. (2022). Genetic adaptation to pathogens and increased risk of inflammatory

394   disorders in post-Neolithic Europe. bioRxiv, 2022.07.02.498543.

395 22. Cuadros-Espinoza, S., Laval, G., Quintana-Murci, L., and Patin, E. (2022). The genomic
396   signatures of natural selection in admixed human populations. Am. J. Hum. Genet.

397 23. Reich, D., Thangaraj, K., Patterson, N., Price, A.L., and Singh, L. (2009).
398   Reconstructing Indian population history. Nature *461*, 489–494.

399 24. Lachance, J., and Tishkoff, S.A. (2013). SNP ascertainment bias in population genetic
400   analyses: why it is important, and how to correct it. Bioessays *35*, 780–786.

401 25. Corbett-Detig, R., and Nielsen, R. (2017). A Hidden Markov Model Approach for
402   Simultaneously Estimating Local Ancestry and Admixture Time Using Next Generation
403   Sequence Data in Samples of Arbitrary Ploidy. PLoS Genet. *13*, e1006529.

404 26. Lamason, R.L., Mohideen, M.-A.P.K., Mest, J.R., Wong, A.C., Norton, H.L., Aros, M.C.,
405   Jurynec, M.J., Mao, X., Humphreville, V.R., Humbert, J.E., et al. (2005). SLC24A5, a
406   putative cation exchanger, affects pigmentation in zebrafish and humans. Science *310*,
407   1782–1786.

408 27. Voight, B.F., Kudaravalli, S., Wen, X., and Pritchard, J.K. (2006). A map of recent
409   positive selection in the human genome. PLoS Biol. *4*, e72.

410 28. Pickrell, J.K., Coop, G., Novembre, J., Kudaravalli, S., Li, J.Z., Absher, D., Srinivasan,
411   B.S., Barsh, G.S., Myers, R.M., Feldman, M.W., et al. (2009). Signals of recent positive
412   selection in a worldwide sample of human populations. Genome Res. *19*, 826–837.

413 29. Enard, D., and Petrov, D.A. (2018). Evidence that RNA Viruses Drove Adaptive
414   Introgression between Neanderthals and Modern Humans. Cell *175*, 360–371.e13.

415 30. Reynolds, A.W., Mata-Míguez, J., Miró-Herrans, A., Briggs-Cloud, M., Sylestine, A.,
416   Barajas-Olmos, F., Garcia-Ortiz, H., Rzhetskaya, M., Orozco, L., Raff, J.A., et al. (2019).
417   Comparing signals of natural selection between three Indigenous North American
418   populations. Proc. Natl. Acad. Sci. U. S. A. *116*, 9312–9317.

419 31. Hider, J.L., Gittelman, R.M., Shah, T., Edwards, M., Rosenbloom, A., Akey, J.M., and
420   Parra, E.J. (2013). Exploring signatures of positive selection in pigmentation candidate
421   genes in populations of East Asian ancestry. BMC Evol. Biol. *13*, 150.

422 32. Quillen, E.E., Bauchet, M., Bigham, A.W., Delgado-Burbano, M.E., Faust, F.X.,
423   Klimentidis, Y.C., Mao, X., Stoneking, M., and Shriver, M.D. (2012). OPRM1 and EGFR
424   contribute to skin pigmentation differences between Indigenous Americans and
425   Europeans. Hum. Genet. *131*, 1073–1080.

426 33. Eaaswarkhanth, M., Xu, D., Flanagan, C., Rzhetskaya, M., Hayes, M.G., Blekhman, R.,
427   Jablonski, N.G., and Gokcumen, O. (2016). Atopic Dermatitis Susceptibility Variants in
428   Filaggrin Hitchhike Hornerin Selective Sweep. Genome Biol. Evol. *8*, 3240–3255.

429 34. Ojeda-Granados, C., Abondio, P., Setti, A., Sarno, S., Gnecchi-Ruscone, G.A.,
430   González-Orozco, E., De Fanti, S., Jiménez-Kaufmann, A., Rangel-Villalobos, H.,
431   Moreno-Estrada, A., et al. (2022). Dietary, Cultural, and Pathogens-Related Selective
432   Pressures Shaped Differential Adaptive Evolution among Native Mexican Populations.
433   Mol. Biol. Evol. *39*.

434 35. Granka, J.M., Henn, B.M., Gignoux, C.R., Kidd, J.M., Bustamante, C.D., and Feldman,
435   M.W. (2012). Limited evidence for classic selective sweeps in African populations.

Genetics *192*, 1049–1064.

36. Mathieson, I., and Terhorst, J. (2022). Direct detection of natural selection in Bronze Age Britain. bioRxiv, 2022.03.14.484330.

37. Olalde, I., Allentoft, M.E., Sánchez-Quinto, F., Santpere, G., Chiang, C.W.K., DeGiorgio, M., Prado-Martinez, J., Rodríguez, J.A., Rasmussen, S., Quilez, J., et al. (2014). Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. Nature *507*, 225–228.

38. Skoglund, P., Malmström, H., Omrak, A., Raghavan, M., Valdiosera, C., Günther, T., Hall, P., Tambets, K., Parik, J., Sjögren, K.-G., et al. (2014). Genomic Diversity and Admixture Differs for Stone-Age Scandinavian Foragers and Farmers. Science *344*, 747–750.

39. Field, Y., Boyle, E.A., Telis, N., Gao, Z., Gaulton, K.J., Golan, D., Yengo, L., Rocheleau, G., Froguel, P., McCarthy, M.I., et al. (2016). Detection of human adaptation during the past 2000 years. Science *354*, 760–764.

40. Abi-Rached, L., Jobin, M.J., Kulkarni, S., McWhinnie, A., Dalva, K., Gragert, L., Babrzadeh, F., Gharizadeh, B., Luo, M., Plummer, F.A., et al. (2011). The shaping of modern human immune systems by multiregional admixture with archaic humans. Science *334*, 89–94.

41. Lenz, T.L. (2018). Adaptive value of novel MHC immune gene variants. Proc. Natl. Acad. Sci. U. S. A. *115*, 1414–1416.

42. Phillips, K.P., Cable, J., Mohammed, R.S., Herdegen-Radwan, M., Raubic, J., Przesmycka, K.J., van Oosterhout, C., and Radwan, J. (2018). Immunogenetic novelty confers a selective advantage in host-pathogen coevolution. Proc. Natl. Acad. Sci. U. S. A. *115*, 1552–1557.

43. Lazaridis, I., Nadel, D., Rollefson, G., Merrett, D.C., Rohland, N., Mallick, S., Fernandes, D., Novak, M., Gamarra, B., Sirak, K., et al. (2016). Genomic insights into the origin of farming in the ancient Near East. Nature *536*, 419–424.

44. Skoglund, P., Malmström, H., Raghavan, M., Storå, J., Hall, P., Willerslev, E., Gilbert, M.T.P., Götherström, A., and Jakobsson, M. (2012). Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. Science *336*, 466–469.

45. Olalde, I., Brace, S., Allentoft, M.E., Armit, I., Kristiansen, K., Booth, T., Rohland, N., Mallick, S., Szécsényi-Nagy, A., Mittnik, A., et al. (2018). The Beaker phenomenon and the genomic transformation of northwest Europe. Nature *555*, 190–196.

46. Brace, S., Diekmann, Y., Booth, T.J., van Dorp, L., Faltyskova, Z., Rohland, N., Mallick, S., Olalde, I., Ferry, M., Michel, M., et al. (2019). Ancient genomes indicate population replacement in Early Neolithic Britain. Nat Ecol Evol *3*, 765–771.

47. Günther, T., Valdiosera, C., Malmström, H., Ureña, I., Rodriguez-Varela, R., Sverrisdóttir, Ó.O., Daskalaki, E.A., Skoglund, P., Naidoo, T., Svensson, E.M., et al. (2015). Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. Proceedings of the National Academy of Sciences *112*, 11917–11922.

48. Günther, T., Malmström, H., Svensson, E.M., Omrak, A., Sánchez-Quinto, F., Kılınç, G.M., Krzewińska, M., Eriksson, G., Fraser, M., Edlund, H., et al. (2018). Population genomics of Mesolithic Scandinavia: Investigating early postglacial migration routes and

479   high-latitude adaptation. PLoS Biol. *16*, e2003703.

480   49. Yaka, R., Mapelli, I., Kaptan, D., Doğu, A., Chyleński, M., Erdal, Ö.D., Koptekin, D.,
481   Vural, K.B., Bayliss, A., Mazzucato, C., et al. (2021). Variable kinship patterns in
482   Neolithic Anatolia revealed by ancient genomes. Curr. Biol. *31*, 2455–2468.e18.

483   50. Coutinho, A., Günther, T., Munters, A.R., Svensson, E.M., Götherström, A., Storå, J.,
484   Malmström, H., and Jakobsson, M. (2020). The Neolithic Pitted Ware culture foragers
485   were culturally but not genetically influenced by the Battle Axe culture herders. Am. J.
486   Phys. Anthropol. *172*, 638–649.

487   51. Sánchez-Quinto, F., Malmström, H., Fraser, M., Girdland-Flink, L., Svensson, E.M.,
488   Simões, L.G., George, R., Hollfelder, N., Burenhult, G., Noble, G., et al. (2019).
489   Megalithic tombs in western and northern Neolithic Europe were linked to a kindred
490   society. Proc. Natl. Acad. Sci. U. S. A. *116*, 9469–9474.

491   52. Immel, A., Pierini, F., Rinne, C., Meadows, J., Barquera, R., Szolek, A., Susat, J.,
492   Böhme, L., Dose, J., Bonczarowska, J., et al. (2021). Genome-wide study of a Neolithic
493   Wartberg grave community reveals distinct HLA variation and hunter-gatherer ancestry.
494   Commun Biol *4*, 113.

495   53. Rivollat, M., Jeong, C., Schiffels, S., Küçükkalıpçı, İ., Pemonge, M.-H., Rohrlach, A.B.,
496   Alt, K.W., Binder, D., Friederich, S., Ghesquière, E., et al. (2020). Ancient genome-wide
497   DNA from France highlights the complexity of interactions between Mesolithic hunter-
498   gatherers and Neolithic farmers. Sci Adv *6*, eaaz5344.

499   54. Skourtanioti, E., Erdal, Y.S., Frangipane, M., Balossi Restelli, F., Yener, K.A., Pinnock,
500   F., Matthiae, P., Özbal, R., Schoop, U.-D., Guliyev, F., et al. (2020). Genomic History of
501   Neolithic to Bronze Age Anatolia, Northern Levant, and Southern Caucasus. Cell *181*,
502   1158–1175.e28.

503   55. Furtwängler, A., Rohrlach, A.B., Lamnidis, T.C., Papac, L., Neumann, G.U., Siebke, I.,
504   Reiter, E., Steuri, N., Hald, J., Denaire, A., et al. (2020). Ancient genomes reveal social
505   and genetic structure of Late Neolithic Switzerland. Nat. Commun. *11*, 1915.

506   56. Villalba-Mouco, V., van de Loosdrecht, M.S., Posth, C., Mora, R., Martínez-Moreno, J.,
507   Rojo-Guerra, M., Salazar-García, D.C., Royo-Guillén, J.I., Kunst, M., Rougier, H., et al.
508   (2019). Survival of Late Pleistocene Hunter-Gatherer Ancestry in the Iberian Peninsula.
509   Curr. Biol. *29*, 1169–1177.e7.

510   57. Mathieson, I., Alpaslan Roodenberg, S., Posth, C., Szécsényi-Nagy, A., Rohland, N.,
511   Mallick, S., Olalde, I., Broomandkhoshbacht, N., Cheronet, O., Fernandes, D., et al.
512   (2017). The Genomic History Of Southeastern Europe. bioRxiv.

513   58. Mathieson, I., Alpaslan-Roodenberg, S., Posth, C., Szécsényi-Nagy, A., Rohland, N.,
514   Mallick, S., Olalde, I., Broomandkhoshbacht, N., Candilio, F., Cheronet, O., et al. (2018).
515   The genomic history of southeastern Europe. Nature *555*, 197–203.

516   59. Sánchez-Quinto, F., Schroeder, H., Ramirez, O., Ávila-Arcos, M.C., Pybus, M., Olalde,
517   I., Velazquez, A.M.V., Marcos, M.E.P., Encinas, J.M.V., Bertranpetit, J., et al. (2012).
518   Genomic Affinities of Two 7,000-Year-Old Iberian Hunter-Gatherers. Curr. Biol. *22*,
519   1494–1499.

520   60. Olalde, I., Schroeder, H., Sandoval-Velasco, M., Vinner, L., Lobón, I., Ramirez, O., Civit,
521   S., García Borja, P., Salazar-García, D.C., Talamo, S., et al. (2015). A Common Genetic
522   Origin for Early Farmers from Mediterranean Cardial and Central European LBK

523   Cultures. Mol. Biol. Evol. *32*, 3132–3142.

524   61. Gamba, C., Jones, E.R., Teasdale, M.D., McLaughlin, R.L., Gonzalez-Fortes, G.,
525   Mattiangeli, V., Domboróczki, L., Kővári, I., Pap, I., Anders, A., et al. (2014). Genome
526   flux and stasis in a five millennium transect of European prehistory. Nat. Commun. *5*.

527   62. Saag, L., Varul, L., Scheib, C.L., Stenderup, J., Allentoft, M.E., Saag, L., Pagani, L.,
528   Reidla, M., Tambets, K., Metspalu, E., et al. (2017). Extensive Farming in Estonia
529   Started through a Sex-Biased Migration from the Steppe. Curr. Biol. *27*, 2185–2193 e6.

530   63. Allentoft, M.E., Sikora, M., Sjogren, K.-G., Rasmussen, S., Rasmussen, M., Stenderup,
531   J., Damgaard, P.B., Schroeder, H., Ahlstrom, T., Vinner, L., et al. (2015). Population
532   genomics of Bronze Age Eurasia. Nature *522*, 167–172.

533   64. Broushaki, F., Thomas, M.G., Link, V., López, S., van Dorp, L., Kirsanow, K.,
534   Hofmanová, Z., Diekmann, Y., Cassidy, L.M., Díez-del-Molino, D., et al. (2016). Early
535   Neolithic genomes from the eastern Fertile Crescent. Science *353*, 499.

536   65. Jones, E.R., Zarina, G., Moiseyev, V., Lightfoot, E., Nigst, P.R., Manica, A., Pinhasi, R.,
537   and Bradley, D.G. (2017). The Neolithic Transition in the Baltic Was Not Driven by
538   Admixture with Early European Farmers. Curr. Biol. *27*, 576–582.

539   66. Marchi, N., Winkelbach, L., Schulz, I., Brami, M., Hofmanová, Z., Blöcher, J., Reyna-
540   Blanco, C.S., Diekmann, Y., Thiéry, A., Kapopoulou, A., et al. (2022). The genomic
541   origins of the world's first farmers. Cell *185*, 1842–1859.e18.

542   67. Delignette-Muller, M.L., Dutang, C., and Others (2015). fitdistrplus: An R package for
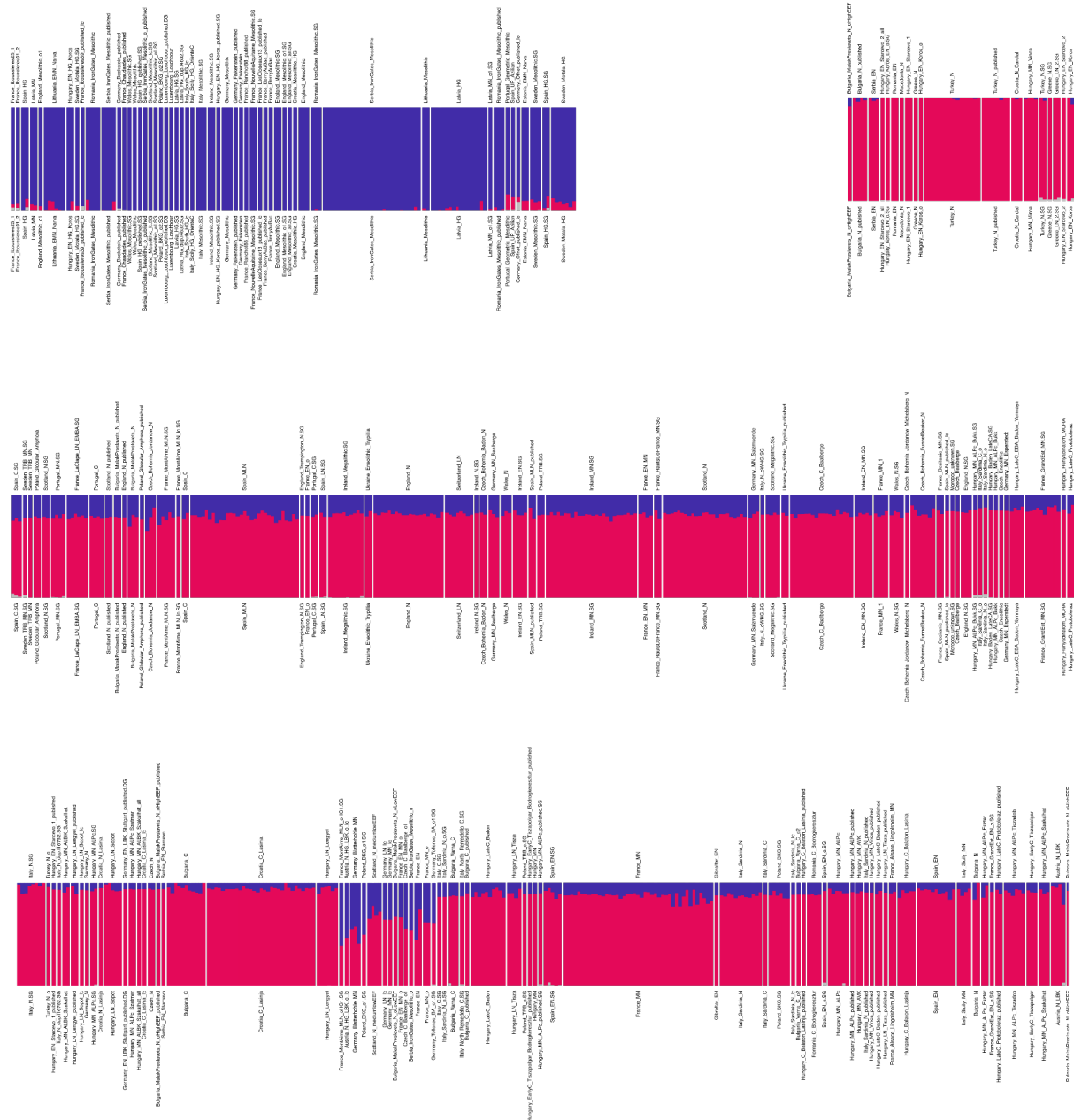543   fitting distributions. J. Stat. Softw. *64*, 1–34.

544

545
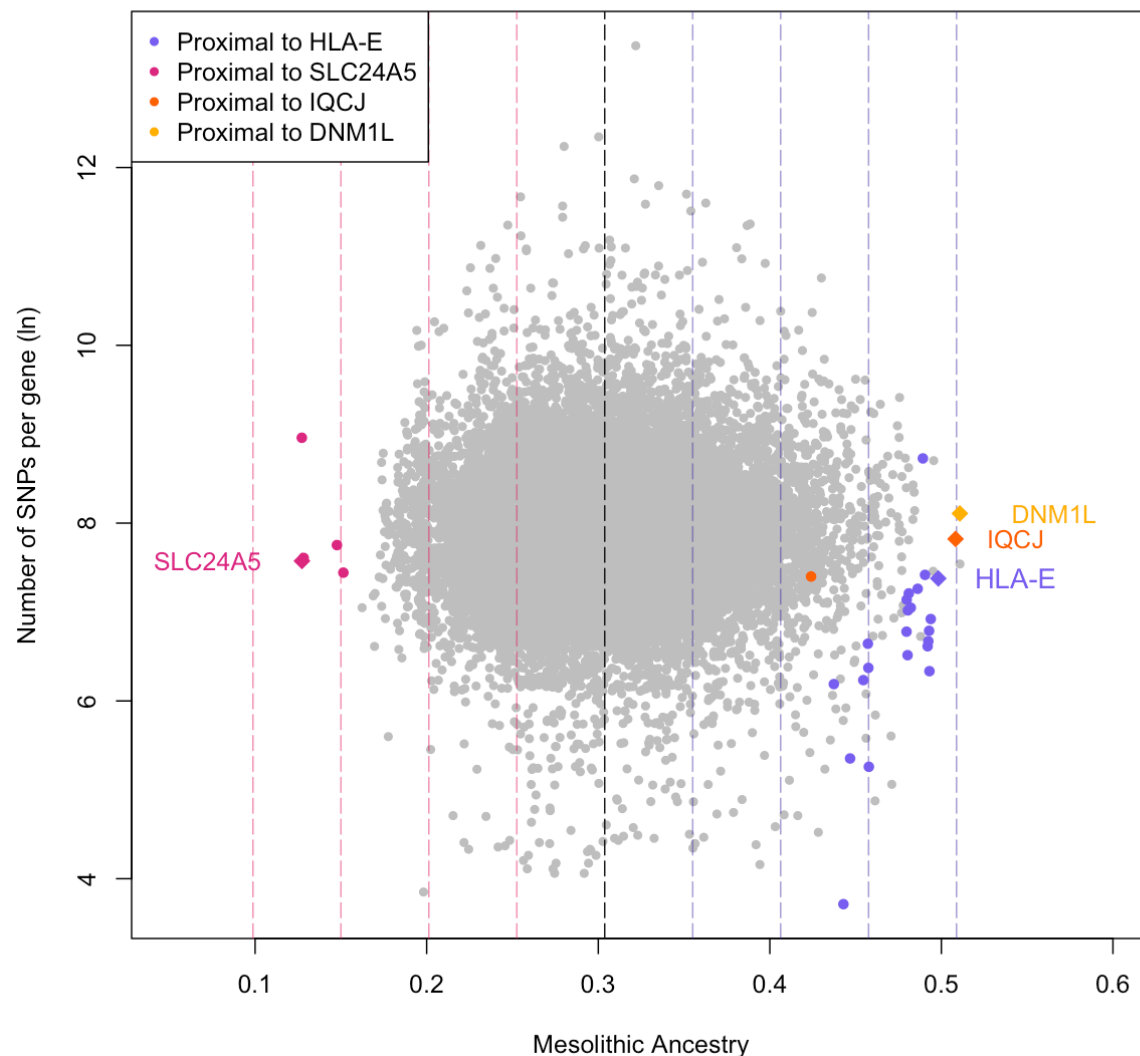
546 **Supplementary Figures**

547
548
549
550
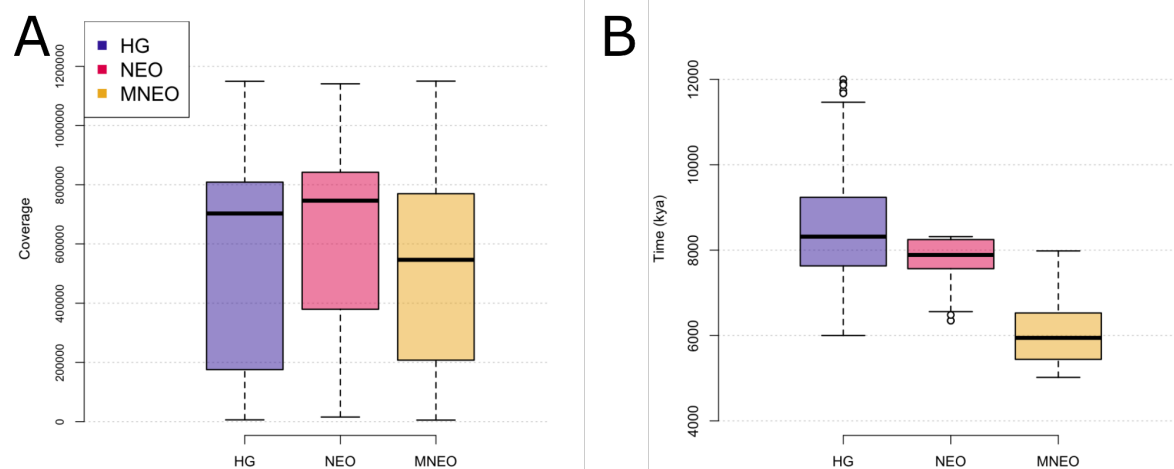
551 **Supplementary Figure 1.** Metadata of individuals used in selection scan. A**)**
552 Coverage in each population, as the number of 1240k sites covered in each
553 individual per population. **B)** Age of individuals in the 1240k selection scan.
554

555

556
Supplementary Figure 2. Clustering results of 728 HG, NEO, and MNEO individuals alongside 40 individuals from the MSL 1000 genomes panel (the latter not shown) obtained using ADMIXTURE with K=3 (Methods).

560
561

562
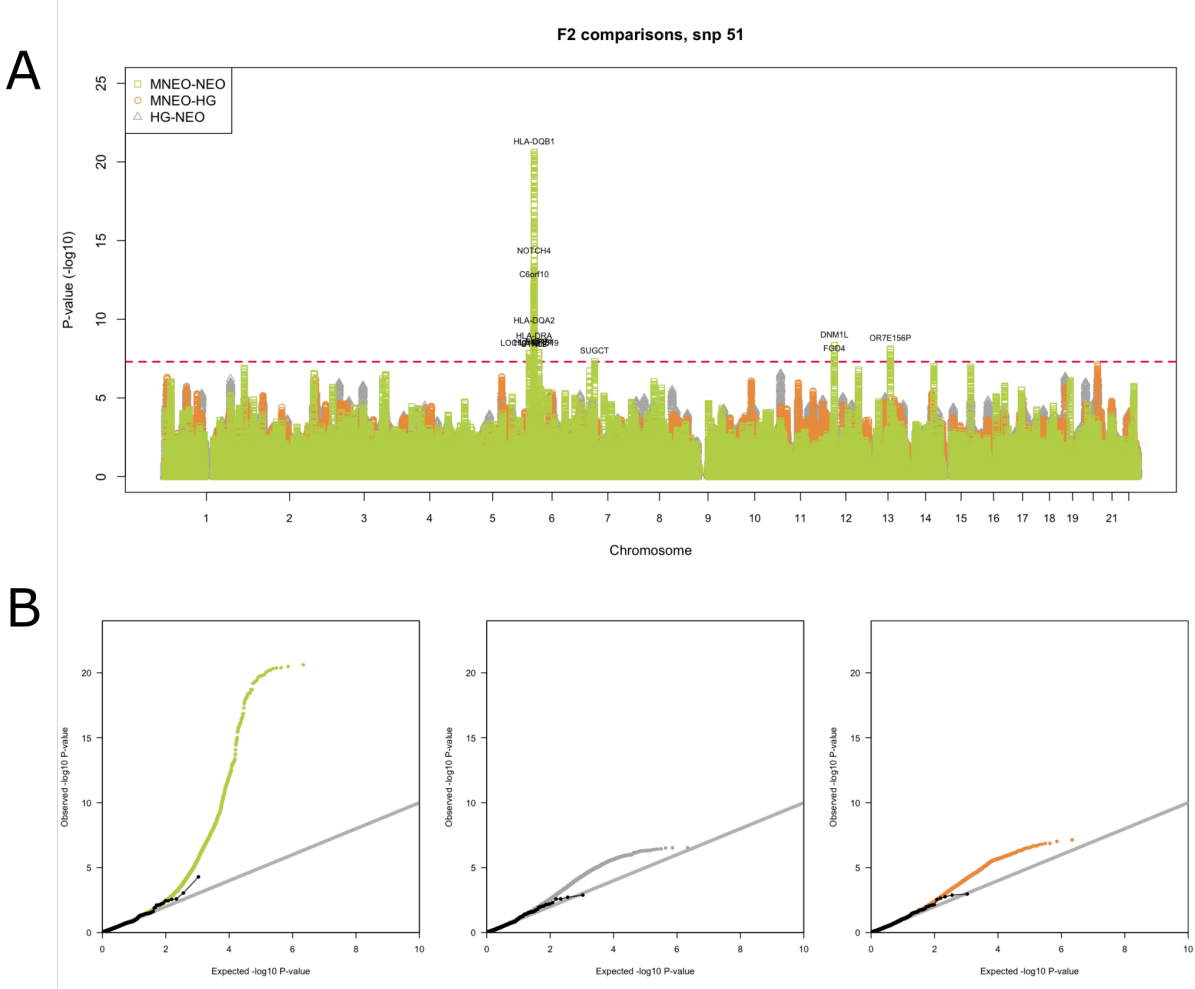
563
564 **Supplementary Figure 3.** f_2 pairwise comparisons of ancient populations. A)
565 Manhattan plot of p-values for each pairwise comparison. B) Quantile-quantile plots
566 for each pairwise comparison.
567

568

569 **Supplementary Figure 4.** Mesolithic ancestry per gene. Proximal genes are defined

570 as being within 300kb from the start or end of the gene. Dotted lines represent

571 standard deviations from the mean.

572
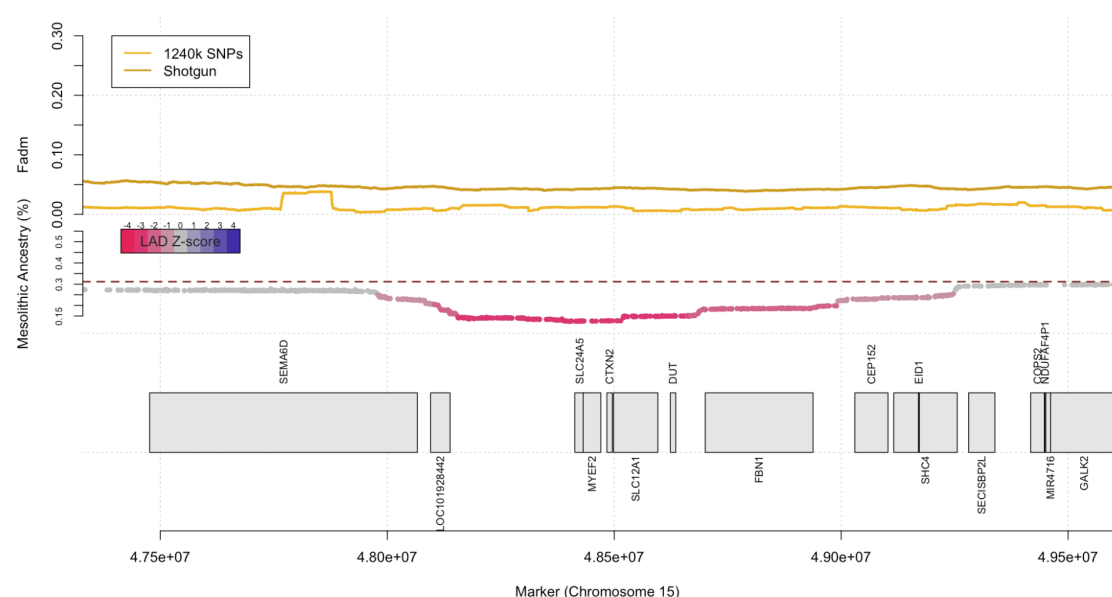
Supplementary Figure 5. Zoomed-in region of chromosome 11, with adaptive admixture statistics derived from 1240k and Shotgun data alongside local ancestry signal throughout the MHC regions I,II & III on chromosome 15.

578

579
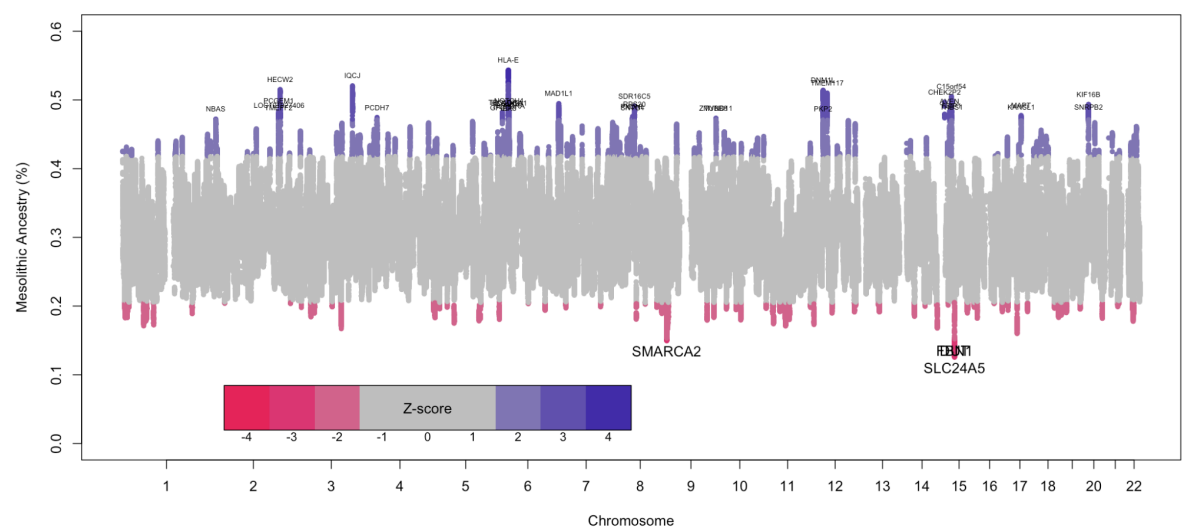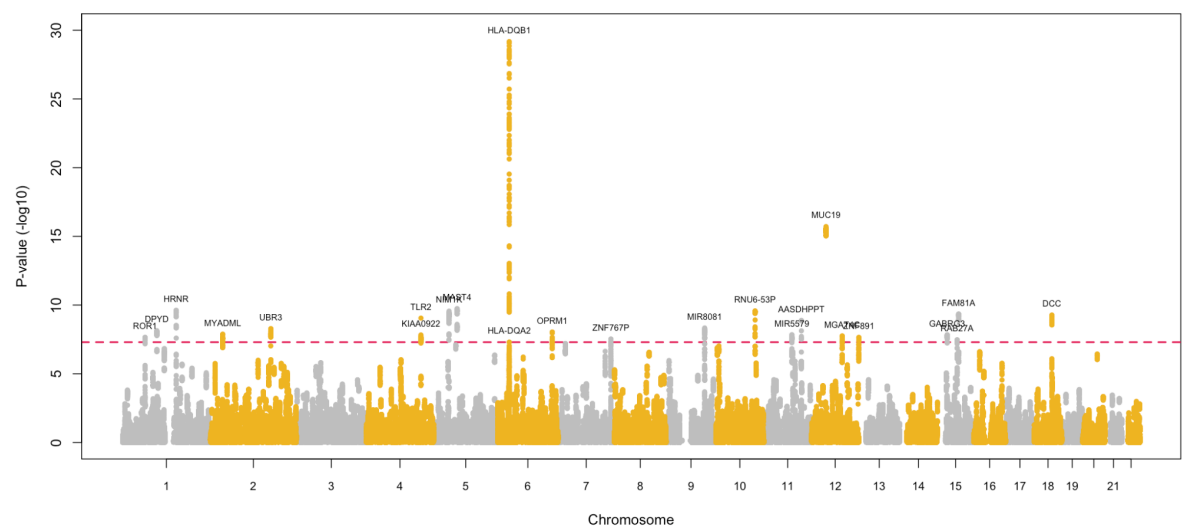
Supplementary Figure 6. Fully annotated versions of adaptive admixture test Manhattan plots as seen in figure 2. A) F_{adm} **B**) Local Ancestry Deviation.

582
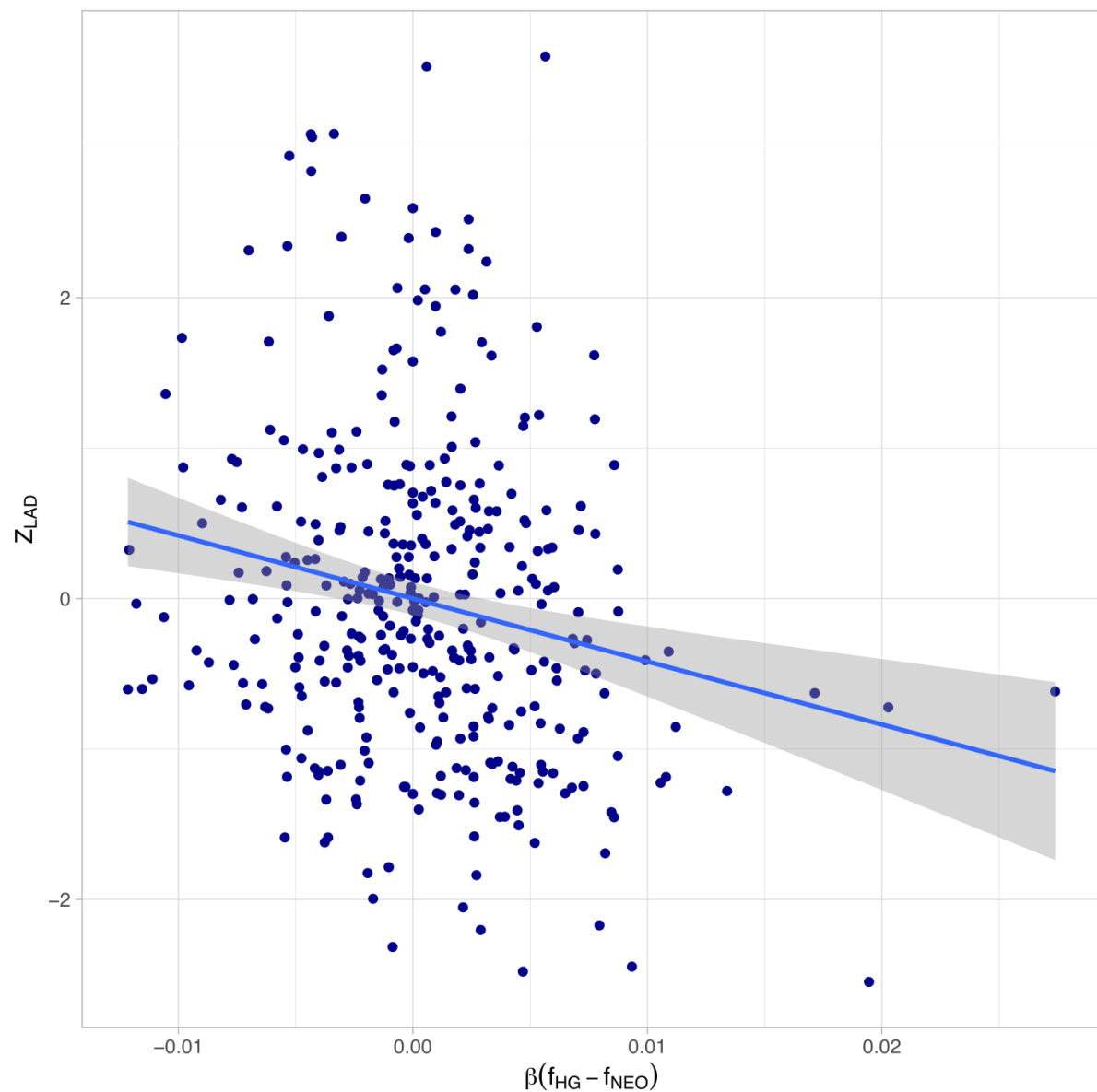
583
584 **Supplementary Figure 7.** Correlation of LAD Z-scores with hip circumference SNP
585 effect size weighted by the signed allele frequency difference between the two
586 source populations.
587