18. Ancient genomics in Neolithic central Anatolia and Çatalhöyük

Reyhan Yaka, Ayça Doğu, Damla Kaptan, Nihan Dilşad Dağtaş, Maciej Chyleński, Kıvılcım Başak Vural, N. Ezgi Altınışık, Igor Mapelli, Dilek Koptekin, Cansu Karamurat, Hasan Can Gemici, Sevgi Yorulmaz, Vendela Kempe Lagerholm, Evrim Fer, Ulaş Işıldak, Ayshin Ghalichi, Gülşah Merve Kılınç, Camilla Mazzucato, Anna Juras, Arkadiusz Marciniak, Clark S. Larsen, Marin Pilloud, Scott D. Haddow, Christopher J. Knüsel, İnci Togan, Anders Götherström, Yılmaz Selim Erdal, Elif Sürer, Füsun Özer, Çiğdem Atakuman and Mehmet Somel

Over multiple millennia, from the earliest traces of long-term occupation of camp sites (ca 20,000 BC) to the development of full-scale farming (ca 8000–6000 BC), the Neolithic transition in southwest Asia gradually shaped human societies in dramatic ways (Nadel 2002; Maher et al. 2012; Asouti, Fuller 2013). Here, we present recent insights from ancient genomics studies into these societies while focusing on two questions: the population processes driving cultural change in Neolithic central Anatolia and genetic kinship among Çatalhöyük co-burials.

Ancient genomics has most frequently been employed in inferring demographic history, which, in turn, can shed light on questions surrounding cultural change. During the Neolithic Transition, many new material cultural traits, raw materials and technologies, as well as new patterns of social organisation and belief systems were created and shared among geographically distinct communities of the Levant, Zagros Mountains, Mesopotamia, Anatolia, Cyprus and the Aegean (Ibáñez et al. 2018; Reingruber 2011). While these exchanges can be partly traced from material culture, the type of human interactions that drove these transitions remains little understood. Did new traits spread purely through cultural interaction among societies, via acculturation, emulation or material exchanges? Or did the new traits spread via human movement between regions? If such human mobility did have a role, was it a steady dispersal process involving conjugal exchange, or did it occur in pulses, as mass migration events? Was human movement sex-biased, or did it involve males and females equally? Today, we can discern among different modes of human mobility and material culture change by studying ancient genome data integrated with archaeological data.

Ancient genomics can further help investigate social traditions of Neolithic societies. Were the earliest settled societies organised as nuclear or extended families, or were they not organised around biological ties at all? Studying the subject solely based on material culture is notoriously difficult (Kuijt, Goring-Morris 2002). To date, a number of studies have estimated biological relatedness among co-buried individuals in southwest Asian Neolithic societies using either dental features (Pilloud, Larsen 2011; Alt et al. 2015) or maternal genetic markers (Chyleński et al. 2019). However, these approaches provide only low-resolution perspectives, as we will discuss later. Kinship has thus remained a largely unwritten chapter for the Neolithisation of southwest Asia.

The potential of ancient genomics in studying demography and kinship

Ancient DNA (aDNA) analysis refers to the study of degraded DNA molecules from ancient remains that contain organic matter, such as bones, teeth, coprolites, dental calculus, or sediments. In certain ways aDNA work today involves the same steps as standard modern DNA analysis: DNA purification (extraction), amplification, sequencing and bioinformatic analysis, aimed at inferring kinship, demographic history, etc. However, unique characteristics of aDNA render its analysis additionally challenging – compared to that of modern DNA. Most important is the fact that DNA molecules break down over time, eventually to the point of non-existence.

The first generation of aDNA

From the 1980s to the early 2010s, human aDNA research was an exceptionally cumbersome and much more precarious endeavour than it is today. The DNA technologies of the late 20th century permitted manipulation and sequencing of only specific short DNA regions (ca 100 base pairs), in a targeted fashion, one by one, rather than the whole set of chromosomes, that is, the genome (ca 3 billion base pairs). Most aDNA work was focused on mitochondrial DNA (mtDNA), which is the small DNA molecule of the cellular organelle passed from mother to offspring. Mitochondrial DNA was an obvious target because it is present at 100–1,000 times

higher copy numbers than nuclear DNA (the main source of DNA in a human cell). Therefore, a random fragment of mtDNA is more likely to have survived in a bone sample than a random fragment of nuclear DNA. It is also a particularly variable part of the genome and even short sequences of mtDNA can contain sufficient information to distinguish between closely related lineages.

However, mtDNA provides information only about the maternal lineage and this constitutes a severe limitation when inferring genetic kinship or population history. For example, an individual will appear as unrelated to her/his father as to any random individual from the perspective of mtDNA analysis. In contrast, nuclear DNA (or whole genome data) is derived from all ancestors, and the information it provides is much more extensive. Demographic history inferred about only maternal ancestors (from mtDNA) and about all ancestors (from genome data) can be strikingly different. For instance, mtDNA analysis on the first Denisovan remains had suggested that humans and Neanderthals were sister-species while Denisovans were an outgroup; in contrast, whole genome analysis showed that for most of their ancestry Denisovans were actually sister-species to Neanderthals (Reich et al. 2010). Being mainly restricted to mtDNA was a serious drawback for aDNA research in early periods.

A second major obstacle hindering the early aDNA studies was posed by the risk of modern-day DNA contamination and the seeming inability to distinguish between authentic ancient DNA and modern human DNA, which is ubiquitous (Hofreiter et al. 2001; Malmström et al. 2005). It can be difficult, if not impossible, to fully determine whether a *single* short DNA molecule extracted from an archaeological sample is authentic (endogenous aDNA) or represents external human DNA contamination (that could have happened in the field or in the lab). Hence, despite the adoption of many precautions against contamination by laboratories, human aDNA analysis was for a long time fraught with scepticism.

The technical challenges, the limited information available per sample and the spectre of modern DNA contamination collectively impeded aDNA research on human archaeological material for about three decades. The field remained confined to a few specialised laboratories until the 2010s. During this period, most aDNA work focused on species other than *Homo sapiens* and particularly on extinct species, such as Neanderthals and woolly mammoths, the DNA sequences of which would not be easily confused with those from their present-day relatives.

Second generation aDNA

The late 2000s and early 2010s saw major shifts in aDNA analysis approaches and the emergence of a new field, that of ancient genomics (Stoneking, Krause 2011; Pickrell, Reich 2014; Skoglund, Mathieson 2018). In this approach, aDNA molecules extracted from organic material are processed and sequenced in parallel, on the order of millions to billions of DNA fragments. This yields sequence information across the whole genome of an organism within a single, relatively simple experiment. The information thus retrieved is orders of magnitude more extensive than that derived from sequencing mtDNA. Indeed, genome-wide data from even a single individual will yield information about all the individual's recent ancestors. For instance, for a non-inbred person this would translate into 128 ancestors seven generations ago, while studying mtDNA would inform about a single ancestral lineage only.

The first Neanderthal genome study was a prime example demonstrating the power of this approach. Comparisons of Neanderthal and modern human mitochondrial DNA sequences had initially revealed no evidence for Neanderthal introgression into the modern human gene pool (Serre et al. 2004). Once the Neanderthal genome was sequenced, however, this new data permitted geneticists to identify traces of recently introduced Neanderthal ancestry in modern-day non-African genomes, strongly indicating the occurrence of low-level introgression of Neanderthals in the ancestors of non-African humans (Green et al. 2010). Since that time, ancient genome analyses have been growing at rapid speed. All this was made possible by innovations in four areas, which we summarise below.

Next generation sequencing

The most fundamental development was the advent of next generation (or second generation) sequencing (NGS) technologies. This permits shotgun sequencing of all available genomic DNA from a sample – in parallel – producing much more extensive information at much lower costs than previously available. Here 'shotgun' refers to the fact that any molecule obtained from the material is processed and sequenced in parallel with others, irrespective of its origin, in contrast to earlier targeted approaches that analysed only specific parts of DNA (such as mtDNA loci). Importantly, the experiments involved in preparing the aDNA extracts for sequencing are technically not more challenging than those of the earlier 'targeted' approaches.

Novel laboratory protocols

Another line of development involved laboratory methods that boosted aDNA retrieval. These were critical for aDNA analyses of material from temperate regions, such as Neolithic bones from Anatolia because DNA decays faster at higher environmental temperatures and therefore little endogenous DNA remains in bones older

than a few millennia in warm climates (Allentoft et al. 2012). Advances included the identification of the petrous temporal bone as a rich source of DNA (Gamba et al. 2014), the development of efficient aDNA isolation and processing protocols tailored for retrieving short degraded molecules (Dabney et al. 2013; Gansauge, Meyer 2013) and removing damage from ancient molecules (Briggs et al. 2010; Rohland et al. 2015). Another technique that strongly influenced ancient genomics was the hybridisation capture procedure for enriching human DNA molecules (Carpenter et al. 2013; Fu et al. 2013). Within a DNA extract from an ordinary Anatolian Neolithic bone, the vast majority (usually >99%) of molecules will not be human but rather environmental DNA, mainly from bacteria. Hybridisation capture, a type of molecular fishing, alleviates this situation by enriching human DNA molecules in total DNA extracts. Enrichment, in turn, reduces the total amount of next generation sequencing needed to achieve the same amount of information. Enrichment has been employed in two ways: targeting the whole genome (Allentoft et al. 2015; Kılınç et al. 2016), or targeting specific variants (single nucleotide polymorphisms, or SNPs) (Haak et al. 2015). The latter method appears particularly efficient in enrichment but is limited to only previously ascertained genetic variants; this kind of data may complicate downstream demographic inferences, such as population diversity measures.

In silico *identification of DNA contamination*

Another major development involved statistical methods for estimating modern DNA contamination from shotgun sequencing experiments. This development was a growing understanding of the types of modifications and post-mortem damage (PMD) that accumulate within DNA through time – specifically, cytosine deamination causing C to T mismatches at the ends of molecules (Hofreiter et al. 2001; Briggs et al. 2007). A number of *in silico* methods were thus developed to estimate the amount of human contamination in an ancient genome dataset using information on PMD or on abnormal patterns of polymorphism at haploid loci (for example, more than one type of mtDNA in the same DNA sample, which does not biologically occur) (Green et al. 2010; Fu et al. 2013). These methods not only permit identification of modern contamination in aDNA datasets, but they can also be used to select molecules with PMD signatures, thus filtering out potential contaminant molecules and saving at least part of the authentic data (Skoglund et al. 2014; Fu et al. 2015). Contamination has thereby become less of a peril and more of an economic issue for human aDNA work, although aDNA laboratories still strive to avoid it as much as possible due to the cost it incurs.

Novel computational protocols

Finally, the last decade has witnessed the development of computational protocols and statistical methods for analysing ancient genomes. This is important as most population genomic methods are tailored for high- or medium-quality modern genomes (for example, covering the full human genome 30 times, or 30X coverage). Conversely, ancient genome datasets are generally of both much lower and also more variable quality compared to modern DNA datasets. Most cover only a small fraction of the whole genome, are comprised of only short sequence reads, and contain post-mortem damage. They thus require special treatment. This added treatment frequently involves (a) considering only variants that have been identified elsewhere, (b) considering only variants that cannot be confounded with post-mortem damage, and (c) removing some of the data to equalise quality among samples (a process called 'pseudohaploidisation') (Skoglund et al. 2014; Mathieson et al. 2015).

Meanwhile, population genomic methods have been developed that can be efficiently applied to such sub-optimal genomic data. These include the outgroup f_3 statistic that measures genome-wide genetic affinity between two groups and the *D*-test (or f_4 statistic) used to test whether a 'Test' population has higher affinity to one of two alternative populations (Green et al. 2010; Patterson et al. 2012). Here, higher genetic affinity is measured as the proportion of shared variants between two populations and it indicates a relatively higher amount of shared ancestry, such as with a sister versus with a cousin. Importantly, such tests can be performed with populations represented by even single individual genomes and can result in significant signals, given the wealth of information in a single genome.

The laboratory routine

We now describe a routine second generation aDNA experiment on a human archaeological sample (fig. 18.1). The initial part is called 'pre-screening' – an initial attempt to gauge the amount of useful DNA in a sample. The first step of pre-screening involves molecular work: well-preserved parts of the human skeleton (for example, the petrous temporal bone, teeth, auditory ossicles or finger bones) are brought into the clean room, powderised, the DNA is extracted, a sequencing library (a modified collection of DNA molecules) prepared and shotgun sequenced on an NGS (Next Generation Sequencing) machine, the latter task frequently performed at specialised centres. Sequencing at this stage is in small volumes; that is, 'low coverage' or 'low depth', restricted to millions of DNA molecules.

The second step of pre-screening is computational: the resulting raw data are cleaned *in silico* and mapped to the human reference genome. This yields two important

statistics: (a) the ratio of human sequences mapping (aligning) to the reference human genome over the whole set of sequences (note that the non-human portion here may be of microbial origin or be DNA artefacts); (b) the PMD signal across the mapped human sequences, that is, the proportion of C to T type of mismatches at the ends of molecules (fig. 18.1).

If the pre-screening results from a sample reveal that the human-specific proportion of DNA is too low (for example, <0.5% of molecules being human) and/or the PMD signal is too weak to be compatible with authenticity, the researcher will normally choose to discard that sample, since it is impossible to gain useful information from such material using current methods. Otherwise, the sample may be further processed. This may involve (a) deep sequencing, that is, sequencing large volumes of molecules, from the same library or newly built libraries from the same sample, with hundreds to thousands of millions of molecules produced; (b) enriching for human molecules via hybridisation capture (whole genome capture or SNP-capture) on the libraries and then performing new sequencing, to increase the efficiency and thus reduce sequencing costs.

The resulting data are then used in computationally estimating the genetic sex and the genotype of that individual. Genetic sex determination is based on the proportion of DNA sequences deriving from the X chromosome to other chromosomes. Calling genotypes refers to estimating the genotype of an individual for variants found across the genome, up to many millions. For instance, if a SNP has two alleles, G or T, individuals in a population can carry three possible genotypes: GG, TT, or GT (one allele from paternal and the other from maternal origin). Theoretically, DNA sequencing data can be used to determine which of these three genotypes an individual carries. However, aDNA data are frequently sparse, and genotype calling involves only determining if the individual carries at least one G or at least one T variant.

The genotype data can then be used to address various questions: to infer demographic history (based on tests of affinity between the genotype of an ancient individual and those of other ancient or present-day individuals), to estimate the level of genetic kinship to other individuals, to estimate certain phenotypic traits such as eye colour, or to determine how inbred the individual was. If the aDNA dataset has been produced

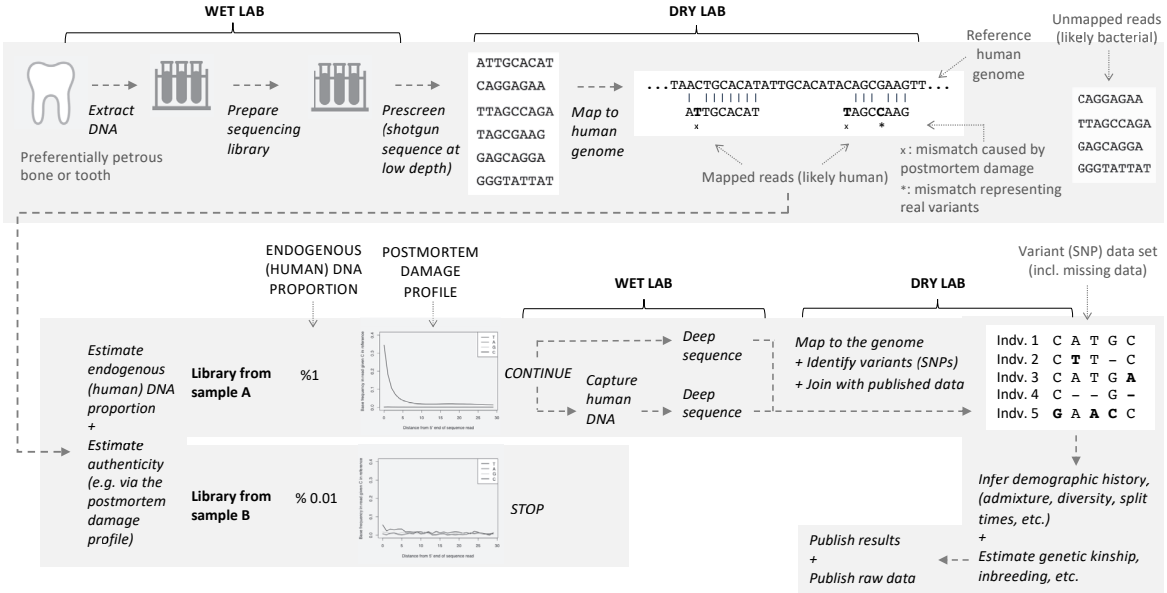

Figure 18.1. The second generation aDNA analysis routine. The first phase of an experiment – pre-screening – aims to estimate the endogenous human DNA proportion and post-mortem damage in the human molecules, as a rough estimate of authenticity. Selected libraries continue to be analysed by direct deep (large volume) sequencing, or hybridisation capture followed by deep sequencing. The resulting data are then used in population genomic analyses for estimating demographic history, genetic kinship levels among individuals, and level of inbreeding, etc. The post-mortem damage figures in the lower panel indicate the proportion of sequenced human DNA molecules that contain a T base at their ends when a C is expected in the reference genome at that same position. Observing a T instead of a C is the most prominent indication of post-mortem damage that accumulates in DNA over time.

from teeth, the DNA molecules that do not map to the human genome can be screened for matches to blood-borne pathogen genomes (Rascovan et al. 2018). Age-at-death of individuals may also be roughly estimated from DNA methylation patterns using specifically processed high quality aDNA datasets (Pedersen et al. 2014).

An important point is the conventional practice of sharing raw and/or processed datasets upon publication among population geneticists, usually by depositing the data in public databases. This enables other researchers to download and compare their new data with previously published datasets and augment knowledge accumulation. As of May 2020, there have already been more than 3,500 ancient human genomes published by various groups, available for analysis.

Sequencing technologies are still evolving, but the experimental procedures described here may still prevail in aDNA laboratories for some time yet because of constraints imposed by the degraded and low-quality nature of aDNA. Instead, we may expect statistical methods to develop even further in the coming years to make even more effective use of the currently available data. The rapidly expanding collection of ancient genomes is expanding the potential uses of the data itself.

Genomic insights into European Neolithisation: demic processes

Since the late 20th century, geneticists in Europe and North America have been tackling the demic versus cultural diffusion question in the context of European Neolithisation. Early studies used genetic data from modern populations (Menozzi et al. 1978; Cavalli-Sforza et al. 1988; Richards 2003), suggesting a role for migration of southeastern origin in the spread of farming across Europe. Later, mitochondrial ancient DNA studies (Haak et al. 2008; Malmström et al. 2009) reported discrepancies between the mtDNA profiles of early Neolithic farmers and contemporaneous European Mesolithic hunter-gatherers, also supporting an ancient migration event. Results from ancient mtDNA analyses from the Neolithic and the Iron Age Southwest Asian groups (Fernández et al. 2014; Yaka et al. 2018) have further been in line with this idea.

However, given the uncertainties pertaining to mtDNA analyses (discussed earlier), it was ancient genomics research using whole genome sequence data that provided the final verdict. These studies consistently showed that European Neolithic farmer communities, across Central Europe, Iberia and Scandinavia, were all genetically highly distinct from European Mesolithic groups from the same localities (Skoglund et al. 2012; Gamba et al. 2014; Lazaridis et al. 2014; Skoglund et al. 2014). It was further shown that early European farmers had genetic profiles highly similar to those of Neolithic and Chalcolithic west Anatolian and Aegean populations (Günther et al. 2015; Mathieson et al. 2015; Cassidy et al. 2016; Hofmanová et al. 2016; Kılınç et al. 2016; Omrak et al. 2016), as would be predicted from the archaeological record. Demic diffusion from the Aegean/Anatolia likely had a major role in driving the Neolithisation of Europe. Genomic data has further confirmed that this human movement process followed two main routes: the land route that reached central Europe and a maritime route that reached Iberia and from there, spread along the Atlantic coast (Brace et al. 2019).

That said, genomic sampling of early farmer populations is currently sparse, and the results of genomic analyses do not suggest that cultural interaction was non-existent or that local foragers were replaced by migrant populations. Genetic profiling of middle Neolithic European populations shows that migrant early farmers in Europe and local hunter-gatherer populations eventually did admix, such that modern Europeans have ancestry from both sources (Lazaridis et al. 2014). Moreover, this admixture may have started quite early, as implied by the finding of a 'genetically Mesolithic' individual buried in an early LBK site in Hungary (Gamba et al. 2014). Finally, genetic data on the Neolithisation of the Baltic region indicates that sedentism and farming there were adopted by locals instead of being brought by migrating farmers (Jones et al. 2017). In fact, as we discuss below, acculturation may be the predominant process in many instances of rapid cultural change.

Genomic insights into Neolithic development in the primary zone: acculturation and exchange

It might appear awkward that genetic studies on European Neolithisation have been more intensive than those on the initial Neolithisation process of Southwest Asia, *inside* the primary zone of incipient Neolithic developments; that is, the Levant, north Mesopotamia, the Zagros Mountains and Central Anatolia. The population dynamics behind the initial evolution of sedentism, farming and herding, and the social re-organisation associated with this transition, have been long discussed (Kuijt 2000b; Price, Bar-Yosef 2011). This discrepancy may partly be attributed to Eurocentric tendencies and the fact that many major aDNA research centres are in Europe. Tighter regulations in southwest Asian countries on archaeological sample export may also have slowed down the influx of archaeological material to European centres. At least equally important is the fact that working with aDNA from southwest Asia is far more formidable than working with European samples, due to relatively high temperatures that promote DNA degradation in the former region.

Nevertheless, multiple groups have recently succeeded in producing genomic data from different regions of the Neolithic primary zone through intensive efforts usually involving processing large numbers of individual remains. The year 2016 saw the publication of the first genome studies on the Neolithic primary zone, including individuals from Central Anatolia (Kılınç et al. 2016), the Zagros (Broushaki et al. 2016; Gallego-Llorente et al. 2016) and both the Levant and Zagros (Lazaridis et al. 2016). This was later followed by re-analyses of these data sets (Kılınç et al. 2017) and recently a new publication on Central Anatolia (Feldman et al. 2019). Overall, this work revealed a number of notable points.

1) At the start of the Holocene, western Eurasia hosted genetically differentiated populations. Before the advent of the full Neolithic package (pre-7000 BCE) we have evidence for three relatively isolated human gene pools in the Levant, in the Caucasus and Zagros regions (fig. 18.2) and in Europe, including mainland and steppe (Kılınç et al. 2017). Of particular note here is the genetic isolation between the Zagros region and the Levant, despite their geographic proximity and evidence for cultural contacts, such as the widespread circulation of Anatolian obsidian in these regions (Goring-Morris, Belfer-Cohen 2011; Ibáñez et al. 2018). It is likely that the gene pools were shaped by the shrinkage of available habitats during the Last Glacial Maximum, driving isolation and differentiation through genetic drift (Brewster et al. 2014).

Meanwhile, Epipalaeolithic and Aceramic Central Anatolians (represented by the populations of Epipalaeolithic Pınarbaşı and Aceramic Boncuklu) shared affinities with all three gene pools (Kılınç et al. 2017; Feldman et al. 2019). In a sense, Central Anatolians were 'in the middle' with connections to all groups. How this Central Anatolian population came into being before the Holocene is an open question that will require studying even earlier genetic material from the region. The genetic profiles of north Mesopotamian and Aegean populations of the pre-7000 BCE period are also as yet unknown.

The studies mentioned here were conducted using genetic data from a small number of individuals from a few settlements, sometimes even a single individual representing a settlement. As mentioned earlier, owing to the fact that the genome of a single individual contains information from that individual's many ancestors, it is possible to make wide inferences with even this limited data, assuming one was not unlucky enough to sample a single fresh migrant in a mainly local population. That said, increasing sample size in the future would facilitate making bolder statements.

2) The Pre-Pottery/Aceramic villagers of the Levant, of the Zagros and of Central Anatolia were, in their genomic profiles, not only distinct from each other, but also closely related to the Epipalaeolithic populations of those same regions (Lazaridis et al. 2016; Feldman et al. 2019). The same pattern is also reflected in the mtDNA profiles of these populations (Chyleński et al. 2019). These results strongly imply that the Neolithic transition within Southwest Asia was largely enacted through cultural interactions between local populations, as previously predicted by material culture studies (Baird 2012). In other words, the inter-regional spread of technology, rituals and domestic species mainly occurred through cultural exchange but not large-scale population dispersal.

Still, this does not rule out any population movement among regions. On the contrary, *D*-tests suggest that the Aceramic Boncuklu population, relative to a single Epipalaeolithic Pınarbaşı individual sequenced, received gene flow from eastern (Zagros and/or Caucasus) regions (Feldman et al. 2019). Feldman and colleagues estimate that 10 per cent of Boncuklu's ancestors may have arrived from the east, although these estimates may vary widely. Of course, it should also be noted that this human movement signal may actually represent gene flow from southeast Anatolia/Mesopotamia, which is not yet genetically profiled, instead of gene flow from the Zagros or the Caucasus.

3) From the Pre-Pottery/Aceramic period to the Pottery/Ceramic Neolithic period, human dispersal between Anatolia and neighbouring regions appears to have gained momentum. Populations of the PPNB Levant (for example, 'Ain Ghazal) show higher genetic affinity to Neolithic Anatolians relative to Natufian groups who lived in the same region 2,000 years earlier (Kılınç et al. 2017). This can be explained by a simple scenario: after the Natufian, certain Anatolian-related groups moved in and mixed with the Levantine locals and therefore PPNB Levantine groups carry more Anatolian variants than did the Natufian groups.

The Pottery Neolithic Central Anatolian population of Tepecik-Çiftlik likewise had closer affinity to Levantine and Zagros populations than did Aceramic Boncuklu people who lived in the same broad region 2,000 years earlier. These changing genetic affinities can also be explained by gene flow into Anatolia from its southern and eastern peripheries. A pattern of accelerating inter-regional gene flow is captured by increasing within-population genetic diversity levels during the Neolithic transition (Kılınç et al. 2016; Lazaridis et al. 2016).

That said, alternative complex scenarios, such as complex population structure (for example, spatial heterogeneity in the population) and variation in admixture tendencies among sub-populations could also explain the observed genomic data without the occurrence of regional admixture events. These more complicated scenarios may be tested in the future when denser data are available.

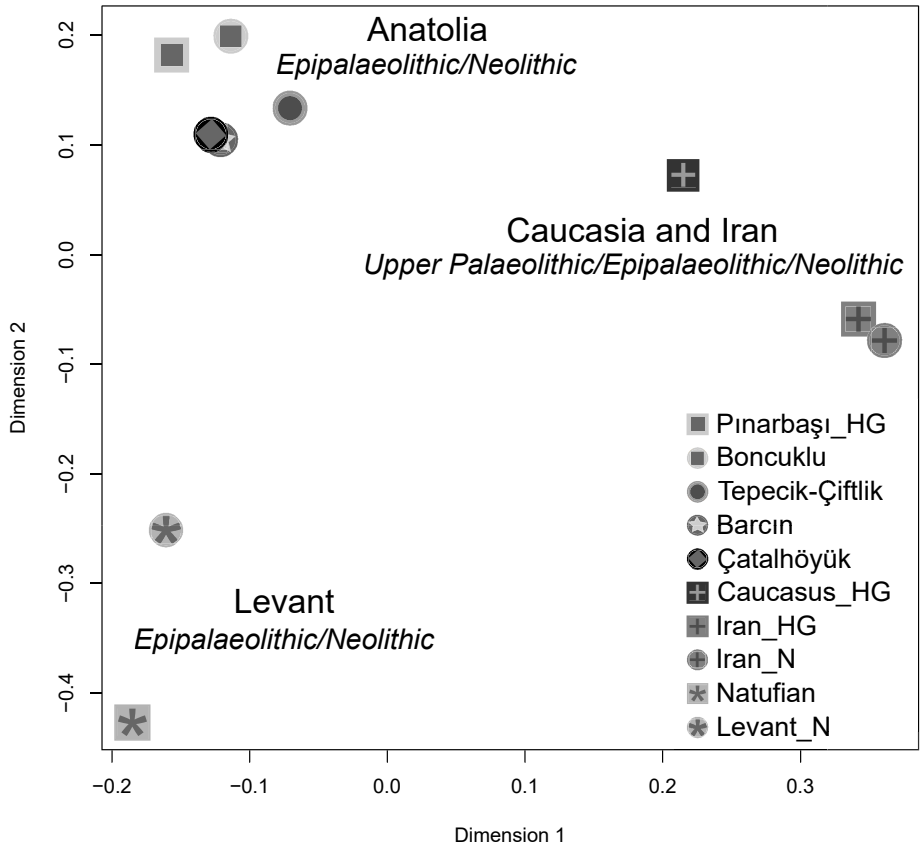

Figure 18.2. Genetic similarities between pre-Neolithic and Neolithic populations of Southwest Asia. The multi-dimensional scaling plot summarises f_3*-statistics that measure genome-wide similarity between pairs of populations (Patterson et al. 2012). The statistics were calculated among regional populations with published genomic data: Pınarbaşı_HG: Epipalaeolithic from Pınarbaşı (Feldman et al. 2019); Boncuklu: Aceramic Boncuklu Höyük (Kılınç et al. 2016; Feldman et al. 2019); Tepecik-Çiftlik: Pottery Neolithic Tepecik-Çiftlik (Kılınç et al. 2016); Barcın: Pottery Neolithic Barcın Höyük (Mathieson et al. 2015; Hofmanová et al. 2016); Caucasus_HG: Upper Palaeolithic Caucasus hunter-gatherer (Jones et al. 2015); Iran_HG: Iran Epipalaeolithic from Hotu Cave (Lazaridis et al. 2016); Iran_N: Iran Pre-Pottery Neolithic from Ganj Dareh, Tepe Abdul Hosein and Wezmeh Cave (Broushaki et al. 2016; Lazaridis et al. 2016); Natufian: Natufian from Raqefet Cave (Lazaridis et al. 2016); Levant_N: Levant Pre-Pottery Neolithic from Ain Ghazal, Mota, Ba'ja and Kfar HaHoresh (Lazaridis et al. 2016; Feldman et al. 2019). Çatalhöyük data are unpublished (Yaka et al. in preparation).*

4) A little-understood point is the source of the Neolithic North Aegean communities that seem to have emerged post-7000 BCE. Neolithic European farmers were genetically closest to these Neolithic North Aegean communities (Barcın in northwest Anatolia and Revenia in northeast Greece) than to any other primary zone population hitherto studied (Kılınç et al. 2017).

So, who were these Neolithic Aegean people? Were they seafaring colonists from the Levant, a composite of Mediterranean and inland Anatolian migrants who merged on the Aegean coast, or local Mesolithic populations acculturated by their maritime or inland connections? The archaeological evidence has remained equivocal (Özdoğan 2011b; Perlès et al. 2011; Reingruber 2011; Baird 2012; Çilingiroğlu, Çakırlar 2013; Horejs et al. 2015), while genetic data now provides interesting clues.

First, the north Aegean population was genetically closer to that of Central Anatolia than to any other potential source population, including the Neolithic Levantine people (Kılınç et al. 2017; Chyleński et al. 2019). This rules out a Levantine-derived colonist origin. Surprisingly, relative to Central Anatolians, the north Aegean Neolithic populations also carried genetic affinities to distant populations, from Europe and the Caucasus to the Levant (fig. 18.2) (Kılınç et al. 2017). The Aegeans thus appeared perhaps 'too diverse' for a migrant group of recent Central Anatolian origin.

We may thus consider an alternative hypothesis to the idea of migration/colonisation of the Aegean by farmers from the Neolithic primary zone. Perhaps Aegean Neolithic societies were largely local forager populations who underwent rapid acculturation via connections with the Neolithic primary zone (Kılınç et al. 2017). This conclusion resonates well with archaeological evidence indicating a heterogeneous cultural background in the Aegean; for example, the intensive use of obsidian from the Aegean island of Melos or the non-existence of Near Eastern traditions of lithic tool production, such as pressure-flaking, before the seventh millennium BCE in the region (Düring 2011b; Perlès et al. 2011; Reingruber 2011; Yakar 2016). If true, acculturation could have had a more prominent role, at least for the early Neolithisation of the Aegean sphere, than currently thought. If so, it becomes plausible to suggest that population movements drove the Neolithisation of Europe from a wider geographical area encompassing the Aegean and Anatolia, rather than merely Anatolia. Full investigation of these scenarios awaits further data, specifically from pre-Neolithic populations in the region.

First genomic insights into Neolithic Çatalhöyük humans

At the METU Ancient DNA Laboratory, genomic analyses of Çatalhöyük human material started in 2015, but it was more than three years before useable amounts of genetic data were obtained. The work was focused on East Mound sub-floor burials. By early 2020, the METU aDNA Laboratory had prepared over 200 Illumina sequencing libraries from 157 individuals' teeth and petrous bones and subjected these to pre-screening (fig. 18.1) in collaboration with Stockholm University; selected libraries were further deep-sequenced after hybridisation capture. In collaboration with METU, the aDNA group at the Adam Mickiewicz University in Poznań (AMUP) also initiated a study on Çatalhöyük genetics focused on mtDNA. This work led to the first DNA-based study of Çatalhöyük humans, published in 2019 (Chyleński et al. 2019).

DNA preservation and the question of adult burial practices

Between 2015 and 2018 the METU team had prepared sequencing libraries based on >20 individuals' remains, and none had provided useful amounts of endogenous DNA (>0.5%). This precluded us from including Çatalhöyük aDNA data in a 2016 publication on Neolithic Anatolian populations, which had included a Boncuklu Höyük (Aceramic Neolithic Central Anatolia) individual with >6% endogenous DNA, and four out of six Tepecik-Çiftlik (Pottery Neolithic Central Anatolia) individuals with endogenous DNA proportions close to 1% (Kılınç et al. 2016). Chyleński et al. also reported low DNA preservation in Çatalhöyük (Chyleński et al. 2019).

This had raised the question of whether Çatalhöyük human bones may have been overall more poorly preserved than those of other Neolithic Central Anatolian sites. However, recent sampling and sequencing efforts showed this not to be the case: the current data from recent pre-screening results suggest no systematic difference among Tepecik-Çiftlik, Boncuklu Höyük and Çatalhöyük in the endogenous DNA proportions (fig. 18.3) (Kruskal-Wallis test $p>0.10$).

Strikingly, analysis of this pre-screening dataset revealed that endogenous DNA proportions clearly depend on the age-at-death of Çatalhöyük individuals, with adults showing significantly lower DNA preservation than sub-adults, including newborns, infants and children (fig. 18.3) (Mann-Whitney U test $p<0.001$). This pattern was wholly unexpected, as it is generally believed that adult bones preserve DNA better due to their compactness. In fact, early sampling efforts in Çatalhöyük had for this reason tried to avoid subadult material, leading to the earlier observation on low DNA preservation at this site.

The reason for lower adult vs subadult DNA preservation at Çatalhöyük is as yet unknown, as well as whether it applies to other Neolithic Anatolian settlements. It is tempting to speculate that unique burial traditions were practised mainly for adults, such as collective burials, secondary burials and interventions, that intervened in the decomposition process in ways such as defleshing (Pilloud et al. 2016a; Haddow, Knüsel 2017) and could underlie DNA preservation differences. Nevertheless, the as yet small sample size and high variation in DNA preservation rates, even among individuals buried in the same location, preclude reaching an absolute conclusion at this point.

Çatalhöyük genetic and genomic data

In the first DNA-based study of Çatalhöyük people published, the aDNA laboratory at AMUP generated whole mitochondrial genomes using pre-screening followed by mitochondrial DNA capture from ten individuals (Chyleński et al. 2019). All were burials within four neighbouring buildings (B.80, B.89, B.96, B.97) belonging to Levels South O-North G, 6700–6500 BCE. Seven were subadults.

As of March 2020, the METU aDNA Laboratory, in collaboration with the AMUP and Stockholm University (SU) laboratories, produced low-coverage genomic data from 12 individuals from the levels South K, M and N

and North G (B.17, B.50, B.89, B.91, B.114) and two from Late-Neolithic phases of Çatalhöyük (TP and TPC). Seven of these individuals' remains were radiocarbon dated and all dates were consistent with the archaeological contexts (ca 6700–6500 BCE for Middle phase and ca 6400–6100 BCE for Late phase individuals).

The genome data were produced using both shotgun sequencing from the pre-screening experiments and whole genome capture experiments followed by deep sequencing. Across the studied individuals the genome coverages range from 0.01x to 0.27x; in other words, one can observe only 1% to 27% of the genome-wide variants of each individual. Despite these low proportions, given the presence of millions of common variants across the human genome, it is still possible to extract information on tens of thousands of variants per individual. This amount of information can be sufficient for simple demographic analyses (for example, comparing affinities of a Çatalhöyük individual to one from the Levant or from Boncuklu in a *D*-test). It also permits estimation of close biological kinship.

All except one of the 14 individuals with sufficient genome data were subadults; the exception was one Late Neolithic individual. Ten were found to be genetically female. Because the sex of subadults cannot be assessed reliably from study of their skeletal remains, we cannot assess the overall match with the anthropological sex assignment.

DNA-based kinship patterns among Çatalhöyük co-burials

The most striking result from the Chyleński et al. study was that, among the ten individuals whose mitochondrial genomes had been sequenced, all carried distinct mtDNA lineages (Chyleński et al. 2019). That is, all ten individuals had different mothers, maternal grandmothers, etc., despite being interred within four neighbouring buildings.

This intriguing result could be explained in two ways: Çatalhöyük burial traditions might have been strictly patrilocal, a possibility that has previously been raised based on the study of similarities in dental markers across larger Çatalhöyük samples (Pilloud, Larsen 2011). Alternatively, individuals buried in the same buildings may not have been biologically closely related kin.

The genomic data currently available (Yaka et al. in preparation), although sparse, supports the latter notion. Out of the 14 individuals we generated genomic data for, ten were buried in B.17, B.50 and B.114. We studied genetic kinship among a total of 14 such co-buried pairs,

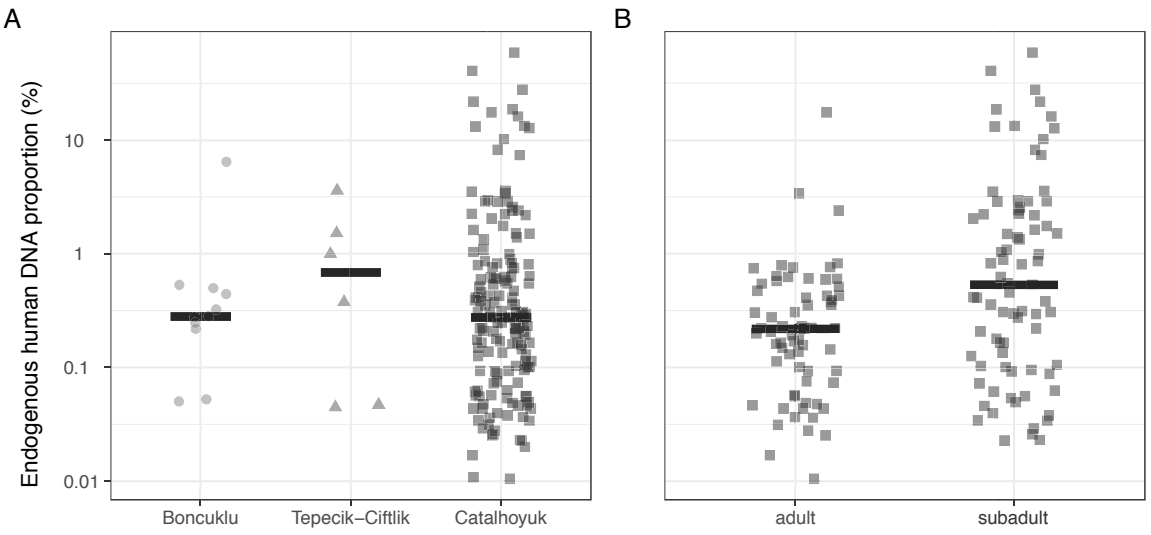

Figure 18.3. Endogenous DNA proportions in Çatalhöyük. (A) Proportions calculated from Boncuklu (n=9), Tepecik-Çiftlik (n=6) and Çatalhöyük (n=157) individuals' bone and teeth samples. (B) Proportions calculated for adult (n=67) and subadult (n=79) Çatalhöyük individuals. Adolescents were not included. Endogenous DNA proportions were measured as the percentage of shotgun sequenced DNA molecules from pre-screening experiments mapping to the reference human genome. Each individual is represented by a single observation. The highest proportion was used for a minority of cases where multiple libraries had been produced per individual. The horizontal black lines indicate the median per group.

where the polymorphism data would allow us to detect up to the third degree of relatedness (that is, cousins or avuncular relations).

Out of the 17 pairs scanned, only one pair of female subadults were found to be siblings (Yaka et al. in preparation). None of the other co-buried subadults were closely related. Although the data are as yet too limited to reach a clear conclusion, they point to the exciting prospect that non-biological social kinship structures influenced co-burial at Çatalhöyük (cf. Pilloud, Larsen 2011) and possibly other contemporaneous Neolithic societies in southwest Asia. We discuss the implications of this finding in the last section.

Population-level similarities between Çatalhöyük and other southwest Asian Neolithic groups

Preliminary population genetics and kinship analyses have revealed a number of patterns that we summarise here (Yaka et al. in prep.):

1) Treated together as a population, the 14 Çatalhöyük individuals with genomic data belong to the Early Holocene central/west Anatolian gene pool. Specifically, the Çatalhöyük population is genetically close to individuals from other Epipalaeolithic and Neolithic Central Anatolian populations (Pınarbaşı, Boncuklu Höyük, Tepecik-Çiftlik) and to those from the northwest Anatolian Neolithic Barcın Höyük, but distinct from Levantine or Zagros populations (fig. 18.2). This overall pattern is supported by *D*-tests (Yaka et al. in prep.). It is also consistent with the report by Chyleński et al., who found that the overall haplogroup composition of this sample was similar to that of other Central Anatolian and northwest Anatolian groups (Chyleński et al. 2019).
2) There is limited genetic differentiation among these Çatalhöyük individuals and no conspicuous difference between Middle- and Late-period Çatalhöyük individuals. That said, we cannot yet rule out possible subtle differences in the genetic profiles between Middle- and Late-period Neolithic Çatalhöyük populations (which could be caused by gene flow into Çatalhöyük from distant regions). More genomes will be needed to address this.
3) In *D*-tests, the Çatalhöyük population appears equally related to Aceramic Boncuklu from the Konya Plain and the Pottery Neolithic Tepecik-Çiftlik from the Cappadocia region. Meanwhile, it appears to carry more Levantine variants compared to the Boncuklu sample, similar to that reported for Tepecik-Çiftlik (Kılınç et al. 2017). This result is compatible with some level of Levantine gene flow into Central Anatolia between 7500 and 6500 BCE.

Conclusion and future directions

Ancient genomics analyses are shedding light on various aspects of the Neolithic transition in Southwest Asia. We can already see that, during the initial development of sedentism and later farming, populations on the east, west and northwest borders of the Neolithic primary zone largely remained genetically distinct, with limited gene flow among them. However, this is only a first glimpse and the picture remains full of gaps. From genetic data we know nothing, or close to nothing, about human movement to or from Neolithic Mesopotamia. The Aegean question – if local Neolithisation occurred in this region – is open. The colonisation of Cyprus also remains an enigma. Still, we can underline some tentative outcomes. First, the Neolithisation across different regions of the Fertile Crescent, including Central Anatolia, involved limited gene flow across regions (Kılınç et al. 2016; Lazaridis et al. 2016). Also, the Neolithisation of West Anatolia and the general Aegean sphere do not seem to be related to migrations from the Levant (*contra* Horejs et al. 2015), whereas during the PPNB some human movement between Central Anatolia and the Levant seems to have occurred. These broad human movement patterns also involve the ancestry of the Çatalhöyük population, which was genetically highly similar to the Aceramic Boncuklu Höyük population, albeit slightly more diverse due to past admixture (Yaka et al. in preparation).

We further remain largely ignorant regarding the role of biological kinship in these early Neolithic societies. What role did biological relatedness have in the organisation of these communities? Social kinship can take on diverse forms and genetic relatedness is not a prerequisite (Sahlins 2013; Johnson, Paul 2016). Despite the widespread assumption that forager groups are genetically kin-based, recent research has shown that within-group genetic relatedness is conspicuously low in many modern-day foragers (Hill et al. 2011) and post-marital residence is not necessarily sex-biased (Dyble et al. 2015). In contrast, in many farmer communities, genetic kinship is central to social organisation and patrilocality is predominant (Stone 2014). Recent ancient genomics work on Upper Palaeolithic Siberian hunter-gatherers (Sikora et al. 2017) and LBK and Bronze Age farmers from Europe (Mittnik et al. 2019; Sánchez-Quinto et al. 2019) also point in the same direction: a limited role of genetic kinship in pre-farming societies and predominant patrilineal patterns once farming had been established. If social organisation of pre-Neolithic societies was indeed akin to those of modern-day hunter-gatherers, then genetic kin-based and patrilocal/patrilineal traditions could have emerged either early or late and variably from one region or community to another during the Neolithic transition.

At least for the case of Çatalhöyük, our preliminary results suggest that co-buried individuals in the same building may have been frequently not closely related. This, along with earlier observations based on morphological data (Pilloud, Larsen 2011) and mitochondrial DNA analyses (Chyleński et al. 2019), raises doubts regarding a fully kin-centred social organisation at Çatalhöyük. Indeed, they contradict the hypothesis that extended kinship-based families lived and were buried within the Neolithic houses (Kuijt et al. 2011). Instead, the results appear more consistent with the 'history house' concept proposed by Ian Hodder and Peter Pels (Hodder and Pels 2010), where biologically non-kin can become related, that is, can become 'practical kin', through their identification with the physical structure of a 'house', which in fact is the primary agent of making and naturalising relations between people and place. Hodder suggests that those people who were buried in a particular building need not even have lived within that architectural structure and may involve 'adoptive, foster or fictive kin held together by memory and history making' (Hodder 2016). At the same time, based on Jessica Pearson's studies (Pearson 2013), it is possible to say that a social 'house' may incorporate more than one building, linked through co-eating and co-burying. Now, the goal before us is to generate a full-scale map of population dynamics and traditions during the Neolithic, through comprehensive and systematic sampling for ancient genomics and integrated analysis of this data with archaeological and bioarchaeological datasets.

Acknowledgements

The authors acknowledge support from TÜBİTAK (grant no. 117Z229), the European Research Council (ERC) (grant no. ERC-772390 NEOGENE), the National Genomics Infrastructure in Stockholm funded by Science for Life Laboratory, the Knut and Alice Wallenberg Foundation and the Swedish Research Council, and SNIC/Uppsala Multidisciplinary Center for Advanced Computational Science for assistance with massively parallel sequencing and access to the UPPMAX computational infrastructure. This research received financial support from the French State via the 'Investments for the Future' framework programme, Initiative d'Excellence de l'Université de Bordeaux (IdEx), (Award No. ANR-10-IDEX-03-02 to CJK).