Transfer of signals by optical fiber is one possible solution. A number of links have been successfully demonstrated for distances of around 100 km, and, with active stabilization of fluctuations in length, can transfer the full performance of optical standards within practical averaging times. The goal is now to reach longer distances: 1000 km or more to link laboratories around Europe, several thousand kilometers to span the United States or Australia, and even longer for intercontinental links. The work of Predehl et al. sets a new distance record (920 km), but more importantly, it shows that the practical challenges such as attenuation and remote operation can be met on this distance scale. Separately, it has also been shown that these comparisons can be performed in parallel with data transfer over the same “lit” fiber—that is, one already in use on other wavelength channels (11). This capability is important, as it extends the reach of networks where “dark” fiber (installed in anticipation of future capacity increases) proves impractical or prohibitively expensive.

Together, these technologies are working toward a new international network of time and frequency standards. This network is not just for metrologists, but supports a whole range of applications, just as did earlier networks linked by telegraph, radio, or satellite. One example is precision geodesy, measuring Earth’s gravitational potential to new levels of precision with optical clocks as “Einstein sensors.” Precise synchronization is also needed for very long baseline interferometry in radio astronomy, with antenna systems such as the proposed Square Kilometre Array extending over continental scales. Here the same fiber network that transfers data can also transfer time; in future, the same can be true for research, industry, and even our homes, with unprecedented access to accurate and reliable time.

In the 60 years or so since the development of the atomic clock, technological infrastructure from telecommunications networks to satellite navigation systems has come to depend intimately not only on the performance of these clocks but also on the ability to compare them. This history tells us to expect the revolution to continue during the next 50 years of timekeeping.

References

10.1126/science.1220826
film populations contained six distinct *Leptospirillum* genotypes (types I to VI). The time series was dominated by *Leptospirillum* genotype III, which appeared to be a recombinant hybrid of genotypes I and VI. Time-series population DNA sequencing of *Leptospirillum* genotype III biofilms sampled over the course of 5 years yielded a single-nucleotide substitution rate of $1.4 \times 10^{-9}$ substitutions per nucleotide per generation. This rate estimate, along with evolutionary history reconstructions, suggested that the six *Leptospirillum* genotypes diverged from a common ancestor in a matter of decades. The data also indicated that the major recombinations reflected single events that led to the domination of specific recombinant genotypes. Population genomic analyses suggested that *Leptospirillum* genotype evolution (see the figure) consisted of a complex menu of immigration, horizontal gene transfer (HGT), homologous recombination, fixation, and selective sweeps that successively generated and replaced different genotypes and, presumably, phenotypes.

The AMD *Leptospirillum* nucleotide substitution rate of $1.4 \times 10^{-9}$ per nucleotide per generation is at the high end of previous estimates, which have ranged from $7.2 \times 10^{-11}$ to $4.0 \times 10^{-9}$ ($5$–$7$). One reason may be the use of a different approach. In perhaps the most elegant and comprehensive study of its kind to date, Denef and Banfield performed cultivation-independent population genomic analyses. In addition, they leveraged a unique combination of data that included in situ proteomic data sets ($8$), three distinct *Leptospirillum* population genotypes (assembled directly from natural populations) and a population genomic time series.

Unlike Denef and Banfield’s study, other estimates of bacterial evolutionary rates have relied on isolation, propagation, or comparison of pure cultures or laboratory populations ($5$, $7$, $9$). These approaches may be selective and may not reflect native community complexity. On the other hand, the AMD habitat is itself unique; its unusual conditions and low biological complexity may affect the evolutionary rate estimates. Denef and Banfield point out, however, that their estimates are consistent with independent theoretical predictions of a universal mutation rate ($10$).

A central goal in microbial population biology is to identify key genomic changes
that lead to adaptation and selection—the “differences that make a difference” (11–13). Denef and Banfield speculate that phenotypic selection in the *Leptospirillum* populations may be due to differences of only a few nucleotides, but there are some caveats. Their approach captured major homologous recombination events, but nonhomologous gene acquisitions from HGT may have gone undetected, because the *Leptospirillum* population analyses relied on comparisons of genes present in reference genomes. Without complete and contiguous genome assemblies from the population genomes, large blocks of newly acquired foreign genes could be invisible to this approach.

Recent studies suggest that HGT and illegitimate recombination can be common and frequent in native microbial populations (13–15).

It turns out that “genome” is a verb, not a noun—a process, not a product. Current and emerging microbial population genomic studies that capture this dynamic (2, 6, 7, 9, 13–15) promise to help paint a more integrated motion picture of microbial ecology and evolution in action.

References

10.1126/science.1221822

**ASTRONOMY**

**Complex Protostellar Chemistry**

Joseph A. Nuth III and Natasha M. Johnson

Two decades ago, our understanding of the chemistry in protostars was simple—matter either fell into the central star or was trapped in planetary-scale objects. Some minor chemical changes might occur as the dust and gas fell inward, but such effects were overwhelmed by the much larger-scale processes that occurred even in bodies as small as asteroids. The chemistry that did occur in the nebula was relatively easy to model because the fall from the cold molecular cloud into the growing star was a one-way trip down a well-known temperature-pressure gradient; the only free variable was time. However, just over 10 years ago it was suggested that some material could be processed in the inner nebula, flow outward, and become incorporated into comets (1, 2). This outward flow was confirmed when the Stardust mission returned crystalline mineral fragments (3) from Comet Wild 2 that must have been processed close to the Sun before they were incorporated into the comet. On page 452 of this issue, Ciesla and Sandford (4) demonstrate that even the outermost regions of the solar nebula can be a chemically active environment. Their finding could have consequences for the rest of the nebula.

Outward flow in the nebula is the natural result of conserving the total angular momentum of the star-forming system as mass continues to accrete onto the rotating disk feeding the growth of the central star (5–7). As mass flows inward along the flared accretion disk, additional mass flows slowly outward along the midplane (see the figure). As an additional complication, because the midplane is hotter than the outer boundary of the disk, convection will also mix materials vertically in the nebula. Ciesla and Sandford show that ice-coated dust grains, moving outward and subject to convection, will be exposed to cosmic radiation that is sufficient to cause the same chemical effects seen in dark cloud cores—that is, the conversion of simple carbon-and nitrogen-containing molecules into more complex organic species—and so will have consequences for nebular chemistry.

It has been nearly 40 years since the discovery that the oxygen in solid bodies throughout the solar system is mass-independently fractionated (8). It appears as if $^{16}$O were added or subtracted from the bulk composition of the dust independently of either $^{17}$O or $^{18}$O. While the first explanation for this observation was the addition of $^{16}$O-rich supernova grains to the forming solar system (8), the currently favored explanation is known as chemical

**Numerical models show that protoplanetary nebulae are sites of chemical activity even in the cold outer disk.**

**Complex reactions.** Large-scale motion driven by conservation of angular momentum, together with more local convective cells above and below the hotter nebular midplane, dynamically mix products of chemical reactions from many different environments throughout the nebula.

**Figures**

**Astronomical units, logarithmic scale.**

Distance from protosun (astronomical units, logarithmic scale) vs. mass accretion rate (solar masses per year) for solar system protoplanetary disk models. The blue and red curves show predictions for models where protoplanetary disk accretion is controlled by either gravitational or viscous forces, respectively. The green curve shows predictions for a model where accretion is controlled by both gravitational and viscous forces. The red curve with error bars shows the range of observed values for the mass accretion rate of the solar system protoplanetary disk. The orange curve shows the range of observed values for the mass accretion rate of other protoplanetary disks. The orange curve with error bars shows the range of observed values for the mass accretion rate of the solar system protoplanetary disk. The orange curve with error bars shows the range of observed values for the mass accretion rate of other protoplanetary disks.

**Collimated bipolar outflow.** Collimated bipolar outflow is a property of protoplanetary disks. The outflow is collimated, meaning that it is directed in a specific direction, and it is bipolar, meaning that it is directed in two opposite directions. The collimated bipolar outflow is thought to be driven by the centrifugal force of the protoplanetary disk, which is the force that acts on the material in the disk to make it move in a circular path. The centrifugal force is greater than the gravitational force of the central star, which is the force that acts on the material in the disk to make it fall into the star. The difference between the centrifugal force and the gravitational force is the force that drives the outflow.
biminerization activity may be sought in the form of Ba- and Sr-rich carbonates. These Ba- and Sr-rich phases therefore add to the list of potential mineral bio signatures in the fossil record.

References and Notes

13. See supplementary materials on Science Online.

Acknowledgments: E.C. was recipient of a fellowship from the French Ministère de la Recherche et de l’Enseignement Supérieur. This project was financed by the CNRS interdisciplin ary program “Environnements plané taires et origines de la vie” (PID EPPO) and Institut National des Sciences de l’Univers (INSU) program “InteractionsTerre/Vie” (InterVie). The SEM facility of the Institut de Minéralogie et de Physique des Milieux Condensés is supported by Région Ile de France grant SESAME 2006-1-D7-593/R, INSU-CNRS, INP-CNRS, Université Pierre et Marie Curie, Paris. Support was also provided by NSF and the U.S. Environmental Protection Agency under cooperative agreement EF-0830093 through the Center for Environmental Implications of Nanotechnology. Advanced Light Source (ALS) Molecular Environmental Science beamline 11.0.2 is supported by the Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences and Materials Sciences Division, U.S. Department of Energy, at the Lawrence Berkeley National Laboratory. Beamline 5.3.2 at ALS is supported by the Office of Basic Energy Sciences of the U.S. Department of Energy under contract DE-AC02-05CH11231. We thank R. Tavera (Universidad Autonoma de Mexico) for providing help as a field guide and the Laboratoire de Geochimie des Eaux at IPGP for providing the facility for Gran measurements. The sequence of the 16S rRNA gene of Candidatus Gloeomargarita lithophora has been deposited in GenBank under accession no. JQ733894. Data and enrichment cultures are available upon request.

Supplementary Materials

www.sciencemag.org/cgi/content/full/336/6080/459/DC1

In Situ Evolutionary Rate Measurements Show Ecological Success of Recently Emerged Bacterial Hybrids

Vincent J. Denef† and Jillian F. Banfield‡

Few data are available on how quickly free-living microorganisms evolve. We analyzed biofilms collected from a well-defined acid mine drainage system over 9 years to investigate the processes and determine rates of bacterial evolution directly in the environment. Population metagenomic analyses of the dominant primary producer yielded the nucleotide substitution rate, which we used to show that proliferation of a series of recombinant bacterial strains occurred over the past few decades. The ecological success of hybrid bacterial types highlights the role of evolutionary processes in rapid adaptation within natural microbial communities.

Microbial communities, which drive Earth’s geochemical cycles (1), can rapidly respond to change, but the proportion of this response that can be attributed to evolutionary processes, rather than species composition or gene expression shifts, remains an unresolved
question (2). Most evolutionary rate estimates are available for nucleotide substitution rates and derive from laboratory measurements (3, 4). It is difficult to know how relevant these rates are for geochemical environments, because studies on natural populations have been restricted to pathogens (5–7) and endosymbions (8).

Our approach to measuring evolutionary rates for free-living bacterial populations involved tracking genome change in a time series of natural microbial community samples. Few systems are suitable for such a study; requirements include the availability of discrete, well-defined, and relatively homogeneous microbial community samples from a location where a reproducible community forms over time, with very restricted microbial input from other regions. The microbial communities catalyzing the formation of acid mine drainage (metal-rich, low-pH solutions) in the Richmond Mine (Iron Mountain, CA) provide such a system. Biofilms that develop at the air-solution interface on standing pools and slow-flowing underground streams have been used for more than a decade to study evolutionary processes and ecological complexity in nature (9). Key to the success of this environment as a model system has been its low species richness, which enables detailed culture-independent community genomics analyses of the genetic structure and dynamics of natural populations (10–13). Most biofilms within this system are dominated by the chemoolithotrophic Leptospirillum group II. Previous metagenomic studies on two biofilms led to the reconstruction of genomes for Leptospirillum group II, types I and VI, the genes of which share ~94% average nucleotide identity (10, 13, 14). Inferences based on proteomics data indicate that large-scale homologous recombination events occurred between two parental Leptospirillum group II types, resulting in six recombinant (hybrid) genotypes (14, 15). One genotypic group, comprising types II through VI, predominates during initial colonization, whereas type I becomes abundant in later successional stages (16).

We have collected biofilms within the Richmond Mine over approximately 9 years. Sampling required field expeditions to locations where ambient conditions are often close to the limit of human endurance (e.g., 100% humidity, ~48°C), with depressed O2 levels at sites only accessible at times of the year when flow rates are low and danger from underground collapse is minimized. Working within these constraints, we collected microbial biofilms (both time series and spatially resolved) from six underground locations between March 2002 and December 2010 at intervals of 3 weeks to more than 1 year (Fig. 1, table S1, and figs. S1 and S2) (17). We selected 13 samples from the C75 location, a site indicated by proteomics-based inferences to be typically dominated by the Leptospirillum group II, type III genotype (15). This allowed us to analyze single-nucleotide substitution accumulation over 5 years, and thereby to estimate the substitution rate. On the basis of the same proteomics-based inferences, we also selected nine biofilms that provided multiple samplings of most of the other Leptospirillum group II genotypes (types I, III, IV, V, and VI). Together with the two sequence data sets from which the type I and type VI reference genomes had been reconstructed, this allowed us to reconstruct a lineage history (fig. S1).

We generated community genomic data sets comprising ~63 billion base pairs of sequence from 24 samples, each of which contained ~108 Leptospirillum group II cells (10). Using sequence obtained from the first five C75 location biofilms, for which longer read lengths were available, we reconstructed a type III genome de novo (17). The genome is indeed a recombinant hybrid of the previously reconstructed type I and type VI genomes, as deduced from proteomic analysis (15). All identified recombination points were located within genes (table S2 and fig. S3). All Leptospirillum group II sequences were accounted for (fig. S4) (17), excluding the possibility of other genotypes in these five samples. The sequencing reads from each of the 13 C75 samples collected over 5 years were recruited to the previously reconstructed Leptospirillum group II, type I and type VI genomes (13, 18) to identify the recombinant block structure of the genome at each sampling point (fig. S1).

The recombinant block structure of the genome was reconstructed using data from the Richmond Mine and type III strain turnover at the C75 location. A Richmond Mine schematic map, with pie charts indicating genotype proportions in 24 samples, estimated on the basis of read recruitment (figs. S5 and S9). (B) Acid mine drainage flow rate (measured at the mine entrance) and community composition at the C75 location, measured by fluorescence in situ hybridization (Arc, Archaea; L3, Leptospirillum group III; L2, Leptospirillum group II). Leptospirillum group II, type III strain transitions, revealed by SNPs analysis, are indicated by an asterisk inside the Leptospirillum group II. (C) Flow data, 2001 to 2011. Shading indicates period displayed (B).

Fig. 1. Leptospirillum group II genotype distribution shows dispersal across the Richmond Mine and type III strain turnover at the C75 location. (A) Richmond Mine schematic map, with pie charts indicating genotype proportions in 24 samples, estimated on the basis of read recruitment (figs. S5 and S9). (B) Acid mine drainage flow rate (measured at the mine entrance) and community composition at the C75 location, measured by fluorescence in situ hybridization (Arc, Archaea; L3, Leptospirillum group III; L2, Leptospirillum group II). Leptospirillum group II, type III strain transitions, revealed by SNP analysis, are indicated by an asterisk inside the Leptospirillum group II. (C) Flow data, 2001 to 2011. Shading indicates period displayed (B).
nucleotide per generation (4, 7, 8). It is similar to the rate of $1.3 \times 10^{-8}$ substitutions per nucleotide per generation predicted for the type III genome size using the universal mutations-per-genome rate suggested by Drake (3). Although the highest rates have been reported for endosymbionts and pathogens, which have small effective population sizes, correspondence with Drake’s prediction may indicate that genetic drift accelerated by population bottlenecks is not the main factor affecting the substitution rate in *Leptospirillum* group II.

To assess potential biases in the measured substitution rate introduced by spatial variation within the contiguous biofilms covering the C75 acid mine drainage pool, we sampled three locations ~1 m apart in June 2008. Samples 1 and 3 were dominated by the same genotype (with four SNPs relative to the type III genome), whereas sample 2 was dominated by a variant of this genotype lacking one high-frequency polymorphism (fig. S7A). The three samples also had only 93 (~8) replicated SNPs at frequency greater than 0.05 (average frequency 0.08). The low levels of variation, both within populations and across space, provide confidence that nucleotide changes in populations at the C75 location over time can be used to calculate the substitution rate.

Read recruitment of community genomic data from sites located across the underground system (five-way, UBA, C75, C10, AB muck, and AB20 locations) to types I and VI yielded six genotypes (Fig. 1 and fig. S9): type I, type VI, and four other types (III, IV, IVa, and V) that each consist of a mosaic of type I and type VI genome blocks tens to hundreds of kilobases in length (Fig. 2, A and B, and figs. S6B and S9). This again confirmed previous strain-resolved proteomic analysis–based inferences (14, 15). The presence of exactly the same transition points between the recombinant blocks in each of the genotypes in every sample indicates that each major recombination event occurred in a single cell, whose descendants rose to fixation (Fig. 2). Each genotype’s fixed mutations were identified by read recruitment to type III (Fig. 3 and fig. S7B) (17) and were used to construct a phylogenetic tree (Fig. 2C). Because read recruitment was performed

---

**Fig. 2.** Reconstruction of the recent evolutionary history of *Leptospirillum* group II. (A) The outer circle plots the counts of peptides unique to the type I (red) and type VI (blue) reference genomes at each protein locus (dots) for the C75 type III population [proteomics-inferred genotyping (PIGT); data from (15)]. Inner circles show Illumina reads recruitment to the reference genomes, revealing the prevalence of type III, IV, IVa, V, and VI recombinants. (B) Correspondence of PIGT and recruitment data, highlighting the identical recombination point in genotypes III to V. (C) Evolutionary history of the sampled genotypes, based on the variant loci (table S7), inferred using the maximum parsimony method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test is shown next to the branches. Dotted arrows indicate recombination events; circle schematics represent the regions affected. The timeline indicates the calculated time ranges of recombination events (table S3) as well as historical events (fig. S2) (18, 21). Branch D* is presented as two strains, each assigned half the total number of UBA 06/05 SNPs, because low incidence of SNPs precluded their linkage. BP, years before the present.
relative to the type III genome, positions within type I–like regions not shared by type IV, V, or VI genomes were designated as gaps for tree construction, and these gaps were treated as missing data in the calculation of tree branch length. Notably, the tree topology indicates that the larger the fraction of type I–like sequence recombined into the type VI background, the more recently the specific genotype emerged (Fig. 2C).

Using the substitution rate calculated from the C75 time series and assuming equal generation times for the different genotypes, we estimated that the times to coalescence for the suite of recombinant genotypes ranged between 2 and 44 years (Fig. 2 and table S4). The causes of the success of these hybrid genotypes are hard to establish. In Bacteria and Archaea, little is known about the importance of hybridization in adaptation to change. However, one recent study hypothesized that agricultural practices led to proliferation of hybrids of Campylobacter jejuni and Campylobacter coli formed by homologous recombination (19). In Eukarya, hybrids are often reported to have little ecological success (20). However, increased fitness relative to parental types has been observed when the adaptive landscape changes as a result of natural- or human-induced environmental alterations (21, 22). The Richmond Mine site is subject to both natural (seasonal and year-to-year variation; Fig. 1C) and human perturbations (Fig. S2) (17), a subset of which may have provided opportunity for hybrid proliferation.

Although we cannot connect specific events to emergence of the hybrid genotypes, we can evaluate the role of neutral versus selection-based evolutionary processes in lineage divergence. For example, extreme population bottlenecks such as flushing events that remove biofilms during the rainy season may purge diversity independent of the fitness of particular strains. However, there are several lines of evidence indicating that hybrid genotypes arose as the result of selective processes. We detected evidence for positive selection (McDonald-Kreitman test, \( P < 0.05 \)) for mutations specific to one type VI population (branch C; see Fig. 2), type V (branch F), and the ancestor of types III to V (branch E) (table S5). Selection was also supported by the disproportionately high number of signal transduction genes (Fig. 3, functional category T) and transcriptional regulation genes (Fig. 3, category K), global regulators in particular, that were affected by fixed mutations (Fig. 2, branches E, F, D*, and H; see also tables S5 and S6). Previous observations in laboratory settings (23–25) and a natural system (16) similarly identified evolution of gene expression as an important factor underpinning early ecological divergence. Finally, an earlier study analyzing the distribution of the different Leptospirillum group II genotypes in the context of geochemical conditions indicated selection among type I through type VI recombinants (15). Evidence for selection also emerged from the current study, where two sites located ~140 m apart along a single flow path were compared (C10 and C75). The type III population, which dominated all upstream C75 biofilms, was found in only one of three biofilms sampled at the same time at the downstream C10 location. Even when the type III population was present at both sites, the genomes were different (Fig. 1 and fig. S10). This result suggests selection between genotypes that differ by only a few nucleotides—a finding consistent with prior laboratory studies showing increased fitness due to adaptive SNPs (4, 25). It also indicates that lack of diversity is not the explanation for dominance of many biofilms by a single genotype.

**Fig. 3.** Overview of the Leptospirillum group II high-frequency variants, based on read recruitment to the C75 June 2006 type III genotype. Dots indicate variant nucleotides relative to the last common ancestor of types III to VI. Functional categories (table S6) are indicated only for nonsynonymous substitutions and only for the first occurrence, starting from the most recent genotype (outside circle). Shading in inner circles indicates regions that were excluded for SNP analysis in a particular genotype because of recombination or indel events.
Our population genomic analyses provided data on the accumulation of genome change in free-living populations that underpin geochemical cycles. Application of these rates revealed that a series of important divergence events occurred over a time scale of years to decades within the natural model system of the Richmond Mine. We attribute these to major selection episodes, although we cannot conclude whether they were mediated by natural or human factors. Rapid adaptive evolution may have assisted in the maintenance of Leptospirillum group II as both the dominant primary producer and most active iron oxidizer responsible for acid mine drainage formation. Our results contribute to the development of a predictive understanding of how microbial systems respond to both natural and anthropogenic change (2).

References and Notes
2. A. Reid, M. Buckley, Microbial Evolution (American Academy of Microbiology, Washington, DC, 2011).
17. See supplementary materials on Science Online.

Acknowledgments: We thank T. W. Arnan (president, Iron Mountain Mines Inc) and R. Sugarek (U.S. Environmental Protection Agency) for site access; R. Carver and M. Jones for on-site assistance; R. Carver for Richmond Mine flow data; J. Sickles (U.S. EPA) for providing site history documentation; Banfield laboratory members for their contributions to sample collection; C. Miller, C. Sun, and B. Thomas (University of California, Berkeley) for assistance with next-generation sequencing data analysis; and C. Miller and K. Vogel (University of Arizona) and three anonymous reviewers for constructive criticism. Supported by the Office of Biological and Environmental Research, U.S. Department of Energy, through the Genomic Sciences (DE-FG02-05ER64134) and Carbon-Cycling (DE-FG02-10ER64996) programs. The authors declare no conflict of interest. All 454 and Illumina sequencing reads generated for this study have been deposited in the NCBI sequence reads archive (accession number SRA047370). The Leptospirillum group II C75 strain genome has been deposited at DDBJ/EMBL/GenBank under the Whole Genome Shotgun project accession number AJM00000000. The version described in this paper, containing the assembled and annotated contigs, is the first version, AJM01000000.

**Supplementary Materials**
www.sciencemag.org/cgi/content/full/336/6080/462/DC1
Materials and Methods
Figs. S1 to S11
Tables S1 to S7
References (26–48)
23 December 2011; accepted 9 March 2012
10.1126/science.1218389

Origins and Genetic Legacy of Neolithic Farmers and Hunter-Gatherers in Europe

Pontus Skoglund, Helena Malmström, Maanasa Raghavan, Jan Storå, Per Hall, Eske Willerslev, M. Thomas P. Gilbert, Anders Götherström, Mattias Jakobsson
†

The farming way of life originated in the Near East some 11,000 years ago and had reached most of the European continent 5000 years later. However, the impact of the agricultural revolution on demography and patterns of genomic variation in Europe remains unknown. We obtained 249 million base pairs of genomic DNA from ~5000-year-old remains of three hunter-gatherers and one farmer excavated in Scandinavia and find that the farmer is genetically most similar to extant southern Europeans, contrasting sharply to the hunter-gatherers, whose distinct genetic signature is most similar to that of extant northern Europeans. Our results suggest that migration from southern Europe catalyzed the spread of agriculture and that admixture in the wake of this expansion eventually shaped the genomic landscape of modern-day Europe.

The transition from the hunter-gatherer lifestyle to a sedentary farming economy spread throughout Europe from the south-east, starting ~8400 years before the present (yr B.P.) (1). Early population genetic studies argued that clines in genetic variation within Europe favored a model in which this expansion occurred in concert with substantial replacement of resident hunter-gatherer populations (2–4), contradicting models emphasizing the culturally mediated spread of farming economy (5). Current results from extant populations are inconclusive (4, 6); however, ancient DNA analyses (7–10) present tentative evidence for population replacement but suffer from uncertainties associated with single-locus studies of the Y chromosome or the mitochondrion. Population genomic analysis of ancient human remains (6, 11) shows promise, but poses technical difficulties due to low DNA yield and the risk of present-day human contamination (6). In addition, geographical and temporal differences between ancient samples make observations of genetic differentiation difficult to interpret. Thus, more robust interpretations of ancient DNA might be gained by analyzing samples from cultural complexes occurring in the same region and during the same time period. In this study, we focused on the Neolithic era of northern Europe, where the relatively late arrival of farming (~6000 yr B.P.) was followed by more than 1000 years of coexistence between hunter-gatherer and farming cultures (12).

We obtained genomic DNA sequences from three samples (Ajv52, Ajv70, and Ire8) from a hunter-gatherer context (“Neolithic hunter-gatherers” associated with the Pitted Ware Culture (PWC)), one sample from a farming context (“Neolithic farmer” associated with the Funnel Beaker Culture or Trichterbecher kultur (TRB)), and two animal remains as contamination controls. The human remains were chosen from a larger panel on the basis of their molecular preservation and previously yielded reproducible single-locus genetic data (8, 13, 14). The Neolithic farmer sample (Gök4) was excavated from a mega-lithic burial structure in Gökhem parish, Sweden, and has been directly 14C-dated to 4921 ± 50 calibrated yr B.P. (cal yr B.P.), similar to the age (5100 to 4900 cal yr B.P.) of the majority of other finds in the area (15). There were no indications from the burial context suggesting that Gök4 was different from other TRB individuals (15, 16), and strontium isotope analyses indicate that Gök4 was born less than 100 km from the