HHSN272201400008C for use on influenza virus. L.C.K. was supported by the Gates Cambridge Scholarship and the NIH Oxford Cambridge Scholars Program. J.M.F. was supported by a Medical Research Council Fellowship (MR/K021885/1) and a Junior Research Fellowship from Homerton College Cambridge. E.C.H. was supported by an National Health and Medical Research Council Australia Fellowship. N.V. and R.B.T were supported by NIH contract HHSN272201000040I/HHSN27200004/D04. The viruses and sera used in this study are covered by standard material transfer agreements at the home institutions of S.S.W., C.P.S., E.H., P.B., J.G.A., J.L.M.J., N.V., and R.B.T. A.D.M.E.O. is a professor and director of Artemis One Health Utrecht, The Netherlands; Chief Scientific Officer Viroclinics Biosciences BV, the Netherlands; and a Board Member of Protein Sciences USA. P.B. performed this work while at the Institut Pasteur in Cambodia, but since June 2015, is with GlaxoSmithKline vaccines in Singapore, and has stock options with GSK. C.P.S. is a paid consultant to GSK Pharma, GSK Vaccines, and Merck and has received a grant and consulting payments to his institution from Sanofi Pasteur. The sequences used in this study are available from GenBank (http://www.ncbi.nlm.nih.gov/genbank/) and are listed in table S1. Files used for genetic analyses are available as supplementary data files. The NIH monovalent DENV vaccines trials (ClinicalTrials.gov identifiers: NCT00473135 NCT00920517, NCT00831012, NCT00831012) were performed under an investigational new drug application reviewed by the U.S. Food and Drug Administration and approved by the Institutional Review Board at the University of Vermont and Johns Hopkins University. Informed consent was obtained in accordance federal and international regulations (21CFR50, ICHE6). The Pediatric Dengue Cohort Study in Managua, Nicaragua, was approved by the Institutional Review Boards of the Nicaraguan Ministry of Health and the University of California, Berkeley. Parents or legal guardians of all subjects provided written informed consent, and subjects 6 years of age and older provided assent.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/349/6254/1338/suppl/DC1
Materials and Methods
Figs. S1 to S29
Tables S1 to S15
Data Files S1 to S9
References (*43*–*66*)

6 May 2015; accepted 6 August 2015
10.1126/science.aac5017

HUMAN GENETICS

Greenlandic Inuit show genetic signatures of diet and climate adaptation

Matteo Fumagalli,[1,2]* Ida Moltke,[3]* Niels Grarup,[4] Fernando Racimo,[2] Peter Bjerregaard,[5,6] Marit E. Jørgensen,[5,7] Thorfinn S. Korneliussen,[8] Pascale Gerbault,[1,9] Line Skotte,[3] Allan Linneberg,[10,11,12] Cramer Christensen,[13] Ivan Brandslund,[14,15] Torben Jørgensen,[10,16,17] Emilia Huerta-Sánchez,[18] Erik B. Schmidt,[17,19] Oluf Pedersen,[4] Torben Hansen,[4]† Anders Albrechtsen,[3]† Rasmus Nielsen[2,20]†

The indigenous people of Greenland, the Inuit, have lived for a long time in the extreme conditions of the Arctic, including low annual temperatures, and with a specialized diet rich in protein and fatty acids, particularly omega-3 polyunsaturated fatty acids (PUFAs). A scan of Inuit genomes for signatures of adaptation revealed signals at several loci, with the strongest signal located in a cluster of fatty acid desaturases that determine PUFA levels. The selected alleles are associated with multiple metabolic and anthropometric phenotypes and have large effect sizes for weight and height, with the effect on height replicated in Europeans. By analyzing membrane lipids, we found that the selected alleles modulate fatty acid composition, which may affect the regulation of growth hormones. Thus, the Inuit have genetic and physiological adaptations to a diet rich in PUFAs.

Previous studies have attempted to understand the genetic basis of human adaptation to local environments, including cold climates and a lipid-rich diet (*1*). A recent study found evidence that a coding variant in *CPT1A*, a gene involved in the regulation of long-chain fatty acid, has been the target of strong positive selection in native Siberians, possibly driven by adaptation to a cold climate or to a high-fat diet (*2*). Another study found evidence that adaptation to the traditional hypoglycemic diet of Greenlandic Inuit may have favored a mutation in *TBC1D4* that affects glucose uptake and occurs at high frequency only among the Inuit (*3*). However, knowledge about the genetic basis of human adaptation to cold climates and lipid-rich diets remains limited.

Motivated by this, we performed a scan for signatures of genetic adaptation in the population of Greenland. The Inuit ancestors of this population arrived in Greenland less than 1000 years ago (*4*), but they lived in the Arctic for thousands of years before that (*5*). As such, they have probably adapted to the cold Arctic climate and to their traditional diet, which has a high content of omega-3 polyunsaturated fatty acids (PUFAs) derived from seafood (*6*) and a content of omega-6 PUFAs that is lower than in Danish controls (*7*).

We analyzed data from previously genotyped Greenlandic individuals (*3*) by using the Illumina MetaboChip (*8*), which is an array enriched with single-nucleotide polymorphisms (SNPs) identified in genome-wide association studies (GWASs) associated with cardiometabolic phenotypes. As a result of recent admixture, modern Greenlanders have, on average, 25% genetic European ancestry (*9*). To get a representative sample of the indigenous Greenlandic Inuit (GI), we analyzed the subset of 191 individuals that had less than 5% estimated European ancestry per individual (0.5% on average) (*9*). We combined the data from these individuals with the MetaboChip data from 60 individuals of European ancestry (CEU) and 44 Han Chinese individuals (CHB) from the HapMap Consortium (fig. S1) (*10*).

To detect signals of positive selection, we used the population branch statistic (PBS) (*11*), which identifies alleles that have experienced strong changes in frequency in one population (GI) relative to two reference populations (CEU and CHB) (*5*). A sliding window analysis identified several SNP windows with high PBS values, indicative of selection (Fig. 1 and table S1).

The strongest signal of selection is located within a region on chromosome 11 (Fig. 1A) and encompasses five genes: two open reading frames, C11orf10 (*TMEM258*) and C11orf9 (*MYRF*); and three fatty acid desaturases, *FADS1*, *FADS2*, and *FADS3*. The SNP with the highest PBS value falls within *FADS2*. The function of *FADS3* is not known; *FADS1* and *FADS2* encode delta-5 and delta-6 desaturases, which are the rate-limiting steps in the conversion of linoleic acid (omega-6) and α-linolenic acid (omega-3) to the longer, more unsaturated and biologically active eicosapentaenoic acid (EPA, omega-3), docosahexaenoic acid (DHA, omega-3), and arachidonic acid (omega-6).

[1]Department of Genetics, Evolution, and Environment, University College London, London WC1E 6BT, UK. [2]Department of Integrative Biology, University of California–Berkeley, Berkeley, CA 94720, USA. [3]The Bioinformatics Centre, Department of Biology, University of Copenhagen, 2200 Copenhagen, Denmark. [4]The Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, Faculty of Health and Medical Sciences, University of Copenhagen, 2100 Copenhagen, Denmark. [5]National Institute of Public Health, University of Southern Denmark, 1353 Copenhagen, Denmark. [6]Greenland Center for Health Research, University of Greenland, Nuuk, Greenland. [7]Steno Diabetes Center, 2820 Gentofte, Denmark. [8]Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, 1350 Copenhagen, Denmark. [9]Department of Anthropology, University College London, London WC1H 0BW, UK. [10]Research Centre for Prevention and Health, Capital Region of Denmark, Copenhagen, Denmark. [11]Department of Clinical Experimental Research, Rigshospitalet, Glostrup, Denmark. [12]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [13]Department of Medicine, Lillebaelt Hospital, Vejle, Denmark. [14]Department of Clinical Biochemistry, Lillebaelt Hospital, Vejle, Denmark. [15]Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark. [16]Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [17]Faculty of Medicine, University of Aalborg, Aalborg, Denmark. [18]School of Natural Sciences, University of California–Merced, Merced, CA 95343, USA. [19]Department of Cardiology, Aalborg University Hospital, 9100 Aalborg, Denmark. [20]Department of Statistics, University of California–Berkeley, Berkeley, CA 94720, USA.
*These authors contributed equally to this work. **†Corresponding author. E-mail: torben.hansen@sund.ku.dk (T.H.); albrecht@binf.ku.dk (A.A.); rasmus_nielsen@berkeley.edu (R.N.)**

Polymorphisms in *FADS1* and *FADS2* are associated with increased levels of plasma and erythrocyte delta-5 desaturases in Alaskan Inuit (*12*) as well as with levels of PUFA in blood and breast milk (*13*, *14*).

We also found signals of selection in a region on chromosome 1 (Fig. 1A), which encompasses *WARS2*, a mitochondrial tryptophanyl-tRNA synthetase, and *TBX15*, a transcription factor member of the T-box family. Within this region, the SNP with the highest PBS value is located upstream of *WARS2*. Polymorphisms in or near *WARS2* and *TBX15* have been shown to be associated with numerous phenotypes among individuals of European descent, including waist-hip ratio (*15*). Based on linkage disequilibrium (LD) patterns in Greenlandic Inuit, the results from (*15*) suggest that the allele that occurs frequently in Greenlandic Inuit may decrease the waist-hip ratio. *TBX15* plays a role in the differentiation of brown (subcutaneous) and brite (typically inguinal) adipocytes (*16*). The latter, upon stimulation by exposure to cold, can differentiate into cells capable of expressing UCP1 (uncoupling protein 1), which produces heat by lipid oxidation. Therefore, *TBX15* may be associated with adaptation to cold in Inuit.

FN3KRP shows evidence of selection as well (Fig. 1A). *FN3KRP* encodes an enzyme that catalyzes fructosamines, psicosamines, and ribulosamines. This protein protects against nonenzymatic glycation, an oxidative process that is associated with various pathophysiologies (*17*). A high intake of PUFAs is associated with increased oxidative stress (*18*); it is possible that the alleles affected by selection in *FN3KRP* counteract the negative fitness caused by a PUFA-rich diet. A list of additional candidate regions under positive selection is presented in tables S2 and S3.

To corroborate our results from the SNP chip–based analysis described above, we also calculated PBS values (table S4) for exome sequencing data from 18 unrelated GI individuals (*3*), combined with data from 85 CEU individuals and 97 CHB individuals from the 1000 Genomes Project (fig. S1) (*19*).

These analyses identified two high-scoring genes (table S5): *DSP*, a gene associated with cardiomyopathy (*20*), and *ANGPTL6*, a gene that counteracts high-fat diet–induced obesity and related insulin resistance through increased energy expenditure (*21*). Gene ontology enrichment analyses of genes under selection revealed enriched muscle- and heart-development categories, similar to those positively selected in polar bears (table S6) (*5*, *22*).

In addition, these analyses reproduced the strong signal observed in the *FADS1-FADS2-FADS3* region, even though the SNPs with the highest PBS values are not detected by the system used for exome capture (Agilent SureSelect; fig. S2), and this region has the SNP with the strongest signal of selection (i.e., highest PBS value) in any of the data analyzed. We therefore focused on this region for the rest of this study. On the basis of an inferred demographic model (*5*), we estimated a divergence time between CHB and GI of 23,250 years before the present (yr B.P.), unidirectional gene flow from GI to CHB at some point in the history of these populations, and a reduced effective population size of GI (effective population size = 1550). The estimated model (fig. S3A) fits the observed joint site frequency spectrum (fig. S4), and the PBS value for the *FADS* region is a strong outlier, corroborating the idea that selection probably has affected this region (fig. S5).

Using an approximate Bayesian computation approach, we also estimated the starting time and intensity of selection, *s* (*5*). Because of the high LD within the region and the fact that our data were from SNP chip (fig. S6), we could not pinpoint the causative SNP(s) by means of population genetic analyses; we therefore used the SNP with the highest PBS value (reference SNP identification number rs74771917) as a proxy. This SNP has a derived allele frequency of 0.98 in GI, 0.025 in CEU, and 0.16 in CHB. Our analyses produced maximum a posteriori probability (MAP) estimates of the selection starting time, 19,751 yr B.P. [95% Bayesian credible interval (BCI): 2499 to 22,771 yr B.P.] (figs. S3B and S7), and of *s*, 3.13% (95% BCI: 0.98 to 19.49%) (fig. S3C). These results suggest that selection began to act on these genes long before the earliest settlement of Inuit in Greenland (*4*). In population samples from the HGDP-CEPH (Human Genome Diversity Project–Centre d'Etude du Polymorphisme Humain) database, the selected allele of rs74771917 has much higher frequencies among Native Americans than it does among East Asians (fig. S8) (*23*), suggesting that selection began to act before the Inuit split from the Native Americans, when their common ancestors lived in or around Beringia (*24*).

Six SNPs in the *FADS* region (Table 1) have PBS values above 2, suggesting that they have been subjected to strong selection. One of these SNPs, rs174570, is associated with circulating high-density lipoprotein, low-density lipoprotein (LDL), and total cholesterol levels in Europeans (*25*). We therefore tested for associations between the top six SNPs and 13 metabolic and anthropometric phenotypes in Greenlanders by analyzing data from the Greenlandic cohorts IHIT (Inuit Health in Transition) and B99 (Greenland Population Study 1999), which include 2733 and 1331 genotyped individuals, respectively (*3*). We analyzed the cohorts separately, combined the results in a meta-analysis (*5*), and found marginally significant associations with multiple phenotypes, including body-mass index, fasting serum insulin, and fasting serum LDL cholesterol (tables S7 to S12). In all cases, the derived (selected)

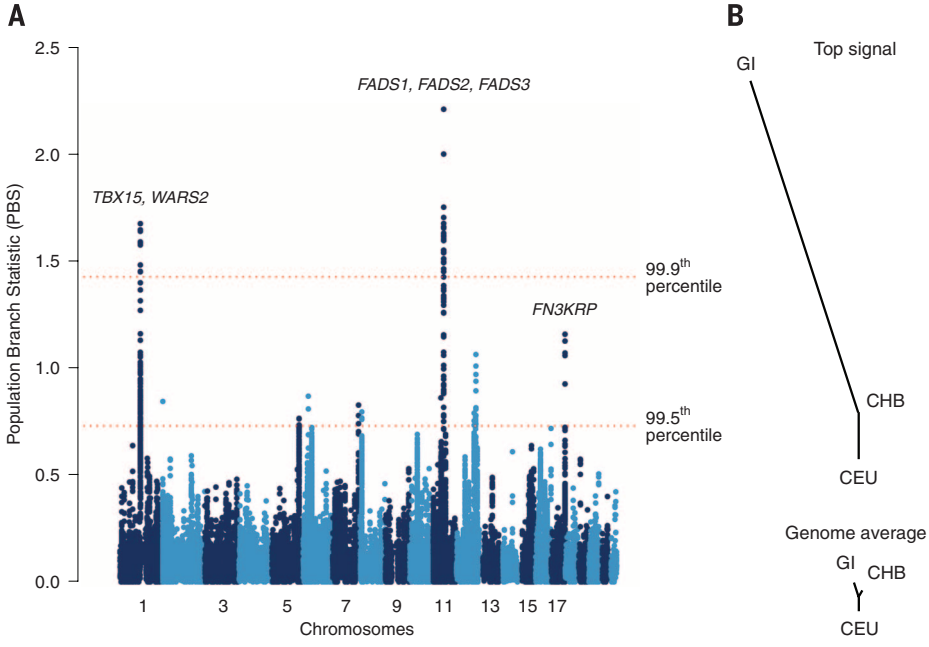

Fig. 1. Results from a genome-wide scan for positive selection. (**A**) PBS values in windows of 20 SNPs, using a step size of 5 SNPs. The 99.5th and 99.9th percentiles of the empirical distribution are shown as red dashed horizontal lines. Names of genes associated with the highest peaks are shown. (**B**) Evolutionary trees underlying the strongest signal of selection. The bottom panel shows genomic-average branch lengths based on F_{ST} (fixation index, a measure of population genetic differentiation) for GI, CEU and CHB branches (bottom); the top panel shows branch lengths for the SNPs in the window with the highest PBS values, indicating substantial changes in allele frequencies along the GI branch.

allele was associated with a reduction in the phenotypic value. The strongest association was with body weight ($P = 1.1 \times 10^{-6}$; rs7115739) and height ($P = 0.00012$; rs7115739) (table S10). Both of these associations remained significant after Bonferroni correction for testing for association between 13 phenotypes and six SNPs. To further validate the association with height, we genotyped an additional Greenlandic cohort, known as BBH, consisting of 541 Greenlandic individuals who live in Denmark and for whom height information is available. When we added these data to the meta-analysis of height, the association signal for rs7115739 became even stronger ($P = 4.6 \times 10^{-7}$). Moreover, the per-allele effect size estimates for the derived allele for height and weight are −0.66 cm and −2.2 kg in IHIT and −1.2 cm and −2.4 kg in B99 (Fig. 2, A and B, and table S10). As mentioned, the statistical method that we used accounts for admixture. Furthermore, we observed an effect both in Greenlanders with little or no European ancestry and in Greenlanders with more than 40% European ancestry when we stratified the data on the basis of ancestry proportions, which we would not expect if the association signal was caused by admixture in our data (fig. S9). These observations indicate that our association results are not caused by insufficient correction for admixture.

The six SNPs with the highest PBS values are also polymorphic in Europeans (Table 1). However,

Table 1. Annotation for the top six SNPs under positive selection in Greenlandic Inuit. DAFs for each population (CEU, CHB, and GI) and PBS values are reported, along with the genomic position for each SNP.

Position*	Reference SNP identification number	Alleles†	DAF			PBS
			CEU	CHB	GI	
chr11:61627960	rs74771917	C/T	0.025	0.16	0.98	2.67
chr11:61631510	rs3168072	A/T	0.017	0.18	0.98	2.64
chr11:61632310	rs12577276	A/G	0.017	0.18	0.98	2.64
chr11:61641717	rs7115739	G/T	0.017	0.22	0.98	2.54
chr11:61624414	rs174602	C/T	0.80	0.73	0.01	2.11
chr11:61597212	rs174570	C/T	0.16	0.34	0.99	2.056

*Positions refer to human genome assembly hg19. †Alleles are coded as ancestral/derived states.

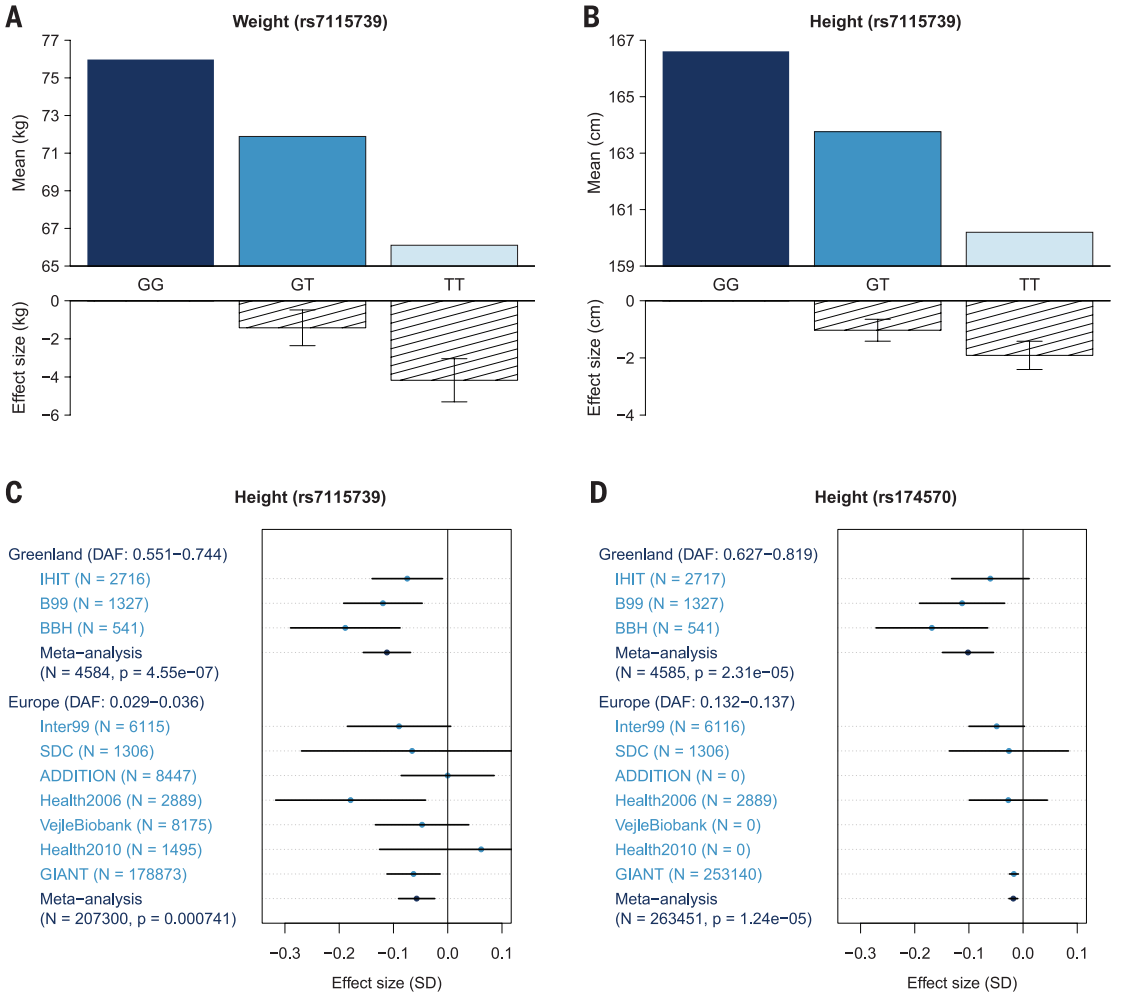

Fig. 2. The effect of rs7115739 and rs174570 on weight and height. (**A**) The effect of rs7115739 on weight in the Greenlandic cohorts IHIT and B99. Shown is the mean value stratified by the genotypes of rs7115739 (top) and the estimated effect of carrying one and two copies of the derived allele, respectively (bottom). The effect-size estimates are adjusted for admixture and other confounding factors and were obtained using a linear mixed model applied to untransformed phenotype measurements. Unlike the estimates in the text and table S10, the estimates shown here were obtained without assuming an additive effect. Error bars, ±1 SE. (**B**) As in (A), but for height. (**C**) Effect sizes for height for the derived allele of rs7115739 in three Greenlandic cohorts and seven European cohorts (SDC, Steno Diabetes Center). Point estimates are shown as points and 95% confidence intervals are shown as horizontal bars. For each of the two geographic regions, the results from a meta-analysis of all the cohorts from the region are also shown. *N* indicates the number of individuals analyzed; DAF, derived allele frequency. For the Greenlandic cohorts, the effect sizes were estimated from height measurements that were quantile-transformed to a standard normal distribution. For the European cohorts, height was analyzed as sex-specific standard (*z*) scores. Hence, the effect sizes from the two geographic regions are not directly comparable. (**D**) As in (C), but for rs174570.

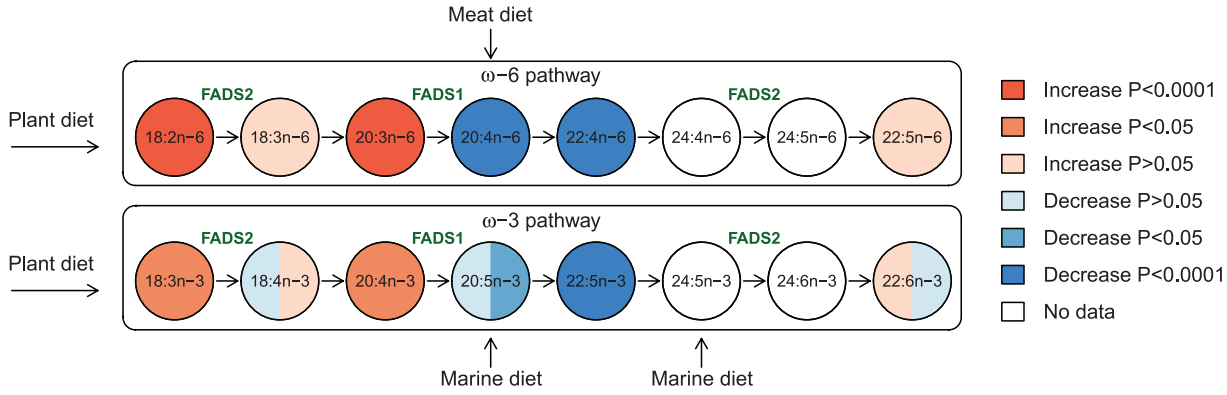

Fig. 3. Results of testing for association between the fatty acids in the omega-3 and omega-6 synthesis pathways and each of two SNPs, rs7115739 and rs174570. The omega-6 (top) and omega-3 (bottom) synthesis pathways are depicted with circles for each fatty acid and arrows for each synthesis step. For each fatty acid, *P* values for tests of association and effect directions for the derived allele are illustrated by the colors on the left (rs7115739) and right (rs174570) halves of the circle. Green text indicates in which of the synthesis steps *FADS1* and *FADS2* play a role. Arrows outside the boxes are simplified indications of where different types of diet enter the two pathways.

because most of the identified SNPs have low allele frequencies in Europeans, they may have been missed by GWAS studies. When combining seven European cohorts, including GIANT (Genetic Investigation of Anthropometric Traits; *26*), we found associations with lower height in carriers of the derived T-allele for rs7115739 (n = 207,300; P = 0.000741) and rs174570 (n = 263,451; $P = 1.24 \times 10^{-5}$) (Fig. 2, C and D, and table S13). The meta-analysis–based effect sizes are equivalent to −0.35 and −0.12 cm for rs7115739 and rs174570, respectively. In contrast, we found no evidence that the six SNPs are associated with weight in Europeans. These results are consistent with results that we obtained when we explicitly tested for differences in effect sizes between Europeans and Greenlandic Inuit (table S14): We found no evidence of a difference in effect size for height for rs7115739 (P = 0.44), but we found significant evidence for a difference in effect size for weight (P = 0.025 and P = 0.012 for rs7115739 and rs174570, respectively), with little or no effect on weight in Europeans. The associations with height in Europeans are unexpected, because this locus was not found to be significant genome-wide in the recent GIANT study of the height of more than 170,000 Europeans (*26*). In addition to the associations with height, we also found known associations with low fasting serum levels of insulin, total cholesterol, and LDL cholesterol for European carriers of low-frequency–derived alleles of *FADS1* variation, suggesting that there may be a protective effect of these variants on cardiometabolic phenotypes (table S13).

To further elucidate the possible functional effects of the alleles of rs7115739 and rs174570, we investigated associations with red blood cell–membrane lipid composition, which reflects fatty-acid intake from the preceding 2 to 4 months and which has previously been measured in IHIT, the largest of our Greenlandic cohorts (*27*). We found significant associations with multiple different fatty acids (fig. S10 and tables S15 and S16). Particularly, we found that the selected alleles are significantly associated with an increase in the concentration of eicosatetraenoic acid (ETA, 20:4n-3) and other omega-3 fatty acids upstream in the omega-3 synthesis pathway, before conversion to EPA (20:5n-3), but a decrease in the concentration of both EPA and omega-3 docosapentaenoic acid (DPA, 22:5n-3), with no significant effect on DHA (22:6n-3) (Fig. 3). These results are consistent with previous observations of linked alleles in Europeans (*28*). The conversion of ETA to EPA is catalyzed by delta-5 desaturases encoded by *FADS1*, and EPA is a major dietary omega-3 fatty acid in the traditional Inuit diet (*18*). Hence, these results suggest that selection affecting the fatty acid desaturases may have compensated for a high dietary intake of EPA.

The changes in the concentration of omega-6 fatty acids mirror those of omega-3 fatty acids (Fig. 3). This might be expected, given that the same enzymes (encoded by *FADS1* and *FADS2*) are involved in both the omega-3 and omega-6 biosynthesis pathways. The similar changes in concentration could therefore be a side effect of selection, driven by a omega-3 PUFA–rich diet. However, selection may also have worked directly on omega-6 fatty acid concentrations early in the ancestral history of Inuit and Native Americans, in the context of a late Paleolithic diet rich in meat from land mammals.

Both rs7115739 and rs174570 show strongly significant associations in conditional analyses where we adjusted for the effects of the other SNP and of rs174602. The remaining three highest-PBS SNPs are in strong LD with rs7115739 in IHIT and would produce similar results. This suggests that there are either multiple causative SNPs or that both rs7115739 and rs174570 are in strong LD with the causal SNP(s).

The challenging environmental conditions of the Arctic have probably imposed strong selective pressures on the Inuit and their ancestors. In all the data that we analyzed, the most pronounced allele-frequency difference between Inuit and other populations was found in a cluster of fatty acid desaturases—*FADS1*, *FADS2*, and *FADS3*—although it is possible that even more extreme differences are present in noncoding regions not covered by our exome data. The *FADS* region has probably been under selection, driven by a diet high in PUFAs. The *FADS* genes have previously been hypothesized to be under selection in other populations in response to dietary changes (*28*, *29*), suggesting that these genes in general play an important role in human adaptation to dietary regimes. Our results also show that genetic variants in fatty acid desaturases have a strong effect on height, probably because of the effect of fatty acid composition and concentration on the regulation of growth hormones (*30*). Previous studies (*31*) have shown that fish oil supplementation is associated with increased concentrations of plasma insulin-like growth factor–1. This study illustrates the utility of evolutionary studies of locally adapted populations for understanding the genetic basis of phenotypic variation among humans.

REFERENCES AND NOTES

1. A. M. Hancock *et al.*, *PLOS Genet.* **7**, e1001375 (2011).
2. F. J. Clemente *et al.*, *Am. J. Hum. Genet.* **95**, 584–589 (2014).
3. I. Moltke *et al.*, *Nature* **512**, 190–193 (2014).
4. H. C. Gulløv, *Grønlands Forhistorie* (Gyldendal, Copenhagen, 2004).
5. Materials and methods are available as supplementary materials on *Science* Online.
6. B. Deutch, J. Dyerberg, H. S. Pedersen, E. Aschlund, J. C. Hansen, *Sci. Total Environ.* **384**, 106–119 (2007).
7. H. O. Bang, J. Dyerberg, H. M. Sinclair, *Am. J. Clin. Nutr.* **33**, 2657–2661 (1980).
8. B. F. Voight *et al.*, *PLOS Genet.* **8**, e1002793 (2012).
9. I. Moltke *et al.*, *Am. J. Hum. Genet.* **96**, 54–69 (2015).
10. D. M. Altshuler *et al.*, *Nature* **467**, 52–58 (2010).
11. X. Yi *et al.*, *Science* **329**, 75–78 (2010).
12. V. S. Voruganti *et al.*, *Front. Genet.* **3**, 86 (2012).
13. L. Xie, S. M. Innis, *J. Nutr.* **138**, 2222–2228 (2008).
14. P. Rzehak *et al.*, *Br. J. Nutr.* **101**, 20–26 (2009).
15. I. M. Heid *et al.*, *Nat. Genet.* **42**, 949–960 (2010).
16. V. Gburcik, W. P. Cawthorn, J. Nedergaard, J. A. Timmons, B. Cannon, *Am. J. Physiol. Endocrinol. Metab.* **303**, E1053–E1060 (2012).
17. F. Collard, G. Delpierre, V. Stroobant, G. Matthijs, E. Van Schaftingen, *Diabetes* **52**, 2888–2895 (2003).
18. A. Jenkinson, M. F. Franklin, K. Wahle, G. G. Duthie, *Eur. J. Clin. Nutr.* **53**, 523–528 (1999).
19. 1000 Genomes Project Consortium, *Nature* **491**, 56–65 (2012).
20. O. Campuzano *et al.*, *Eur. J. Med. Genet.* **56**, 541–545 (2013).
21. K. Mirzaei, A. Hossein-Nezhad, M. Chamari, S. Shahbazi, *Minerva Endocrinol.* **36**, 13–21 (2011).
22. S. Liu *et al.*, *Cell* **157**, 785–794 (2014).
23. N. Patterson *et al.*, *Genetics* **192**, 1065–1093 (2012).

24. M. Raghavan *et al*., *Science* **345**, 1255832 (2014).
25. Y. S. Aulchenko *et al*., *Nat. Genet.* **41**, 47–55 (2009).
26. A. R. Wood *et al*., *Nat. Genet.* **46**, 1173–1186 (2014).
27. C. Jeppesen, M. E. Jørgensen, P. Bjerregaard, *Int. J. Circumpolar Health* **71**, 18361 (2012).
28. A. Ameur *et al*., *Am. J. Hum. Genet.* **90**, 809–820 (2012).
29. I. Mathieson *et al*., http://biorxiv.org/content/early/2015/03/13/016477.
30. H.-J. Quabbe, H.-J. Bratzke, U. Siegers, K. Elban, *J. Clin. Invest.* **51**, 2388–2398 (1972).
31. C. T. Damsgaard, C. Mølgaard, J. Matthiessen, S. N. Gyldenløve, L. Lauritzen, *Pediatr. Res.* **71**, 713–719 (2012).

ACKNOWLEDGMENTS

We thank the Greenlandic participants and the funding agencies and research centers that made this study possible: the Human Frontiers in Science Program Organization (grant LT00320/2014); the Danish Council for Independent Research (grant DFF-YDUN); the Villum Foundation; the Steno Diabetes Center; NIH (grant R01-HG003229); the Leverhulme Programme Grant (grant RP2011-R-045); the University of California–Merced startup funds; Karen Elise Jensen's Foundation and NunaFonden, which supported the collection of data from the Greenlandic cohorts; and the Novo Nordisk Foundation Center for Basic Metabolic Research, which is an independent research center at the University of Copenhagen and is partially funded by an unrestricted donation from the Novo Nordisk Foundation (www.metabol.ku.dk). We also thank T. Lauritzen and A. Sandbæk for the use of the ADDITION (Anglo-Danish-Dutch Study of Intensive Treatment In People with Screen Detected Diabetes in Primary Care) cohort. The Vejle Diabetes Biobank was funded by the Danish Medical Research Council and Vejle Hospital. The genotyping and exome sequencing data from this project are available to researchers who have received ethics approval from the Greenland Research Ethics Committee (nun@nanoq.gl) and can be obtained by contacting T.H.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/349/6254/1343/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S14
Tables S1 to S17
References (*32*–*67*)

30 March 2015; accepted 17 August 2015
10.1126/science.aab2319

STRUCTURAL BIOLOGY

Cryo-EM shows the polymerase structures and a nonspooled genome within a dsRNA virus

Hongrong Liu[1]*† and Lingpeng Cheng[2]*†

Double-stranded RNA (dsRNA) viruses possess a segmented dsRNA genome and a number of RNA-dependent RNA polymerases (RdRps) enclosed in a capsid. Until now, the precise structures of genomes and RdRps within the capsids have been unknown. Here we report the structures of RdRps and associated RNAs within nontranscribing and transcribing cypoviruses (NCPV and TCPV, respectively), using a combination of cryo–electron microscopy (cryo-EM) and a symmetry-mismatch reconstruction method. The RdRps and associated RNAs appear to exhibit a pseudo-D_3 symmetric organization in both NCPV and TCPV. However, the molecular interactions between RdRps and the genomic RNA were found to differ in these states. Our work provides insight into the mechanisms of the replication and transcription in dsRNA viruses and paves a way for structural determination of lower-symmetry complexes enclosed in higher-symmetry structures.

The family Reoviridae causes disease in humans, livestock, insects, and plants. The virions have 10 to 12 segments of dsRNA enclosed in a single-, double-, or triple-layered capsid. The inner capsids (cores) remain intact after the viruses are delivered into the host cell's cytoplasm, and the RNA-dependent RNA polymerases (RdRps) repeatedly transcribe RNA from the minus-strand RNA genome within the core (*1*, *2*). Assembly of the reovirus cores requires encapsidation of the genomic RNA plus strands, along with a roughly equal number of RdRps. The maturation of the reoviruses is accompanied by RdRps-driven synthesis of RNA minus strands complementary to the plus strands, in turn forming genomic double-stranded RNA (dsRNA) segments within the mature virions (*3*, *4*).

Although structures of viral capsids and isolated RdRp complexes have been studied extensively for more than two decades (*3*–*13*), the structures of genomes and RdRps within viral capsids have thus far evaded determination. In this study, we used cryo–electron microscopy (cryo-EM), in combination with our symmetry-mismatch reconstruction method, to report the structures of RdRps and associated RNAs for nontranscribing and transcribing cypoviruses (NCPV and TCPV, respectively) in the family Reoviridae.

Cypovirus particles were isolated and purified, and viral transcription was assayed (*14*, *15*). We reconstructed the structures of the NCPV and TCPV without imposing any symmetry (see supplementary materials and methods). Our analysis of the NCPV showed that the genomic RNAs and RdRps are located inside the capsid within a region of 510 Å radius. The structure of the genomic RNAs is of spherical outline, is composed of regularly distributed layers that are formed by discontinuous dsRNA fragments running in parallel, and is associated with RdRps (Fig. 1, A to C; fig. S1; movie S1). Each RdRp is anchored at the inner surface of the capsid and surrounded by multiple layers of dsRNA (Fig. 1, B and C). The distance between two adjacent dsRNA fragments within the same layer is fixed at ~25 Å, whereas two adjacent layers are ~30 Å apart. The double helices of both dsRNA fragments located close to the inner capsid surface and interacting with the RdRps have a measured helix pitch of ~28 Å (Fig. 1C). The dsRNA fragment structures located closer to the spherical center are not as well resolved as the those at the periphery (fig. S1). Each RdRp density anchors to the inner surface of the capsid, slightly off-center from the fivefold axis (Fig. 1B) (*16*). These RdRps and the associated dsRNA fragments appear to exhibit a pseudo-D_3 symmetric organization (Fig. 1A and figs. S1 to S3), allowing for 12 distinct locations of RdRps inside a viral capsid: Two groups containing three RdRps (threefold RdRps) each approach and are symmetrically arranged about the threefold axes on opposite sides of the virion, and three groups containing two RdRps (twofold RdRps) each approach and are symmetrically arranged about the twofold axes that encircle the center of the virion (fig. S4). Within the three-dimensional density maps, the average density value of the twofold RdRps amounts to approximately two-thirds of the average density value of the threefold RdRps. In contrast, the dsRNA densities surrounding the twofold and the threefold RdRps are all of similar intensity. We reason that this reflects six RdRps occupying the six positions of the threefold RdRps and only four RdRps occupying the six positions of the twofold RdRps (thus, two-thirds of the average density). Therefore, the total number of RdRps within the capsid is 10, in tentative agreement with the observation that each cypovirus genome contains only 10 RNA segments, with each genome segment being specifically associated with one RdRp (*17*). Our structural analysis also revealed that TCPV and NCPV have almost identical genome structures (figs. S2 and S5), except for those genome regions that interact with RdRps. Given the great variations of size and the encoded genes of the 10 different genomic RNA segments in each cypovirus, it is likely that the observed D_3 symmetry in the dsRNA organization does not reflect the true organization of the RNA genome. The layers of the dsRNA fragment resemble the organization of the cholesteric liquid crystal (*18*) (fig. S1 and movie S2), which is consistent with earlier evidence that the dsRNA genome forms liquid crystalline arrays within the highly condensed capsid (*5*). The liquid crystalline model of genome

[1]College of Physics and Information Science, Hunan Normal University, Changsha, Hunan 410081, China. [2]School of Life Sciences, Tsinghua University, Beijing 100084, China.
*These authors contributed equally to this work. **†Corresponding author. E-mail: hrliu@hunnu.edu.cn (H.L.); lingpengcheng@mail.tsinghua.edu.cn (L.C.)**

Greenlandic Inuit show genetic signatures of diet and climate adaptation

Matteo Fumagalli *et al.*

Science **349**, 1343 (2015);
DOI: 10.1126/science.aab2319