

Diet and the evolution of human amylase gene copy number variation

George H Perry^{1,2}, Nathaniel J Dominy³, Katrina G Claw^{1,4}, Arthur S Lee², Heike Fiegler⁵, Richard Redon⁵, John Werner⁴, Fernando A Villanea³, Joanna L Mountain⁶, Rajeev Misra⁴, Nigel P Carter⁵, Charles Lee^{2,7,8} & Anne C Stone^{1,8}

Starch consumption is a prominent characteristic of agricultural societies and hunter-gatherers in arid environments. In contrast, rainforest and circum-arctic hunter-gatherers and some pastoralists consume much less starch^{1–3}. This behavioral variation raises the possibility that different selective pressures have acted on amylase, the enzyme responsible for starch hydrolysis⁴. We found that copy number of the salivary amylase gene (*AMY1*) is correlated positively with salivary amylase protein level and that individuals from populations with high-starch diets have, on average, more *AMY1* copies than those with traditionally low-starch diets. Comparisons with other loci in a subset of these populations suggest that the extent of *AMY1* copy number differentiation is highly unusual. This example of positive selection on a copy number-variable gene is, to our knowledge, one of the first discovered in the human genome. Higher *AMY1* copy numbers and protein levels probably improve the digestion of starchy foods and may buffer against the fitness-reducing effects of intestinal disease.

Hominin evolution is characterized by significant dietary shifts, facilitated in part by the development of stone tool technology, the control of fire and, most recently, the domestication of plants and animals^{5–7}. Starch, for instance, has become an increasingly prominent component of the human diet, particularly among agricultural societies⁸. It stands to reason, therefore, that studies of the evolution of amylase in humans and our close primate relatives may provide insight into our ecological history. Because the human salivary amylase gene (*AMY1*) shows extensive variation in copy number^{9,10}, we first assessed whether a functional relationship exists between *AMY1* copy number and the amount of amylase protein in saliva. We then determined if *AMY1* copy number differs among modern human populations with contrasting amounts of dietary starch.

We estimated diploid *AMY1* gene copy number for 50 European Americans using an *AMY1*-specific real-time quantitative PCR (qPCR) assay. We observed extensive variation in *AMY1* copy number in this population sample (Fig. 1a and Supplementary Table 1 online), consistent with previous studies^{10,11}. Next, we performed protein blot experiments with saliva samples from the same individuals in order to estimate salivary amylase protein levels (Fig. 1b). These experiments showed a significant positive correlation between salivary amylase gene copy number and protein expression ($P < 0.001$; Fig. 1c).

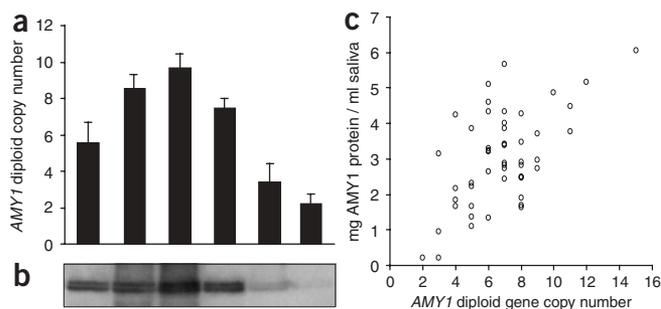


Figure 1 *AMY1* copy number variation and salivary amylase protein expression. (a,b) For the same European American individuals, we estimated diploid *AMY1* gene copy number with qPCR (a) and estimated amylase protein levels in saliva by protein blot (b). Error bars indicate s.d. (c) Relationship between *AMY1* diploid copy number and salivary amylase protein level ($n = 50$ European Americans). A considerable amount of variation in *AMY1* protein expression is not explained by copy number ($R^2 = 0.351$), which may reflect other genetic influences on *AMY1* expression, such as regulatory region SNPs or nongenetic factors that may include individual hydration status, stress level and short-term dietary habits.

¹School of Human Evolution and Social Change, Arizona State University, Tempe, Arizona 85287, USA. ²Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. ³Department of Anthropology, University of California, Santa Cruz, California 95064, USA. ⁴School of Life Sciences, Arizona State University, Tempe, Arizona 85287, USA. ⁵The Wellcome Trust Sanger Institute, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK. ⁶Department of Anthropological Sciences, Stanford University, Stanford, California 94305, USA. ⁷Harvard Medical School, Boston, Massachusetts 02115, USA. ⁸These authors contributed equally to this work. Correspondence should be addressed to N.J.D. (njdominy@ucsc.edu).

Received 9 May; accepted 3 August; published online 9 September 2007; doi:10.1038/ng2123

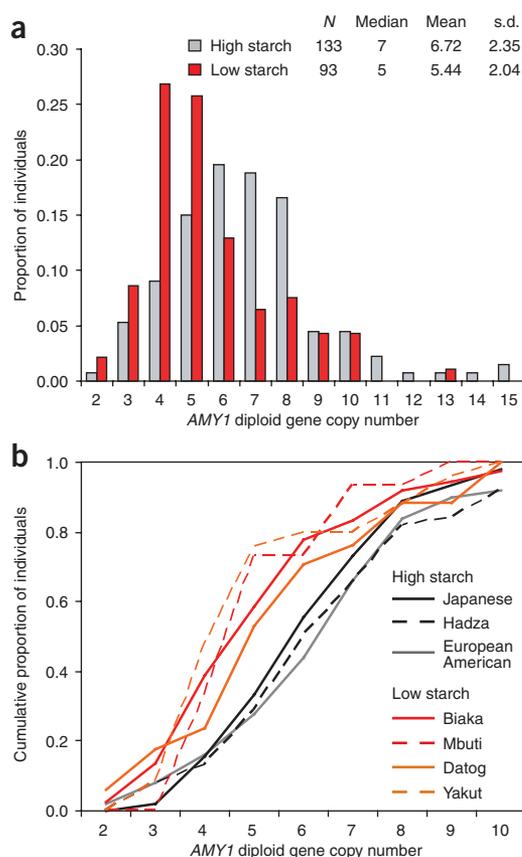


Figure 2 Diet and *AMY1* copy number variation. (a) Comparison of qPCR-estimated *AMY1* diploid copy number frequency distributions for populations with traditional diets that incorporate many starch-rich foods (high-starch) and populations with traditional diets that include little or no starch (low-starch). (b) Cumulative distribution plot of diploid *AMY1* copy number for each of the seven populations in the study.

Supplementary Fig. 1 online). Notably, the proportion of individuals from the combined high-starch sample with at least six *AMY1* copies (70%) was nearly two times greater than that for low-starch populations (37%). To visualize the allele-specific number and orientation of *AMY1* gene copies, we performed high-resolution FISH on stretched DNA fibers (fiber FISH); these results were consistent with diploid *AMY1* copy number estimates from our qPCR experiments (**Fig. 3**).

The among-population patterns of *AMY1* copy number variation do not fit expectations under a simple geographical region-based model of genetic drift: our high- and low-starch samples include both African and Asian populations, suggesting that diet more strongly predicts *AMY1* copy number than geographic proximity. Based on this observation, we hypothesized that natural selection may have influenced *AMY1* copy number in certain human populations. However, we cannot rigorously test such a hypothesis on the basis of our qPCR results alone, in part because we lack comparative data from other loci. Therefore, we next performed array-based comparative genomic hybridization (aCGH) on the Yakut population sample with a whole-genome tile path (WGTP) array platform that was previously used¹¹ to describe genome-wide patterns of copy number variation in 270 individuals (the HapMap collection), including the same Japanese population sample as in our study. For the Yakut aCGH experiments, we used the same reference DNA sample (NA10851) as in the previous study¹¹, facilitating comparisons of Japanese and Yakut relative intensity \log_2 ratios for the 26,574 BAC clones on the array, including two clones mapped to the *AMY1* locus.

Although there is a considerable range of variation in dietary starch intake among human populations, a distinction can be made between 'high-starch' populations for which starchy food resources comprise a substantial portion of the diet and the small fraction of 'low-starch' populations with traditional diets that incorporate relatively few starchy foods. Such diets instead emphasize proteinaceous resources (for example, meats and blood) and simple saccharides (for example, from fruit, honey and milk). To determine if *AMY1* copy number differs among populations with high- and low-starch diets, we estimated *AMY1* copy number in three high-starch and four low-starch population samples. Our high-starch sample included two agricultural populations, European Americans ($n = 50$) and Japanese ($n = 45$), and Hadza hunter-gatherers who rely extensively on starch-rich roots and tubers ($n = 38$)¹². Low-starch populations included Biaka ($n = 36$) and Mbuti ($n = 15$) rainforest hunter-gatherers, Datog pastoralists ($n = 17$) and the Yakut, a pastoralist, fishing society ($n = 25$). Additional details on the diets of these populations are provided in **Supplementary Table 2** online. We found that mean diploid *AMY1* copy number was greater in high-starch populations (**Fig. 2** and

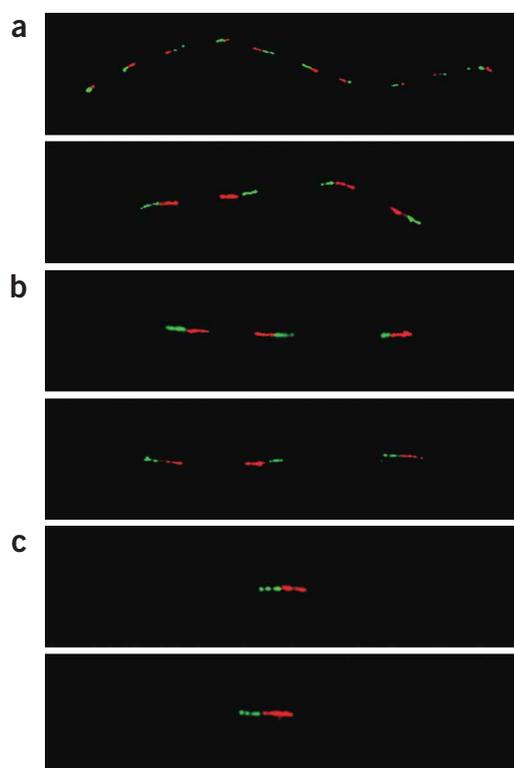


Figure 3 High-resolution fiber FISH validation of *AMY1* copy number estimates. Red (~10 kb) and green (~8 kb) probes encompass the entire *AMY1* gene and a retrotransposon directly upstream of (and unique to) *AMY1*, respectively. (a) Japanese individual GM18972 was estimated by qPCR to have 14 (13.73 ± 0.93) diploid *AMY1* gene copies, consistent with fiber FISH results showing one allele with ten copies and the other with four copies. (b) Biaka individual GM10472 was estimated by qPCR to have six (6.11 ± 0.17) diploid *AMY1* gene copies, consistent with fiber FISH results. (c) The reference chimpanzee (Clint; S006006) was confirmed to have two diploid *AMY1* gene copies.

Results from the two *AMY1*-mapped clones on the WGTP array supported our original observations: the \log_2 ratios were strongly correlated with the qPCR estimates of *AMY1* diploid copy number (Supplementary Fig. 1), and the population mean \log_2 ratios for both clones were greater for the Japanese sample (Fig. 4a and Supplementary Fig. 1). More importantly, with the WGTP data, we were able to compare the extent of population differentiation at the *AMY1* locus to other loci in the genome for the two Asian population samples in our study. We would expect the magnitude and direction of the Japanese-Yakut mean \log_2 ratio difference for the *AMY1*-mapped clones to be similar to those for other copy number-variable clones, if these CNVs have experienced similar evolutionary pressures. However, the two *AMY1*-mapped clones are significant outliers in this distribution (Fig. 4b and Supplementary Fig. 2 online), leading us to reject this null hypothesis. In addition, we considered a database of genotypes for 783 genome-wide microsatellites for the same Yakut individuals and a different Japanese population sample¹³, because microsatellite loci are usually multiallelic (as is the *AMY1* locus). We found that the extent of Japanese-Yakut differentiation at the *AMY1* locus exceeded that for >97% of the microsatellite loci (Supplementary Fig. 3 online). Although this result should be interpreted with caution because we do not know whether *AMY1* copy number and microsatellite mutation rates and patterns are similar, this finding is consistent with our results from the WGTP comparison.

These observations suggest that natural selection has shaped *AMY1* copy number variation in either the Japanese or the Yakut or in both populations. We cannot fully test the null hypothesis for the other high- and low-starch populations in our study, but the patterns of copy number variation we observed in these populations are similar to those for the Japanese and Yakut and thus may also reflect non-neutral evolution. We favor a model in which *AMY1* copy number has been subject to positive or directional selection in at least some high-starch populations but has evolved neutrally (that is, through genetic drift) in low-starch populations. Although it is possible that lower *AMY1* gene copy numbers have been favored by selection in low-starch populations, such an interpretation is less plausible for the simple reason that excessive amylase production is unlikely to have a significant negative effect on fitness. Furthermore, several lines of evidence offer mechanisms by which higher salivary amylase protein expression may confer a fitness advantage for individuals with a high-starch diet. First, a significant amount of starch digestion occurs in the mouth during mastication¹⁴. For example, blood glucose has been shown to be significantly higher when high-starch foods such as corn, rice and potatoes (but not low-starch foods such as apples) are first chewed and then swallowed, rather than swallowed directly¹⁵. In addition, it has been suggested that oral digestion of starch is critically important for energy absorption during episodes of diarrhea⁴. Diarrheal diseases can have a significant effect on fitness; for example, such diseases caused 15% of worldwide deaths among children younger than 5 years as recently as 2001 (ref. 16). Last, salivary amylase persists in the stomach and intestines after swallowing¹⁷, thereby augmenting the enzymatic activity of pancreatic amylase in the small intestine. Higher *AMY1* copy number and a concomitant increase in salivary amylase protein thus are likely to improve the efficiency with which high-starch foods are digested in the mouth, stomach and intestines and may also buffer against the potential fitness-reducing effects of intestinal disease.

To understand better the evolutionary context of human *AMY1* copy number variation, we analyzed patterns of *AMY1* copy number variation in chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*). In contrast to the extensive copy number variation

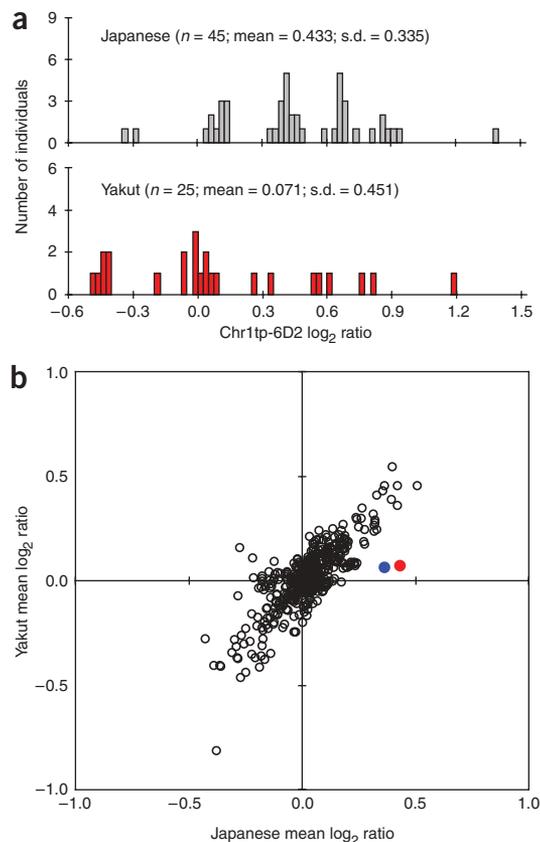


Figure 4 Japanese-Yakut copy number differentiation at *AMY1* versus other genome-wide loci. (a) Frequency distributions of WGTP aCGH relative intensity \log_2 ratios from *AMY1*-mapped clone Chr1tp-6D2 for Japanese and Yakut individuals. (b) Relationship between Japanese and Yakut mean \log_2 ratios for all autosomal WGTP clones that were copy number variable in both populations. *AMY1*-mapped clones Chr1tp-6D2 and Chr1tp-30C7 are depicted as filled red and blue circles, respectively.

we observed in humans, each of 15 wild-born western chimpanzees (*P. t. verus*) showed evidence of only two diploid *AMY1* copies (Fig. 3c and Supplementary Fig. 4 online), consistent with previous findings^{18–21}. Although we observed evidence of a gain in *AMY1* copy number in bonobos relative to chimpanzees (Supplementary Fig. 4), our sequence-based analyses suggest that each of these *AMY1* copies has a disrupted coding sequence and may be nonfunctional (Supplementary Fig. 5 online). Therefore, the average human has roughly three times more *AMY1* copies than chimpanzees, and bonobos may not have salivary amylase at all. Outgroup comparisons with other great apes suggest that *AMY1* copy number was probably gained in the human lineage, rather than lost in chimpanzees^{21,22}. Given that *AMY1* copy number is positively correlated with salivary amylase protein expression in humans, it stands to reason that the human-specific increase in copy number may explain, at least in part, why salivary amylase protein levels are approximately six to eight times higher in humans than in chimpanzees²³. These patterns are consistent with the general dietary characteristics of *Pan* and *Homo*; chimpanzees and bonobos are predominantly frugivorous and ingest little starch relative to most human populations²⁴.

Among other primates, New World monkeys do not produce salivary amylase and tend to consume little starch, but cercopithecines (a subfamily of Old World monkeys including macaques and

mangabeys) have relatively high salivary amylase expression, even compared to humans²³. Although the genetic mechanisms are unknown, this expression pattern may have evolved to facilitate the digestion of starchy foods (such as the seeds of unripe fruits) stowed in the cheek pouch, a trait that, among primates, is unique to cercopithecines²⁵.

The initial human-specific increase in *AMY1* copy number may have been coincident with a dietary shift early in hominin evolutionary history. For example, it is hypothesized that starch-rich plant underground storage organs (USOs) such as bulbs, corms and tubers were a critical food resource for early hominins^{26,27}. Changes in USO consumption may even have facilitated the initial emergence and spread of *Homo erectus* out of Africa^{5,28}. Yet such arguments are difficult to test, mainly because direct evidence for the use of USOs is difficult to obtain, particularly for more remote time periods. USOs themselves are perishable, as are many of the tools used to collect and process them. Therefore, understanding the timing and nature of the initial human-lineage *AMY1* duplications may provide insight into our ecological and evolutionary history. The low amount of nucleotide sequence divergence among the three *AMY1* gene copies found in the human genome reference sequence (hg18; $d = 0.00011$ to 0.00056) implies a relatively recent origin that may be within the time frame of modern human origins (that is, within the last $\sim 200,000$ years; based on human-chimpanzee *AMY1* $d = 0.027$ and using an estimate of 6 million years ago for divergence of the human and chimpanzee lineages). However, given the possibility for gene conversion, we do not necessarily consider this estimate to be reliable. The generation of *AMY1* sequences from multiple humans may ultimately help to shed light on this issue.

In summary, we have shown that the pattern of variation in copy number of the human *AMY1* gene is consistent with a history of diet-related selection pressures, demonstrating the importance of starchy foods in human evolution. Although the amylase locus is one of the most variable in the human genome with regard to copy number¹⁰, it is by no means unique; a recent genome-wide survey identified 1,447 copy number-variable regions among 270 phenotypically normal human individuals¹¹, and many more such regions are likely to be discovered with advances in copy number variation detection technology. It is reasonable to speculate that copy number variants other than *AMY1* are or have been subject to strong pressures of natural selection, particularly given their potential influence on transcriptional and translational levels (for example, see ref. 29). The characterization of copy number variation among humans and between humans and other primates promises considerable insight into our evolutionary history.

METHODS

Samples. Buccal swabs and saliva were collected under informed consent from 50 European Americans ages 18–30 (Arizona State University institutional review board (IRB) protocol number 0503002355). Saliva was collected for 3 min from under the tongue. Buccal swabs were collected from the Hadza ($n = 38$) and Datog ($n = 17$) from Tanzania (Stanford University IRB protocol number 9798-414). Genomic DNA samples from the Biaka (Central African Republic; $n = 32$), Mbuti (Democratic Republic of Congo; $n = 15$) and Yakut (Siberia; $n = 25$) are from the HGDP-CEPH Human Genome Diversity Cell Line Panel. Lymphoblastoid cell lines from 45 Japanese, 4 additional Biaka and the donor for the chimpanzee genome sequence (Clint) were obtained from the Coriell Institute for Medical Research. Whole blood samples were collected during routine veterinary examinations from chimpanzees and bonobos housed at various zoological and research facilities. Two additional bonobo samples were obtained from the Integrated Primate Biomaterials and Information Resource. DNA was isolated using standard methods.

Copy number estimation. Primers for qPCR (Supplementary Table 3 online) were designed to be specific to *AMY1* (that is, to have sequence mismatches with *AMY2A* and *AMY2B*) based on the human and chimpanzee reference genome sequences. A previous study reported a single (haploid) copy of *AMY1* for one chimpanzee¹⁸, and a recent analysis¹⁹ did not find any evidence of recent *AMY1* duplication for Clint. We used fiber FISH to confirm that Clint has two diploid copies of *AMY1* (Fig. 3c). Therefore, we were able to estimate diploid copy number based on relative *AMY1* quantity for human DNA compared to a standard curve constructed from the DNA of Clint. A fragment from the *TP53* gene was also amplified to adjust for DNA dilution quantity variation. Samples were run in triplicate and standards in duplicate. Experiments were performed and analyzed as described previously²⁰.

Protein blot analysis. Protein samples were prepared by solubilizing saliva samples in 2% SDS and heating at 100 °C for 5 min. These samples were analyzed on mini SDS-polyacrylamide gels and transferred to polyvinylidene difluoride (PVDF) membranes (Immobilon-Millipore). For quantification purposes, a human salivary amylase protein sample of known quantity (Sigma) was run on each gel, with 5 μ L of saliva for each sample. After transfer, the membranes were incubated for 1.5 h with primary antibodies raised against human salivary amylase (Sigma). The membranes were washed and goat anti-rabbit alkaline phosphatase-conjugated IgG secondary antibodies (Pierce) were added for 1 h. The membranes were exposed to ECF substrate (Amersham Biosciences) for 5 min and then analyzed using a phosphorimager. Quantification of protein bands was performed using ImageQuant software (Molecular Dynamics).

Fiber FISH. DNA fibers were prepared by gently lysing cultured lymphoblast cells with 300 μ L Cell Lysis Buffer (Gentra Systems) per 5 million cells. 10 μ L of lysate was placed on a poly-L-lysine-coated slide (LabScientific) and mechanically stretched with the edge of a coverslip. After 30 s, 300 μ L of 100% methanol was applied to fix the fibers. Slides were dried at 37 °C for 5 min and then stored at room temperature (22–25 °C).

PCR product probes were made from (i) the entire *AMY1* gene itself (~ 10 kb; red in images) and (ii) the retrotransposon found directly upstream of all *AMY1* copies but not of pancreatic amylase genes or amylase pseudogenes (~ 8 kb; green in images). The gene probe may not be specific to *AMY1* under all hybridization conditions (*AMY1* sequence divergence with *AMY2A* and *AMY2B* = 7.5% and 7.1%, respectively), but the upstream probe is. We used long-range PCR followed by nested PCR for each region (primers and conditions are provided in Supplementary Table 3). PCR products were purified with DNA Clean and Concentrator columns (Zymo).

For each nested PCR product, 750 ng was combined with 20 μ L 2.5 \times random primer (BioPrime aCGH Labeling Module, Invitrogen) in a total volume of 39 μ L. Samples were incubated at 100 °C for 5 min and were then placed on ice for 5 min. Next, 5 μ L 10 \times dUTP and 1 μ L Exo-Klenow Fragment (BioPrime Module) and either 5 μ L (5 nmol) Biotin-16-dUTP (Roche; gene probe) or 5 μ L (5 nmol) digoxigenin-11-dUTP (Roche; upstream probe) were added, and samples were incubated at 37 °C for 5 h. Labeled products were purified with Microcon Centrifugal Filter Devices (Millipore) using three washes of 300 μ L 0.1 \times SSC, eluted with 50 μ L H₂O. For each 1 μ g of labeled DNA, we added 10 μ g human Cot-1 DNA (Invitrogen).

For each experiment, 500 ng of labeled DNA from each of the nested PCR reactions was combined, lyophilized, reconstituted in 10 μ L hybridization buffer (50% formamide, 20% dextran sulfate, 2 \times SSC) and added to the slide (18 \times 18 mm cover glass; Fisher). Fibers and probes were denatured together (95 °C for 3 min) and hybridized in a humidified chamber (37 °C for 40 h). The slide was washed in 0.5 \times SSC at 75 °C for 5 min followed by three washes in 1 \times PBS at room temperature (22–25 °C) for 2 min each. Next, fibers were incubated with 200 μ L CAS Block (Zymed) and 10% (vol/vol) normal goat serum (Zymed) for 20 min at room temperature (22–25 °C) under a HybriSlip (Invitrogen). We used a three-step detection and amplification (with reagents in 200 μ L CAS Block with 10% (vol/vol) normal goat serum). Each step was conducted for 30 min at room temperature under a HybriSlip followed by three washes in 1 \times PBS for 2 min each at room temperature (22–25 °C). Reagents were as follows for each step: step (i): 1:500 anti-digoxigenin-fluorescein, Fab fragments (Roche) and 1:500 Streptavidin, Alexa Fluor 594

conjugate (Invitrogen); step (ii): 1:250 rabbit anti-FITC (Zymed) and 1:500 biotinylated anti-streptavidin (Vector Laboratories); step (iii): 1:100 goat anti-rabbit IgG-FITC (Zymed) and 1:500 streptavidin, Alexa Fluor 594 conjugate. Images were captured on an Olympus BX51 fluorescent microscope with an Applied Imaging camera and were analyzed with Applied Imaging's Genus software.

aCGH analysis. For aCGH experiments, we used a large-insert clone DNA microarray covering the human genome in tiling path resolution³⁰. Test genomic DNA samples (from Yakut individuals) and reference genomic DNA samples (from NA10851) were labeled with Cy3-dCTP and Cy5-dCTP, respectively (NEN Life Science Products) and were cohybridized to the array. For each sample, a duplicate experiment was performed in dye-swap to reduce false-positive error rates. Labeling, hybridization, washes and analyses were performed as described^{11,30}.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We are grateful to all our study participants. We thank H. Cann and C. de Toma of the Fondation Jean Dausset (CEPH), the Cincinnati Zoo, the Lincoln Park Zoo, the New Iberia Research Center, the Primate Foundation of Arizona, the Southwest Foundation for Biomedical Research, the Coriell Institute for Medical Research and the Integrated Primate Biomaterials and Information Resource for samples. C. Tyler-Smith and Y. Gilad provided comments on a previous version of the manuscript. We would also like to thank the Wellcome Trust Sanger Institute Microarray Facility for printing the arrays and T. Fitzgerald and D. Rajan for technical support. This study was funded by grants from the L.S.B. Leakey Foundation and Wenner-Gren Foundation (to N.J.D.), the Department of Pathology, Brigham & Women's Hospital (to C.L.), the National Institutes of Health (to the University of Louisiana at Lafayette New Iberia Research Center; numbers RR015087, RR014491 and RR016483) and the Wellcome Trust (H.F., R.R. and N.P.C.).

AUTHOR CONTRIBUTIONS

G.H.P. and N.J.D. contributed equally to this work. G.H.P., N.J.D., C.L. and A.C.S. designed the study; G.H.P., F.A.V., J.L.M. and A.C.S. collected the samples; G.H.P. and A.S.L. performed qPCR experiments; J.W. performed protein blot experiments; G.H.P. performed fiber FISH experiments; H.F. and R.R. performed and analyzed aCGH experiments; K.G.C. performed nucleotide sequencing experiments; G.H.P. performed data analyses; R.M., N.P.C., C.L. and A.C.S. supervised the experiments and analyses and G.H.P. and N.J.D. wrote the paper.

Published online at <http://www.nature.com/naturegenetics>

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions>

1. Tokarev, S.A. & Gurvich, I.S. The Yakuts. in *Peoples of Siberia* (eds. Levin, M.G. & Potapov, L.P.) 243–304 (University of Chicago Press, Chicago, 1956).
2. Draper, H.H. The aboriginal Eskimo diet in modern perspective. *Am. Anthropol.* **79**, 309–316 (1977).
3. Hart, T.B. & Hart, J.A. The ecological basis of hunter-gatherer subsistence in African rain forests: the Mbuti of Eastern Zaire. *Hum. Ecol.* **14**, 29–55 (1986).

4. Leubenthal, E. Role of salivary amylase in gastric and intestinal digestion of starch. *Dig. Dis. Sci.* **32**, 1155–1157 (1987).
5. Wrangham, R.W., Jones, J.H., Laden, G., Pilbeam, D. & Conklin-Brittain, N. The raw and the stolen. Cooking and the ecology of human origins. *Curr. Anthropol.* **40**, 567–594 (1999).
6. Klein, R.G. Archeology and the evolution of human behavior. *Evol. Anthropol.* **9**, 17–36 (2000).
7. Diamond, J. Evolution, consequences and future of plant and animal domestication. *Nature* **418**, 700–707 (2002).
8. Zohary, D. & Hopf, M. *Domestication of Plants in the Old World* (Oxford Science Publications, Oxford, 2000).
9. Groot, P.C. *et al.* The human alpha-amylase multigene family consists of haplotypes with variable numbers of genes. *Genomics* **5**, 29–42 (1989).
10. Iafrate, A.J. *et al.* Detection of large-scale variation in the human genome. *Nat. Genet.* **36**, 949–951 (2004).
11. Redon, R. *et al.* Global variation in copy number in the human genome. *Nature* **444**, 444–454 (2006).
12. Vincent, A.S. Plant foods in savanna environments: a preliminary report of tubers eaten by the Hadza of northern Tanzania. *World Archaeol.* **17**, 131–148 (1984).
13. Rosenberg, N.A. *et al.* Clines, clusters, and the effect of study design on the inference of human population structure. *PLoS Genet.* **1**, e70 (2005)(doi:10.1371/journal.pgen.0010070).
14. Valdez, I.H. & Fox, P.C. Interactions of the salivary and gastrointestinal systems. I. The role of saliva in digestion. *Dig. Dis.* **9**, 125–132 (1991).
15. Read, N.W. *et al.* Swallowing food without chewing: a simple way to reduce post-prandial glycaemia. *Br. J. Nutr.* **55**, 43–47 (1986).
16. Lopez, A.D., Mathers, C.D., Ezzati, M., Jamison, D.T. & Murray, C.J. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* **367**, 1747–1757 (2006).
17. Fried, M., Abramson, S. & Meyer, J.H. Passage of salivary amylase through the stomach in humans. *Dig. Dis. Sci.* **32**, 1097–1103 (1987).
18. Samuelson, L.C., Wiebauer, K., Snow, C.M. & Meisler, M.H. Retroviral and pseudogene insertion sites reveal the lineage of human salivary and pancreatic amylase genes from a single gene during primate evolution. *Mol. Cell. Biol.* **10**, 2513–2520 (1990).
19. Cheng, Z. *et al.* A genome-wide comparison of recent chimpanzee and human segmental duplications. *Nature* **437**, 88–93 (2005).
20. Perry, G.H. *et al.* Hotspots for copy number variation in chimpanzees and humans. *Proc. Natl. Acad. Sci. USA* **103**, 8006–8011 (2006).
21. Wilson, G.M. *et al.* Identification by full-coverage array CGH of human DNA copy number increases relative to chimpanzee and gorilla. *Genome Res.* **16**, 173–181 (2006).
22. Fortna, A. *et al.* Lineage-specific gene duplication and loss in human and great ape evolution. *PLoS Biol.* **2**, e207 (2004)(doi:10.1371/journal.pbio.0020207).
23. McGeachin, R.L. & Akin, J.R. Amylase levels in the tissues and body fluids of several primate species. *Comp. Biochem. Physiol. A* **72**, 267–269 (1982).
24. Hohmann, G., Robbins, M. & Boesch, C. *Feeding Ecology in Apes and Other Primates* (Cambridge Univ. Press, Cambridge, 2006).
25. Lambert, J.E. Competition, predation, and the evolutionary significance of the cercopithecine cheek pouch: the case of *Cercopithecus* and *Lophocebus*. *Am. J. Phys. Anthropol.* **126**, 183–192 (2005).
26. Coursey, D.G. Hominid evolution and hypogeous plant foods. *Man* **8**, 634–635 (1973).
27. Laden, G. & Wrangham, R. The rise of the hominids as an adaptive shift in fallback foods: plant underground storage organs (USOs) and australopithecine origins. *J. Hum. Evol.* **49**, 482–498 (2005).
28. Hawkes, K., O'Connell, J.F., Jones, N.G., Alvarez, H. & Charnov, E.L. Grandmothering, menopause, and the evolution of human life histories. *Proc. Natl. Acad. Sci. USA* **95**, 1336–1339 (1998).
29. Stranger, B.E. *et al.* Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* **315**, 848–853 (2007).
30. Fiegler, H. *et al.* Accurate and reliable high-throughput detection of copy number variation in the human genome. *Genome Res.* **16**, 1566–1574 (2006).