

# Limited Evidence for Selection at the *FADS* Locus in Native American Populations

Iain Mathieson<sup>\*,1</sup>

<sup>1</sup>Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

\*Corresponding author: E-mail: mathi@pennmedicine.upenn.edu.

Associate editor: Daniel Falush

## Abstract

The *FADS* locus contains the genes *FADS1* and *FADS2* that encode enzymes involved in the synthesis of long-chain polyunsaturated fatty acids. This locus appears to have been a repeated target of selection in human evolution, likely because dietary input of long-chain polyunsaturated fatty acids varied over time depending on environment and subsistence strategy. Several recent studies have identified selection at the *FADS* locus in Native American populations, interpreted as evidence for adaptation during or subsequent to the passage through Beringia. Here, we show that these signals are confounded by independent selection—postdating the split from Native Americans—in the European and, possibly, the East Asian populations used in the population branch statistic test. This is supported by direct evidence from ancient DNA that one of the putatively selected haplotypes was already common in Northern Eurasia at the time of the separation of Native American ancestors. An explanation for the present-day distribution of the haplotype that is more consistent with the data is that Native Americans retain the ancestral state of Paleolithic Eurasians. Another haplotype at the locus may reflect a secondary selection signal, although its functional impact is unknown.

**Key words:** diet, selection, ancient DNA, *FADS*, PBS.

Long-chain polyunsaturated fatty acids are essential for many aspects of cellular and organismal function (Marszalek and Lodish 2005; Darios and Davletov 2006). Although they can be obtained from dietary sources, they can also be synthesized from short-chain polyunsaturated fatty acid (SC-PUFA) through the  $\omega$ -3 and  $\omega$ -6 biosynthesis pathways. Some of the steps in these pathways are catalyzed by the fatty acid desaturase genes *FADS1* and *FADS2* (Nakamura and Nara 2004), which are located close to each other on human chromosome 11. This locus (which we refer to as the *FADS* locus) has been targeted by selection multiple times in human evolution (Ameur et al. 2012; Mathias et al. 2012; Mathieson et al. 2015; Buckley et al. 2017; Ye et al. 2017; Mathieson and Mathieson 2018). There are two LD blocks at the locus, but most studies have focused on the two major haplotypes at LD block 1 (Ameur et al. 2012), which we refer to as the *ancestral* (A) and *derived* (D) haplotypes. Haplotype D increases expression of *FADS1* and is hypothesized to be an adaptation to a diet relatively low in PUFA, whereas haplotype A is hypothesized to be advantageous in a PUFA-rich environment (Ameur et al. 2012; Fumagalli et al. 2015). LD block 1 is also where the strongest genome-wide association study signals for lipid levels are detected in European ancestry populations (Teslovich et al. 2010). Haplotype D appears to have been under selection in Africa—indeed, this selection may have preceded the out-of-Africa bottleneck (Mathias et al. 2012; Mathieson and Mathieson 2018)—and it is virtually fixed in present-day African populations (Ameur et al. 2012; Mathias et al. 2012). Given this, it is surprising that early Eurasian populations appear to have largely carried the ancestral haplotype, suggesting selection for the ancestral haplotype at

some time after the split of present-day African and non-African lineages (Ye et al. 2017; Mathieson and Mathieson 2018). By the Mesolithic—around 10,000 years before present (BP)—the ancestral haplotype was fixed in Europe (Mathieson et al. 2015). The derived haplotype was reintroduced to Europe in the Neolithic (around 8,400 BP) by the migration of Early Farming populations from Anatolia (Mathieson et al. 2015), experienced strong positive selection during the Bronze Age (Buckley et al. 2017; Mathieson and Mathieson 2018), and is now at a frequency of around 60%. The trajectory of the derived haplotype in East Asian populations is less clear. It is common today in East Asia (~40%), and the locus does exhibit a signal of selection in East Asian populations (Suo et al. 2012; 1000 Genomes Project Consortium 2015; Chiang et al. 2018). But, with limited ancient DNA evidence, it is unclear whether this represents recent selection or ancient shared ancestry with African lineages.

## The Ancestral Haplotype Was Common in Upper Paleolithic Eurasia

We examined ancient DNA from 16 individuals from Early Upper Paleolithic Eurasia (Fu et al. 2014, 2016; Raghavan et al. 2014; Sikora et al. 2017, 2019; Yang et al. 2017). Of these individuals' 32 haplotypes, 4 are derived and 28 are ancestral (3 vs. 23 supported by more than 6 reads; fig. 1b and supplementary table 1, Supplementary Material online). This confirms that the derived LD block 1 haplotype was uncommon, though not completely absent, throughout Upper Paleolithic Eurasia. With a sufficiently intense bottleneck, genetic

© The Author(s) 2020. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Open Access

drift could fix the ancestral haplotype in the ancestors of Native Americans even if it was not under selection (Harris et al. 2019).

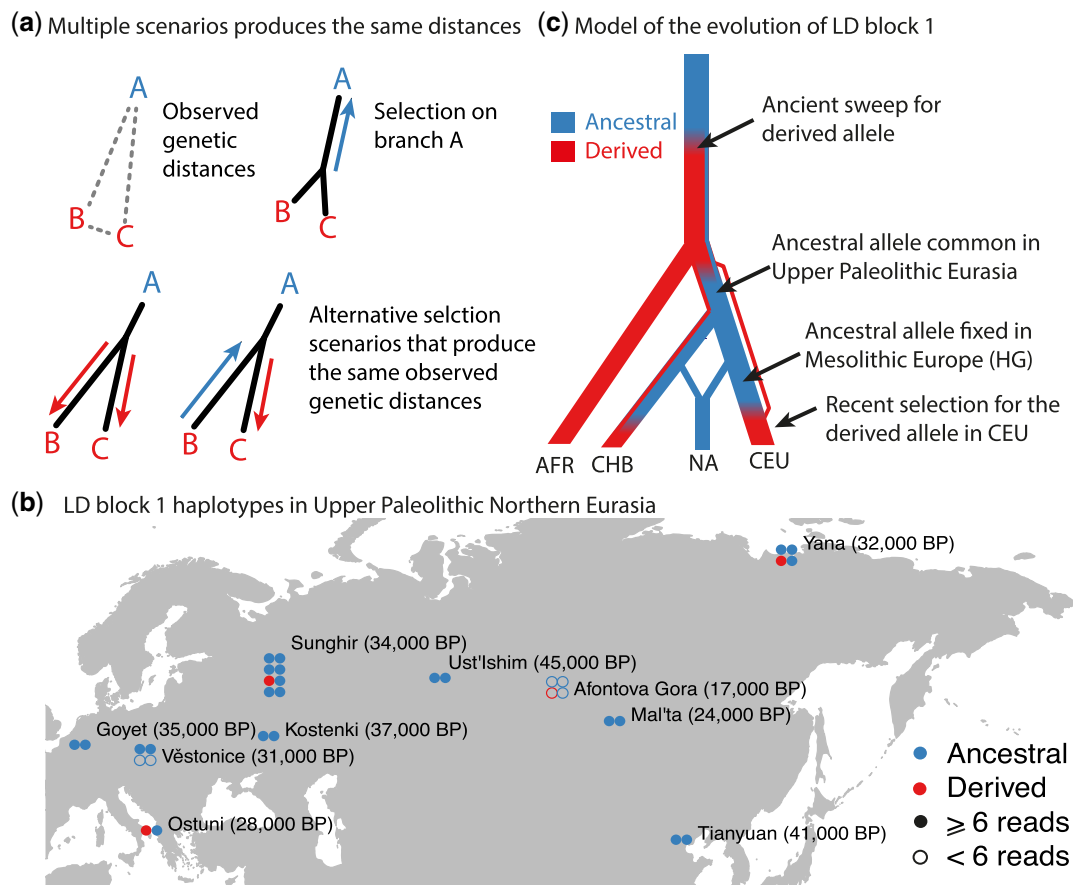
## Selection Scans at the Locus Are Confounded by Independent Adaptation

Excluding the effects of recent admixture, present-day Native American, Inuit, and Siberian populations are fixed for the ancestral haplotype (Ameur et al. 2012; Fumagalli et al. 2015; Harris et al. 2019). Fumagalli et al. (2015) found a strong signal of selection at the locus in the Greenlandic Inuit population, which they interpret as an adaptation to a PUFA-rich Arctic diet at least 20,000 years ago in the common ancestors of present-day Inuit and Siberian populations. Subsequent studies detected a similar signal in Native American populations (Amorim et al. 2017; Harris et al. 2019). Some studies interpret this as evidence for selection for the ancestral haplotype relatively early on the Native American lineage (Hlusko et al. 2018), although it could also represent a shared signal from the population ancestral to both Native Americans and Inuit.

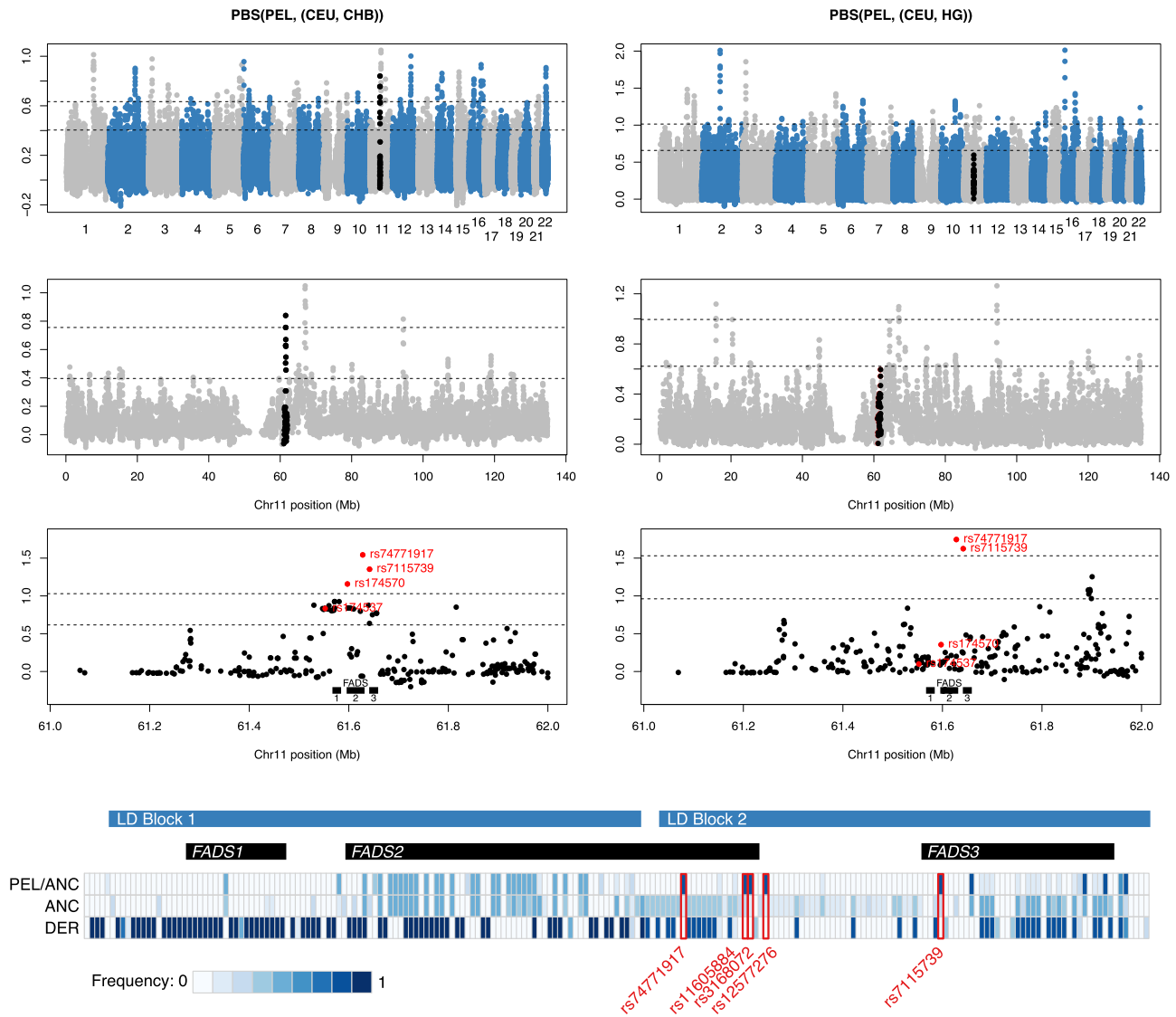
Both these signals extend across LD blocks 1 (tagging the ancestral haplotype) and 2.

These studies used the population branch statistic (PBS) (Yi et al. 2010) to compare Native American (NA), European (EUR), and East Asian (EAS) populations. The PBS compares genetic differentiation between three populations and identifies which, if any, of the three branches has excess differentiation that indicates selection. We write  $PBS(A, (B, C))$  to denote the statistic that is testing for excess differentiation on the A branch, relative to B and C. However, the PBS makes the implicit assumption that each of the three branches is independent. If there is parallel selection on the same haplotype in two population, say B and C, then B and C will be similar to each other, but each will be highly diverged relative to A. Thus, the PBS will misattribute selection to branch A (fig. 1a). Given low frequency of the derived allele in Upper Paleolithic Eurasia (fig. 1b) and high frequencies in present-day European and East Asian populations, we expect that  $PBS(NA, (EUR, EAS))$  would give a spurious signal of selection in the Native American population (fig. 1c).

We tested this by computing the PBS using Native American, European, and Mesolithic European populations.



**Fig. 1.** (a) The PBS compares genetic differentiation (branch length) between three populations. If a new mutation (blue) is under selection in one population (A), the branch leading to A will be longer—a signal of selection. If a haplotype that exists in the ancestral population (red) is under selection in both populations B and C then, because B and C look very similar, the PBS misattributes the long branch to A, instead of to B and C. (b) LD block 1 haplotypes in Upper Paleolithic Eurasia (Fu et al. 2014, 2016; Raghavan et al. 2014; Sikora et al. 2017, 2019; Yang et al. 2017). Ancestral and derived haplotype defined as haplotypes A and C (Mathieson and Mathieson 2018). (c) Model for the evolution of present-day African (AFR), East Asian (CHB), Native American (NA), and European (CEU) population showing where derived (red) and ancestral (blue) haplotypes are common.



**Fig. 2.** PBS on the Native American branch. Left column: PBS(PEL,(CEU, CHB)). Right column: PBS(PEL,(CEU, HG)). Upper row: genome-wide PBS in overlapping 20-SNP windows, shifted by 5 SNPs. Black points indicate the region Chr11:61-62Mb (hg19). Second row: chromosome 11 PBS in overlapping 20-SNP windows. Third row: per-SNP PBS in the region Chr11:61-62Mb. Horizontal lines indicate upper 0.01 and 0.001 genome-wide PBS quantiles. Red labeled points indicate SNPs previously identified as targets of selection (Fumagalli et al. 2015; Amorim et al. 2017). Top three rows restricted to 903,961 autosomal SNPs present on the 1,240k capture array with a minor allele frequency of at least 5% in at least one of the four populations. Fourth row: frequencies for all SNPs at >1% frequency in at least one population in CEU and CHB individuals carrying the derived haplotype (DER), the ancestral haplotype (ANC), and for PEL individuals carrying the ancestral haplotype (PEL/ANC). Color indicates the frequency of the variant that is rarer on the ancestral haplotype. Highlighted in red are five LD block 2 SNPs that have >50% difference in frequency between ANC and PEL/ANC, and at most 10% frequency in DER.

We used the Peruvians from Lima (PEL) 1000 Genomes population to represent Native Americans (1000 Genomes Project Consortium 2015). For each PEL individual, we restricted to regions of homozygous Native American ancestry (Martin et al. 2017) (93 chromosomes, on average). We used CEU (Northern and Western European ancestry) to represent present-day Europeans, CHB (Chinese from Beijing) to represent East Asians and 150 ancient European hunter-gatherers (HG, 68 chromosomes, on average) to represent Mesolithic Europe (Skoglund et al. 2012; Gamba et al. 2014; Lazaridis et al. 2014; Olalde et al. 2014, 2018; Allentoft et al. 2015; Jones et al.

2015; Mathieson et al. 2015, 2018; González-Forbes et al. 2017; Lipson et al. 2017; Saag et al. 2017). We replicate the elevated PBS(PEL, (CEU, CHB)) statistic at the *FADS* locus (fig. 2 left column; upper 0.0002 quantile), but find that it largely disappears if we replace CHB with HG (fig. 2 right column; upper 0.016 quantile). We see similar results if we replace CEU instead of CHB (supplementary fig. 1, Supplementary Material online), or replace PEL with a diverse panel of 28 unadmixed Native American genomes from the Simons Genome Diversity Project (Mallick et al. 2016) (supplementary fig. 2, Supplementary Material online). Two of the LD

block 2 single-nucleotide polymorphisms (SNPs) originally identified by Fumagalli et al. (2015)—rs74771917 and rs7115739—still have extreme PBS values, but PBS is no longer elevated across the locus (fig. 2). Conversely, both PBS(CEU, (PEL, HG)) and PBS(CHB, (PEL, HG)) do show elevated values (upper 0.00004 and 0.0006 quantiles; supplementary fig. 3, Supplementary Material online). These results are consistent with selection in CEU and either selection or retention of the derived haplotype in CHB.

## A Potential Secondary Signal of Selection

Although the Native American PBS signal in LD block 1 (containing rs174570, rs174556, and rs174537) disappears when Mesolithic HG are used as an outgroup, two SNPs (rs74771917 and rs7115739) in LD block 2 remain significant (fig. 2). We also identified three additional SNPs that are highly differentiated between present-day Native American and Eurasian ancestral allele carriers, using sequence data from the 1000 Genomes Project (fig. 2, lower panel). Among individuals who carry the ancestral haplotype at LD block 1, this LD block 2 haplotype has a frequency of around 100% in the Greenland Inuit (Fumagalli et al. 2015), 82% in PEL, 34% in CHB, 9% in HG, and 4% in CEU. The Anzick individual (Rasmussen et al. 2014) carries two copies, whereas the 40,000-year old East Asian Tianyuan individual (Yang et al. 2017) carries one. It therefore remains possible that the high frequency of this haplotype represents a secondary signal of selection in the common ancestor of Inuit and Native Americans. On the other hand, this region is not a genome-wide outlier in the standard window-based PBS analysis, and the frequency of this haplotype may have been driven by linked selection on LD block 1. Further, the LD block 2 haplotype has not been shown to affect expression of any of the *FADS* genes or any other phenotype, independent of the LD block 1 haplotype. Within-population, the two blocks are highly correlated, so it would be necessary to perform conditional analysis at the locus in East Asian populations to identify an independent effect of the block 2 haplotype.

## Conclusion

We demonstrated that the apparent signal of selection for the ancestral allele at the *FADS* locus in Native American populations could actually reflect high frequencies of the derived allele in Europe and East Asia. Direct evidence from ancient DNA demonstrates recent (Holocene) selection for the derived allele in Europe, whereas its high frequency in East Asia could be the result of either recent selection or more ancient shared ancestry. The ancestral haplotype at LD block 1 may have been selected in Upper Paleolithic Eurasian populations but this likely took place around or before the split of East and West Eurasian populations 40–60,000 years ago and certainly before the Native American and Siberian lineages split. There remains some evidence of a secondary signal of selection in LD block 2 but this is shared by Inuit and Siberians and not specific to Native Americans. The complex history of selection at this locus likely confounds selection scans in other

populations as well. Finally, this analysis demonstrates a limitation of the commonly used PBS selection test, and the ability of direct evidence from ancient DNA to resolve complex evolutionary histories that may not be identifiable using present-day data.

## Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

## Acknowledgments

I.M. was supported by a Research Fellowship from the Alfred P. Sloan Foundation (FG-2018-10647), a New Investigator Research Grant from the Charles E. Kaufman Foundation (KA2018-98559), and NIGMS award number (R35GM133708). The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health. The author thanks Shai Carmi, Matteo Fumagalli, Rasmus Nielsen, Fernando Racimo, and Pontus Skoglund for helpful comments on earlier drafts.

## References

- 1000 Genomes Project Consortium. 2015. A global reference for human genetic variation. *Nature* 526:68–74.
- Allentoft ME, Sikora M, Sjogren K-G, Rasmussen S, Rasmussen M, Stenderup J, Damgaard PB, Schroeder H, Ahlstrom T, Vinner L, et al. 2015. Population genomics of Bronze Age Eurasia. *Nature* 522(7555):167–172.
- Ameur A, Enroth S, Johansson A, Zaboli G, Igl W, Johansson AC, Rivas MA, Daly MJ, Schmitz G, Hicks AA, et al. 2012. Genetic adaptation of fatty-acid metabolism: a human-specific haplotype increasing the biosynthesis of long-chain omega-3 and omega-6 fatty acids. *Am J Hum Genet.* 90(5):809–820.
- Amorim CE, Nunes K, Meyer D, Comas D, Bortolini MC, Salzano FM, Hunemeier T. 2017. Genetic signature of natural selection in first Americans. *Proc Natl Acad Sci U S A.* 114:2195–2199.
- Buckley MT, Racimo F, Allentoft ME, Jensen MK, Jonsson A, Huang H, Hormozdiari F, Sikora M, Marnetto D, Eskin E, et al. 2017. Selection in Europeans on fatty acid desaturases associated with dietary changes. *Mol Biol Evol.* 34(6):1307–1318.
- Chiang CWK, Mangul S, Robles C, Sankararaman S. 2018. A comprehensive map of genetic variation in the world's largest ethnic group—Han Chinese. *Mol Biol Evol.* 35(11):2736–2750.
- Darios F, Davletov B. 2006. Omega-3 and omega-6 fatty acids stimulate cell membrane expansion by acting on syntaxin 3. *Nature* 440(7085):813–817.
- Fu Q, Li H, Moorjani P, Jay F, Slepchenko SM, Bondarev AA, Johnson PL, Aximu-Petri A, Prufer K, de Filippo C, et al. 2014. Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* 514(7523):445–449.
- Fu Q, Posth C, Hajdinjak M, Petr M, Mallick S, Fernandes D, Furtwängler A, Haak W, Meyer M, Mittnik A, et al. 2016. The genetic history of Ice Age Europe. *Nature* 534(7606):200–205.
- Fumagalli M, Moltke I, Grarup N, Racimo F, Bjerregaard P, Jorgensen ME, Korneliusen TS, Gerbault P, Skotte L, Linneberg A, et al. 2015. Greenlandic Inuit show genetic signatures of diet and climate adaptation. *Science* 349(6254):1343–1347.
- Gamba C, Jones ER, Teasdale MD, McLaughlin RL, Gonzalez-Fortes G, Mattiangeli V, Domboróczki L, Kóvári I, Pap I, Anders A, et al. 2014. Genome flux and stasis in a five millennium transect of European prehistory. *Nat Commun.* 5(1):5257.
- González-Fortes G, Jones ER, Lightfoot E, Bonsall C, Lazar C, Grandal-d'Anglade A, Garralda MD, Drak L, Siska V, Simalcsik A, et al. 2017.

- Paleogenomic evidence for multi-generational mixing between Neolithic farmers and Mesolithic hunter-gatherers in the lower Danube Basin. *Curr Biol.* 27(12):1801–1810.
- Harris DN, Ruczinski I, Yanek LR, Becker LC, Becker DM, Guio H, Cui T, Chilton FH, Mathias RA, O'Connor T. 2019. Evolution of hominin polyunsaturated fatty acid metabolism: from Africa to the New World. *Genome Biol Evol.* 11(5):1417–1430.
- Hlusko LJ, Carlson JP, Chaplin G, Elias SA, Hoffecker JF, Huffman M, Jablonski NG, Monson TA, O'Rourke DH, Pilloud MA, et al. 2018. Environmental selection during the last ice age on the mother-to-infant transmission of vitamin D and fatty acids through breast milk. *Proc Natl Acad Sci U S A.* 115(19):E4426–E4432.
- Jones ER, Gonzalez-Fortes G, Connell S, Siska V, Eriksson A, Martiniano R, McLaughlin RL, Gallego Llorente M, Cassidy LM, Gamba C, et al. 2015. Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat Commun.* 6(1):8912.
- Lazaridis I, Patterson N, Mittnik A, Renaud G, Mallick S, Kirsanow K, Sudmant PH, Schraiber JG, Castellano S, Lipson M, et al. 2014. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* 513(7518):409–413.
- Lipson M, Szecsenyi-Nagy A, Mallick S, Posa A, Stegmar B, Keerl V, Rohland N, Stewardson K, Ferry M, Michel M, et al. 2017. Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* 551(7680):368–372.
- Mallick S, Li H, Lipson M, Mathieson I, Gymrek M, Racimo F, Zhao M, Chennagiri N, Nordenfelt S, Tandon A, et al. 2016. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* 538(7624):201–206.
- Marszalek JR, Lodish HF. 2005. Docosahexaenoic acid, fatty acid-interacting proteins, and neuronal function: breastmilk and fish are good for you. *Annu Rev Cell Dev Biol.* 21(1):633–657.
- Martin AR, Gignoux CR, Walters RK, Wojcik GL, Neale BM, Gravel S, Daly MJ, Bustamante CD, Kenny EE. 2017. Human demographic history impacts genetic risk prediction across diverse populations. *Am J Hum Genet.* 100(4):635–649.
- Mathias RA, Fu W, Akey JM, Ainsworth HC, Torgerson DG, Ruczinski I, Sergeant S, Barnes KC, Chilton FH. 2012. Adaptive evolution of the FADS gene cluster within Africa. *PLoS One* 7(9):e44926.
- Mathieson I, Alpaslan-Roodenberg S, Posth C, Szecsenyi-Nagy A, Rohland N, Mallick S, Olalde I, Broomandkhoshbacht N, Candilio F, Cheronet O, et al. 2018. The genomic history of southeastern Europe. *Nature* 555(7695):197–203.
- Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, Roodenberg SA, Harney E, Stewardson K, Fernandes D, Novak M, et al. 2015. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* 528(7583):499–503.
- Mathieson S, Mathieson I. 2018. FADS1 and the timing of human adaptation to agriculture. *Mol Biol Evol.* 35(12):2957–2970.
- Nakamura MT, Nara TY. 2004. Structure, function, and dietary regulation of delta6, delta5, and delta9 desaturases. *Annu Rev Nutr.* 24(1):345–376.
- Olalde I, Allentoft ME, Sanchez-Quinto F, Santpere G, Chiang CW, DeGiorgio M, Prado-Martinez J, Rodriguez JA, Rasmussen S, Quilez J, et al. 2014. Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature* 507(7491):225–228.
- Olalde I, Brace S, Allentoft ME, Armit I, Kristiansen K, Booth T, Rohland N, Mallick S, Szecsenyi-Nagy A, Mittnik A, et al. 2018. The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature* 555(7695):190–196.
- Raghavan M, Skoglund P, Graf KE, Metspalu M, Albrechtsen A, Moltke I, Rasmussen S, Stafford TW Jr, Orlando L, Metspalu E, et al. 2014. Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* 505(7481):87–91.
- Rasmussen M, Anzick SL, Waters MR, Skoglund P, DeGiorgio M, Stafford TW Jr, Rasmussen S, Moltke I, Albrechtsen A, Doyle SM, et al. 2014. The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature* 506(7487):225–229.
- Saag L, Varul L, Scheib CL, Stenderup J, Allentoft ME, Saag L, Pagani L, Reidla M, Tambets K, Metspalu E, et al. 2017. Extensive farming in Estonia started through a sex-biased migration from the Steppe. *Curr Biol.* 27(14):2185–2193.e6.
- Sikora M, Pitulko VV, Sousa VC, Allentoft ME, Vinner L, Rasmussen S, Margaryan A, de Barros Damgaard P, de la Fuente C, Renaud G, et al. 2019. The population history of northeastern Siberia since the Pleistocene. *Nature* 570(7760):182–188.
- Sikora M, Seguin-Orlando A, Sousa VC, Albrechtsen A, Korneliusen T, Ko A, Rasmussen S, Dupanloup I, Nigst PR, Bosch MD, et al. 2017. Ancient genomes show social and reproductive behavior of early Upper Paleolithic foragers. *Science* 358(6363):659–662.
- Skoglund P, Malmström H, Raghavan M, Storå J, Hall P, Willerslev E, Gilbert MTP, Götherström A, Jakobsson M. 2012. Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* 336(6080):466–469.
- Suo C, Xu H, Khor CC, Ong RT, Sim X, Chen J, Tay WT, Sim KS, Zeng YX, Zhang X, et al. 2012. Natural positive selection and north-south genetic diversity in East Asia. *Eur J Hum Genet.* 20(1):102–110.
- Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, Pirruccello JP, Ripatti S, Chasman DI, Willer CJ, et al. 2010. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 466(7307):707–713.
- Yang MA, Gao X, Theunert C, Tong H, Aximu-Petri A, Nickel B, Slatkin M, Meyer M, Pääbo S, Kelso J, et al. 2017. 40,000-year-old individual from Asia provides insight into early population structure in Eurasia. *Curr Biol.* 27(20):3202–3208.e3209.
- Ye K, Gao F, Wang D, Bar-Yosef O, Keinan A. 2017. Dietary adaptation of FADS genes in Europe varied across time and geography. *Nat Ecol Evol.* 1(7):167.
- Yi X, Liang Y, Huerta-Sanchez E, Jin X, Cuo ZX, Pool JE, Xu X, Jiang H, Vinckenbosch N, Korneliusen TS, et al. 2010. Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* 329:75–78.