

The American Journal of Human Genetics, Volume 93

Supplemental Data

Diversity of Lactase Persistence Alleles in Ethiopia:

Signature of a Soft Selective Sweep

Bryony L. Jones, Tamiru O. Raga, Anke Liebert, Pawel Zmarz, Endashaw Bekele, E. Thomas Danielsen, Anders Krüger Olsen, Neil Bradman, Jesper T. Troelsen, and Dallas M. Swallow

	G-30210C	D-30203I	A-30182G	G-29949C	C-14011T	G-14010C	T-14009G	A-13957G	T-13915G	T-13913C	C-13910T	C-13907G	A-13806G	G-13800T	T-13730G	C-13603T	C-958T	TC-942/3ΔΔ	G-875A	A-678G	recombinant	Digester	Non-digester
G	1	A	G	C	G	T	A	T	T	C	C	A	G	T	C	C	1	G	A	no	31	86	
G	1	A	G	C	G	T	A	T	T	C	C	A	G	T	C	C	2	G	A	no	10	30	
G	1	A	G	C	G	T	A	T	T	C	C	A	G	T	C	T	2	G	A	yes	0	1	
G	1	A	G	C	G	T	A	T	T	C	C	A	G	T	T	C	1	G	A	no	3	2	
G	1	A	G	C	G	T	A	T	T	C	C	A	G	G	C	C	1	G	A	no	9	25	
G	1	A	G	C	G	T	A	T	T	T	C	A	G	T	C	C	1	G	A	no	52	1	
G	1	A	G	C	G	T	A	G	T	C	C	A	G	T	C	C	1	G	A	no	2	0	
G	1	A	G	C	G	T	A	G	T	C	C	A	G	T	C	C	1	G	G	yes	2	0	
G	1	A	G	C	G	G	A	T	T	C	C	A	G	T	C	C	2	G	A	no	23	5	
G	1	A	G	C	G	G	A	T	T	C	C	A	G	T	C	T	1	G	A	yes	1	1	
G	1	A	C	C	G	T	A	T	T	C	C	A	G	T	C	C	1	G	A	no	2	1	
G	1	A	C	C	G	T	A	T	T	C	C	A	G	T	C	C	2	G	A	no	2	0	
G	1	G	G	C	G	T	A	T	T	C	C	A	G	T	C	C	1	G	A	yes	0	2	
G	1	G	G	C	G	T	A	T	T	C	C	A	G	T	C	C	1	G	G	no	14	41	
G	1	G	G	C	G	T	A	T	T	C	C	G	G	T	C	C	1	G	G	no	5	11	
G	1	G	G	C	G	T	A	G	T	C	C	A	G	T	C	C	1	G	G	no	76	3	
G	1	G	C	C	G	T	A	G	T	C	C	A	G	T	C	C	1	G	G	no	1	0	
G	2	A	G	C	G	T	A	T	T	C	C	A	G	T	C	C	1	G	A	no	0	5	
C	1	A	G	C	G	T	A	T	T	C	C	A	G	T	C	C	1	G	A	yes	12	28	
C	1	A	G	C	G	T	A	T	T	C	C	A	G	T	C	C	2	G	A	yes	0	2	
C	1	A	G	C	G	T	A	T	T	C	C	A	G	T	C	T	1	G	A	no	59	194	
C	1	A	G	C	G	T	A	T	T	C	C	A	G	T	C	T	1	G	G	yes	1	1	
C	1	A	G	C	G	T	A	T	T	C	C	A	G	T	C	T	1	A	A	no	3	2	
C	1	A	G	C	G	T	A	T	T	C	C	A	T	T	C	T	1	G	A	no	0	1	
C	1	A	G	C	G	T	A	T	C	C	C	A	G	T	C	T	1	G	A	no	1	7	
C	1	A	G	C	G	T	G	T	T	C	C	A	G	T	C	C	1	G	A	yes	1	1	
C	1	A	G	C	C	T	A	T	T	C	C	A	G	T	C	C	1	G	A	yes	2	0	
C	1	G	G	C	G	T	A	T	T	C	C	A	G	T	C	T	1	G	A	yes	1	0	
G	1	A	G	C	G	T	A	T	T	C	C	A	G	T	C	C	1	G	A	no	0	1	
C	1	A	G	T	G	T	A	T	T	C	C	A	G	T	C	T	1	G	A	no	1	1	
C	1	A	G	C	G	T	A	T	T	C	G	A	G	T	C	T	1	G	A	?yes	0	1	
C	1	A	G	C	G	T	A	T	T	C	C	A	G	T	C	T	1	G	A	no	0	1	
Total																					314	454	

Percent recombination	
Digester	6.4 %
Non-digester	8.1 %
Enhancer allele carrier	3.9 %

Figure S1. Full set of haplotypes derived from PHASE. Using DnaSP to assess recombination by a 4 gametes based test, the Rm was 5 (minimum of 5 recombination events). Visual inspection of the phased data shows that if one excludes all the enhancer alleles and all other low frequency alleles (pale yellow) there are 3 main haplotypes for control region 1 and also three for control region 2 (coloured in red, yellow and blue). Counts of the digester and non-digester chromosomes show that just 6.4% of chromosomes of digesters are recombined, versus 8.1% for non-digesters. It should be appreciated that some or most of these recombinational events are likely to predate the enhancer mutations. The percentage of enhancer allele carrying chromosomes that are recombined is indeed lower at 3.9% , Note that the penultimate inferred haplotype on the table is likely to be a gene conversion or possibly a *de novo* mutation.

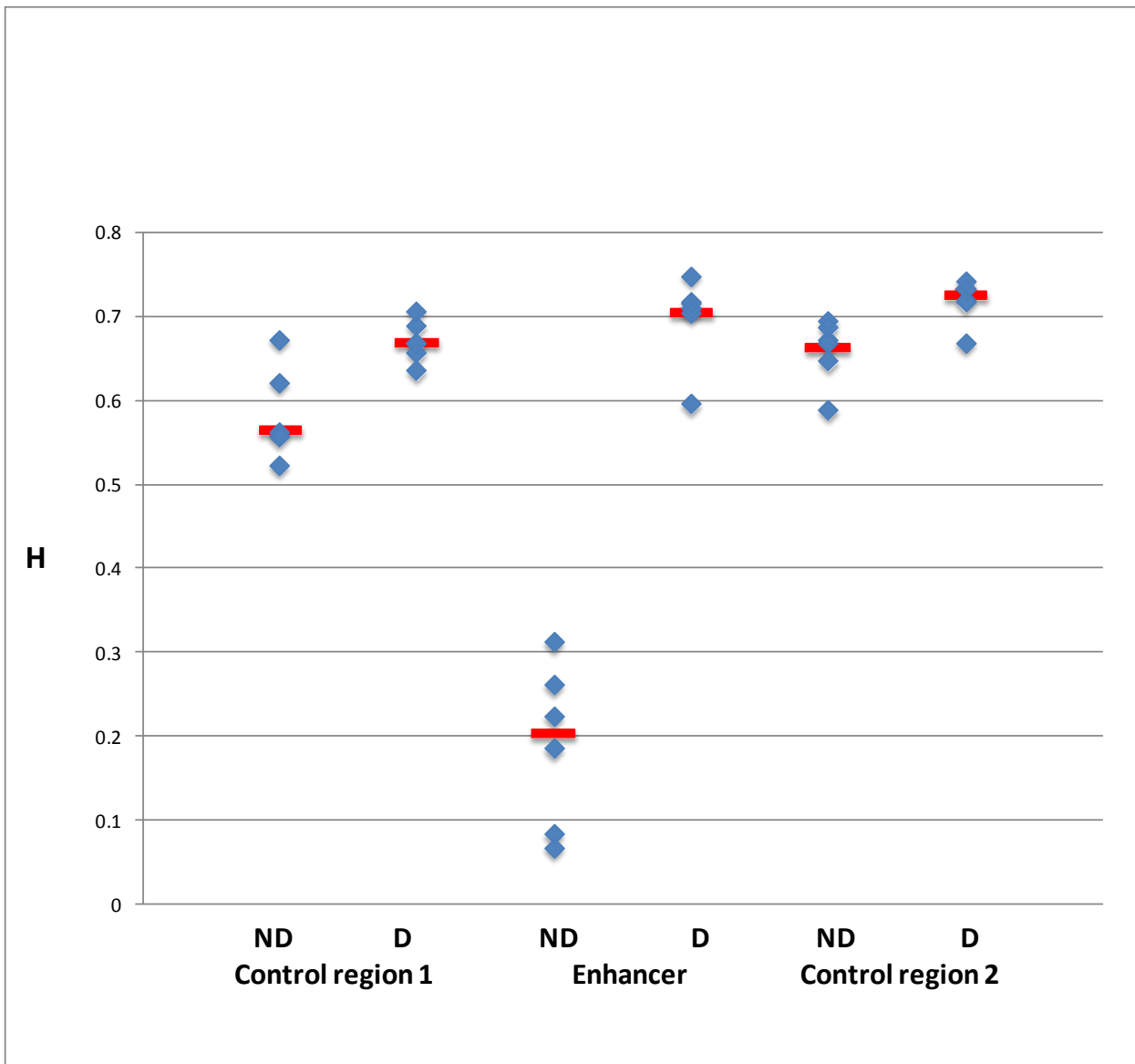


Figure S2. Comparison, in non-digesters and digesters, of the haplotype heterozygosity H measured across the three sequence regions. Data points (diamonds, some overlapping) show the five ethnic groups tested, Amhara, Tigre, Oromo, Wolayita, Somali as well as ‘other Ethiopians’ as a single group. The red horizontal bars show the median values. See Table 2 and Supplementary Table 5 for N values. D=digester, ND= non-digester.

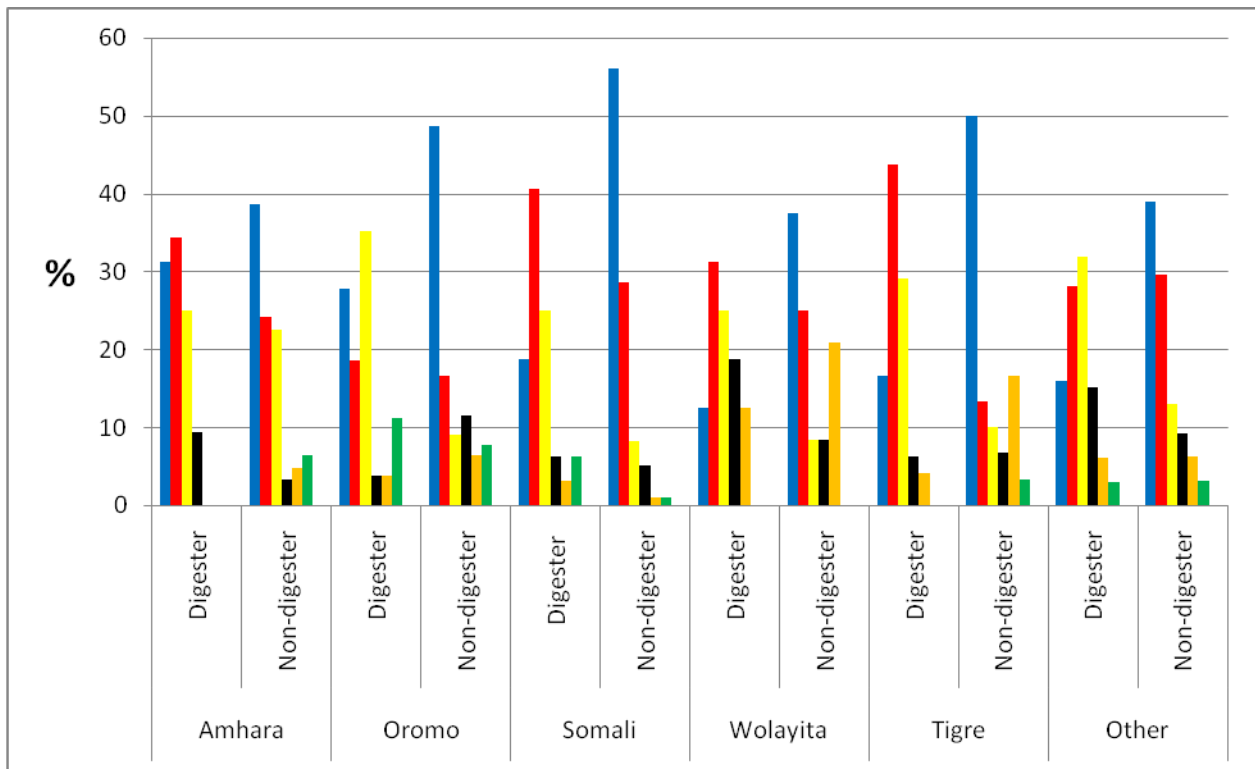


Figure S3 Flanking haplotype distribution in digesters and non-digesters by ethnic group.

The 5 most frequent combined flanking haplotypes (blue, red, yellow, orange, black) and all others combined (green), are shown as percentage frequency. In this analysis the derived variants in the enhancer region are ignored. The blue haplotype, which is the most frequent in Ethiopians, see Figure 2 of the main paper, carries no LP associated alleles, unlike the other four frequent flanking haplotypes, and is unsurprisingly more frequent in non-digesters than digesters in all groups. The most likely *LCT* gene region haplotypes in this colour coding are: blue, B; red, A or K; yellow, C; orange, P and black, X, see Figure 2.

hg19 AGGTATCAGAGTCACTTTGATATGATGAGAGCAGAGATAAACAGATTTGTTGC----- 233
panTro4 AGGTATCAGAGTCACTTTGATATGATGAGAGCAGAGATAAACAGATTTGTTGC----- 233
gorGor3 AGGTATCAGAGTCACTTTGATATGATGAGAGCAGAGATAAACAGATTTGT----- 230
nomLeu2 AGGTATCAGAGTCACTTTGATATGATGAGAGCAGAGATAAACAGATTTGTTGC----- 233
ponAbe2 AGGTATCAGAGTCACTTTGATATGATGAGAGCAGAGATAAACAGATTTGTTGC----- 233
rheMac3 AGGTATCAGAGTCACTTTGATATGATGAGAGCAGAGATAAACAGATTTGTTGCCTGTAG 240
papAnu2 AGGTATCAGAGTCACTTTGATATGATGAGAGCAGAGATAAACACATTTGTTGCCTGTAG 240

-14186

hg19 --ATGTTTTTAATCTTTGGTATGGGACATACTAGAAATTCAGTGCAAATACATTTTATGT 291
panTro4 --ATGTTTTTAATCTTTGGTATGGGACATACTAGAAATTCAGTGCAAATACATTTTATGT 291
gorGor3 -----TTTAAATCTTTGGTATGGGACATACTAGAAATTCAGTGCAAATACATTTTATGT 284
nomLeu2 --ATGTTTTTAATCTTTGGTATGGGACATACTAGAAATTCAGTGCAAATATATTTTATGT 291
ponAbe2 --ATGTTTTTAATCTTTGGTATGGGACATACTAGAAATTCACCGCAAATATATTTTATGT 291
rheMac3 GCATGTTTTTAATCTTTGGTATGGGACGTAAGAAATTCAGTGCAAATATATTTTATGT 300
papAnu2 GCATGTTTTTAATCTTTGGTATGGGACGTAAGAAATTCAGTGCAAATATATTTTATGT 300

-14126

hg19 AACTGTTGAATGCTCATACGACCATGGAATTCCTCCCTTAAAGAGCTTGGTAAGCATT 351
panTro4 AACTGTTGAATGCTCATACGACCATGGAATTCCTCCCTTAAAGAGCTTGGTAAGCATT 351
gorGor3 AACTGTTGAGTGCTCATATGACCATGGAATTCCTCCCTTAAAGAGCTTGGTAAGCATT 344
nomLeu2 AACTGTTGAGTGCTCATAGGACCATGGAATTCCTCCCTTAAAGAGCTTGGTAAGCATT 351
ponAbe2 AACTGTTGAGTGCTCATAGGACCATGGAATTCCTCCCTTAAAGAGCTTGGTAAGCATT 351
rheMac3 AACTGTCGAGTGCTCACAGGACCATGGAATTCCTCCCTTAAAGAGCTTGGTAAGCATT 360
papAnu2 AACTGTCGAGTGCTCACAGGACCATGGAATTCCTCCCTTAAAGAGCTTGGTAAGCATT 360

-14066

hg19 GAGTGTAGTTGTTAGACGGAGACGATCAGTTCATAGTTTATAGAGTGCATAAAGACCTAA 411
panTro4 GAGTGTAGTTGTTAGACGGAGACGATCAGTTCATAGTTTATAGAGTGCATAAAGACGTA 411
gorGor3 GAGTGTAGTTGTTAGACGGAGACGATCAGTTCATAGTTTATAGAGTGCATAAAGACGTA 404
nomLeu2 GAGTGTAGTTGTTAGACGGAGATGATCATGTCATAGTTTATAGAGTGCATAAAGACGTA 411
ponAbe2 GAGTGTAGTTGTTAGACGGAGATGATCAGTTCATAGTTTATAGAGTGCATAAAGACGTA 411
rheMac3 GAGTGTAGTTGTTAGACGGAGATGATCAGTTCATAGTTTATAGAGTGCATAAAGACATA 420
papAnu2 GAGTGTAGTTGTTAGACGGAGATGATCAGTTCATAGTTTATAGAGTGCATAAAGACATA 420

-14006

hg19 GTTACCATTAAATACCTTTCATTTCAGGAAAAATGACTTAGACCCACAAATGACTAGTA 471
panTro4 GTCACCATTAAATACCTTTCATTTCAGGAAAAATGACTTAGACCCACAAATGACTAGTA 471
gorGor3 GTTACCATTAAATACCTTTCATTTCAGGAAAAATGACTTAGACCCACAAATGACTAGTA 464
nomLeu2 GTTACCATTAAATACCTTTCATTTCAGGAAAAATGACTTAGACCCACAAATGACTAGTA 471
ponAbe2 GTTACCATTAAATACCTTTCATTTCAGGAAAAATGACTTAGACCCACAAATGACTAGTA 471
rheMac3 GTTACCATTAAATACCTTTCATTTCAGGAAAAATGACTTAGACCCACAAAGGACTAGTA 480
papAnu2 GTTACCATTAAATACCTTTCATTTCAGGAAAAATGACTTAGACCCACAAAGGACTAGTA 480
** *****

-13946

hg19 GGCTCTGCGCTGGCAATACAGATAAGATAATGTAGCCCTGGCTCAAAGGAACTCTCC 531
panTro4 GGCTCTGCGCTGGCAATACAGATAAGATAATGTAGCCCTGGCTCAAAGGAACTCTCC 531
gorGor3 GGCTCTGCGCTGGCAATACAGATAAGATAATGTAGCCCTGGCTCAAAGGAACTCTCC 524
nomLeu2 GGCTCTGCGCTGGCAATACAGATAAGATAATGTAGCCCTGGCTCAAAGGAACTCTCC 531
ponAbe2 GGCTCTGCGCTGGCAATACAGATAAGATAATGTAGCCCTGGCTCAAAGGAACTCTCC 531
rheMac3 GGCTCTGCGCTGGCAATACAGATAAGATAATGTAGCCCTGGCTCAAAGGAACTCTCC 540
papAnu2 GGCTCTGCGCTGGCAATACAGATAAGATAATGTAGCCCTGGCTCAAAGGAACTCTCC 540

-13886

hg19 TCCTTAGGTTGCATTTGTATAATGTTTATGATTTTATGATTTGTTCTTTGAGCCCTGCATTCC 591
panTro4 TCCTTAGGTTGCATTTGTATAATGTTTATGATTTTATGATTTGTTCTTTGAGCCCTGCATTCC 591
gorGor3 TCCTTAGGTTGCATTTGTATAATGTTTATGATTTTATGATTTGTTCTTTGAGCCCTGCATTCC 584
nomLeu2 TCCTTAGGTTGCATTTGTATAATGTTTATGATTTTATGATTTGTTCTTTGAGCCCTGCATTCC 591
ponAbe2 TCCTTAGGTTGCATTTGTATAATGTTTATGATTTTATGATTTGTTCTTTGAGCCCTGCATTCC 591
rheMac3 TCCTTAGGTTGCATTTGTATAATGTTTATGATTTTATGATTTGTTCTTTGAGCCCTGCATTCC 600
papAnu2 TCCTTAGGTTGCATTTGTATAATGTTTATGATTTTATGATTTGTTCTTTGAGCCCTGCATTCC 600

-13826

hg19 ACGAGGATAGGTCAGTGGGTATTAACAGGTAAGGGGAGTAGTACGAAAGGGCATTCA 651
panTro4 ACGAGGATAGGTCAGTGGGTATTAACAGGTAAGGGGAGTAGTACGAAAGGGCATTCA 651
gorGor3 ACGAGGATAGGTCAGTGGGTATTAACAGGTAAGGGGAGTAGTACGAAAGGGCATTCA 644
nomLeu2 ACGAGGATAGGTCAGTGGGTATTAACAGGTAAGGGGAGTAGTACGAAAGGGCATTCA 651
ponAbe2 ACGAGGATAGGTCAGTGGGTATTAACAGGTAAGGGGAGTAGTACGAAAGGGCATTCA 651
rheMac3 ACGAGGATAGGTCAGTGGGTATTAACAGGTAAGGGGAGTAGTACGAAAGGGCAATCA 660
papAnu2 ACGAGGATAGGTCAGTGGGTATTAACAGGTAAGGGGAGTAGTACGAAAGGGCAATCA 660

```

-13746
hg19      AGCGTCCCATCTTCGCTTCAACCAAAGCAGCCCTGCGTTTTTCCTAGTTTTATTAATAGGT 711
panTro4   AGCGTCCCATCTTCGCTTCAACCAAAGCAGCCCTGCTTTTTTCCTAGTTTTATTAATAGGT 711
gorGor3   AGCGTCCCATCTTCGCTTCAACCAAAGCAGCCCTGCTTTTTTCCTAGTTTTATTAATAGGT 704
nomLeu2   AGCGTCCCATCTTCGCTTCAACCAAAGCAGCCCTGCTTTTTTCCTAGTTTTATTAATAGGT 711
ponAbe2   AGCGTCCCATCTTCGCTTCAACCAAAGCAGCCCTGCTTTTTTCCTAGTTTTATTAATAGGT 711
rheMac3   AGCGTCCCATCTTCGCTTCAACCAAAGCAGCCCTGCTTTTTTCCTAGTTTTATTAATAGGT 720
papAnu2   AGCGTCCCATCTTCGCTTCAACCAAAGCAGCCCTGCTTTTTTCCTAGTTTTATTAATAGGT 720
          *****

-13686
hg19      TTGATGTAAGGTCGTCTTTGAAAAGGGGGTTGGCTTTTTTTTACAGTGTGACTGAGGTA 771
panTro4   TTGATGTAAGGTCGTCTTTGAAAAGGGGGTTGGCTTTTTTTTACAGTGTGACTGAGGTA 771
gorGor3   TTGATGTAAGGTCGTCTTTGAAAAGGGGGTTGGCTTTTTTTTACAGTGTGACTGAGGTA 764
nomLeu2   TTGATGTAAGGTCGTCTTTGAAAAGGGGGTTGGCTTTTTTTTACAGTGTGACTGAGGTA 771
ponAbe2   TTGATGTAAGGTCGTCTTTGAAAAGGGGGTTGGCTTTTTTTTACAGTGTGACTGAGGTA 771
rheMac3   TTGGTGTAAAGTTTTCTTTGAAAAGGGGGTTGGCTTTTTTTT----- 769
papAnu2   TTGGTGTAAAGTTTTCTTTGAAAAGGGGGTTGGCTTTTTTTT----- 765
          *** **

-13626
hg19      TAATTTATAAAAAGGGAAATGTATGGCATGGTGAGTTTTTTTCAATACATCCTTGTGAAT 831
panTro4   TAATTTATAAAAAGGGAAATGTATGGCATGGTGAGTTTTTTTACATACATCCTTGTGAAT 831
gorGor3   TAATTTATAAAAAGGGAAATGTATGGCATGGTGAGTTTTTTTCA--TACATCCTTGTGAAT 822
nomLeu2   TAATTTATAAAAAGGGAAATGTATGGCATGGTGAGTTTTTTTACATACATCCTTGTGAAT 831
ponAbe2   TAATTTATAAAAAGGGAAATGTATGGCATGGTGAGTTTTTTTACATACATCCTTGTGAAT 831
rheMac3   -----ACAGTATGATGAGTTTTTTTACATACATCCTTATGAAT 807
papAnu2   -----TACAGTATGATGAGTTTTTTTACATACATCCTTATGAAT 804
          * * *

-13566
hg19      ACCCAGCTCAAGATCCAAAACATTTCCATAATTTTCAGAAAGTTCCAAACCCCTGCCTCTT 891
panTro4   ACCCAGCTCAAGATCCAAAACATTTCCATAATTTTCAGAAAGTTCCAAACCCCTGCCTCTT 891
gorGor3   ACCCAGCTTAAGATCCAAAACATTTCCATAATTTTCAGAAAGTTCCAAACCCCTGCCTCTT 882
nomLeu2   ACCCAGCTCAAGATCCAAAACATTTCCATAATTTTCAGAAAGTTCCAAACCCCTGCCTCTT 891
ponAbe2   ACCCAGCTCAAGATCCAAAACATTTCCATAATTTTCAGAAAGTTCCAAACCCCTGCCTTTT 891
rheMac3   ACCCAGCTCAAGATCCAAAACATTTCCATAATTTTCAGAAAGTTCC---CCCCTGCCTCTT 863
papAnu2   ACCCAGCTCAAGATCCAAAACATTTCCATAATTTTCAGAAAGTTCC---CCCCTGCCTCTT 860
          ** *****

-13506      -13495C>T-not typed
hg19      TTCAGTCTTAGCCCTCTTCCCCTGAAGTAACACTGTTCCGACTTCAATCACTACTTTTA 951
panTro4   TTCAGTCTTAGCCCTCTTCCCCTGAAGTAACACTGTTCCGACTTCAATCACTACTTTTA 951
gorGor3   TTCAGTCTTAGCCCTCTTCCCCTGAAGTAACACTGTTCTGACTTCAATCACTACTTTTA 942
nomLeu2   TTCAGTCTTAGCCCTCTTCCCCTGAAGTAACACTGTTCTGACTTCAATCACTACTTTTC 951
ponAbe2   TTCAGTCTTAGCCCTCTTCCCCTGAAGTAACACTGTTCTGACTTCAATCACTACTTTTA 951
rheMac3   TTCAGTCTTAGCCCTCTTCCCCTGAAGTAACACTGTTCTGACTTCAATCACTACTTTTA 923
papAnu2   TTCAGTCTTAGCCCTCTTCCCCTGAAGTAACACTGTTCTGACTTCAATCACTACTTTT 920
          *****

-13446
hg19      TCCCACAGGTTAATTTTTTGGCTTTTTTC----- 980
panTro4   TCCCACAGGTTAATTTTTTGGCTTTTTTC----- 980
gorGor3   TCCCACAGGTTAATTTTTTGGCTTTTTTC----- 971
nomLeu2   TCCCACAGGTTAATTTTTTGGCTTTTTTC----- 980
ponAbe2   TCCCACAGGTTAATTTTTTGGCTTTTTTC----- 980
rheMac3   TCCCACAGGTTAATTTTTTGGCTTTTTTC----- 952
papAnu2   TTTTTTTTTTTTTTTGAGGGGAGTCTTCACTCTGTCCCCAGGCTGGAGTGCAGTGGCA 980
          * **

```

Figure S4. Sequence alignment of human MCM6 intron 13 region with 6 other primate species. Human genome sequence hg19, chimpanzee: panTro4, gorilla: gorGor3, orang-utan: ponAbe2, gibbon: nomLeu2, Rhesus macaque: rheMac3, baboon: papAnu2. Note that between position -14028 and -13800 the sequences are ~93.5% identical but that either side of this region there is increasingly greater divergence. PCR primers that delimit the fragment sequenced are shown in red. Functional SNPs mentioned in the paper are highlighted in yellow; others in grey.

Table S1. Oligonucleotide primers and PCR cycling conditions used for Sanger sequencing of the three DNA fragments.

<i>Primer</i>	<i>Primer sequence</i>	<i>Primer concentration (μM)</i>	<i>Annealing temperature</i>	<i>Number of cycles</i>
Enhancer (f)	ggacatactagaattcactgcaaatac	0.5	58 °C	39
Enhancer (r)	cctgtgggataaaagtagtgattg	0.5	58°C	39
Control region 1 (f)	accctcagatttcagcaggac	0.5	56 °C	39
Control region 1 (r)	actccatgatgattcaagcagc	0.5	56 °C	39
Control region 2 (f)	atccacattctacaggtgacaa	0.5	58 °C	39
Control region 2 (r)	ttggtgaccgggtctcactctg	0.5	58 °C	39

Table S2. Enhancer region derived allele frequency by ethnicity. Ethnic groups clustered by language family. In this analysis all unrelated individuals were included, irrespective of diagnosis. A.A. =Afro-Asiatic

Ethnic Group	Chromosome	Language family group	-14011 C>T	-14010 G>C	-14009 T>G	-13957 A>G	-13938 C>T	-13915 T>G	-13913 T>C	-13910 C>T	-13907 C>G	-13806 A>G	-13800 G>T	-13730 T>G	-13603 C>T
Amhara	108	A.A., Semitic	0	0	0.046	0	0	0.056	0.019	0	0.019	0.028	0	0.056	0
Tigre	88		0	0	0.034	0	0	0.114	0	0	0.193	0	0	0.023	0
other	42		0	0	0.048	0	0	0.17	0.048	0	0.095	0	0	0	0.024
Oromo	150	A.A., Cushitic	0	0.007	0.033	0	0.007	0.133	0.02	0	0.053	0.04	0	0.027	0.02
other	94		0.011	0.011	0.021	0.011	0	0.192	0.011	0	0.117	0.032	0	0.053	0
Wolayita	52	A.A., Omotic	0	0.019	0.039	0	0	0.115	0	0	0.078	0.019	0	0.019	0
other	12		0	0	0	0.083	0	0.25	0	0	0	0	0	0.083	0
	16	Nilotic	0	0	0	0	0	0	0	0	0	0	0	0.13	0
Other Ethiopians	178	unclassified	0.006	0	0.09	0.006	0	0.13	0.023	0.006	0.051	0.034	0.006	0.051	0

Table S3. Association of digester status with flanking markers. Data set details in supplementary Table 4. Fishers Exact 2 sided P-values of allele counts shown. Bonferroni threshold of significance (10 tests) is 0.005. Four of the flanking markers show significant differences in allele frequency between digesters and non-digesters. The two associations in which the derived allele is the more frequent in digesters (namely -678*G and -30182*G) can be attributed to the fact that in this case they are both markers of the C haplotype, which carries the -13915*G allele. -30210*C and -958*T which are both more frequent in non-digesters are both markers of a prevalent haplotype (B, F, G or L) that carries no LP alleles, the first haplotype in Figure 1 of the main paper, and the blue haplotype in Supplementary Figure 3.

		<i>Control region 1-Intron 4 MCM6</i>					<i>Control region 2 Upstream of LCT</i>			
<i>SNP</i>	<i>-30210*C</i>	<i>-30203Ins</i>	<i>-30182*G</i>	<i>-30177*G</i>	<i>-29949*C</i>	<i>-958*T</i>	<i>-942/3*ΔΔ</i>	<i>-875*A</i>	<i>-815*G</i>	<i>-678*G</i>
<i>Rs/ss</i>	<i>1435577</i>	<i>ss825678875</i>	<i>56263017</i>	<i>ss825678876</i>	<i>4988172</i>	<i>56064699</i>	<i>148142676</i>	<i>78205226</i>	<i>ss825678877</i>	<i>56211644</i>
Digester Status										
Digester	0.242	0	0.274	0	0.016	0.206	0.113	0.010	0	0.316
Non-digester	0.522	0.011	0.124	.002	.002	0.440	0.084	0.004	0.002	0.115
p-value	1.43 e-13	0.08	1.82 e-7	-	0.044	3.70 e -11	0.210	0.653	-	9.65 e -12

Table S4. Statistical comparisons of diversity between digesters and non-digesters. H values were compared as in Thomas *et al.* 2002 and π as described in the text of the main paper. N= number of chromosomes. DnaSP was used for obtaining measures of diversity.

Region		Control region 1-intron 4 MCM6					Enhancer - Intron 13 MCM6				Control region 2-upstream of LCT			
Ethnic group	Digester status	N	H	H diff p-value	π	π diff p-value	H	H diff p-value	π	π diff p-value	H	H diff p-value	π	π diff p-value
Amhara	Non-digester	62	0.672	0.952	0.0019	0.277	0.185	0.0028	0.00038	0.0021	0.694	0.457	0.0026	0.467
	Digester	32	0.669		0.0017		0.595		0.00141		0.736		0.0028	
Oromo	Non-digester	78	0.556	0.0028	0.0013	0.00016	0.261	3.1 e-09	0.00055	0.00024	0.671	0.406	0.0027	0.449
	Digester	54	0.706		0.0021		0.705		0.00185		0.716		0.0029	
Tigre	Non-digester	30	0.522	0.253	0.0014	0.6667	0.067	0	0.00013	0	0.646	0.765	0.0023	0.973
	Digester	48	0.635		0.0015		0.704		0.00182		0.668		0.0023	
Wolayita	Non-digester	24	0.565	0.353	0.0013	0.375	0.083	2.0 e-04	0.00017	0.0013	0.663	0.472	0.0022	0.503
	Digester	16	0.667		0.0016		0.717		0.00180		0.742		0.0027	
Somali	Non-digester	98	0.559	0.087	0.0013	0.081	0.223	6.9 e-10	0.00046	0	0.588	0.082	0.0021	0.166
	Digester	32	0.688		0.0017		0.714		0.00182		0.718		0.0026	
All other Ethiopians	Non-digester	162	0.621	0.211	0.0015	0	0.312	0	0.00067	0	0.686	0.070	0.0025	0.072
	Digester	132	0.652		0.0016		0.746		0.00203		0.732		0.0029	