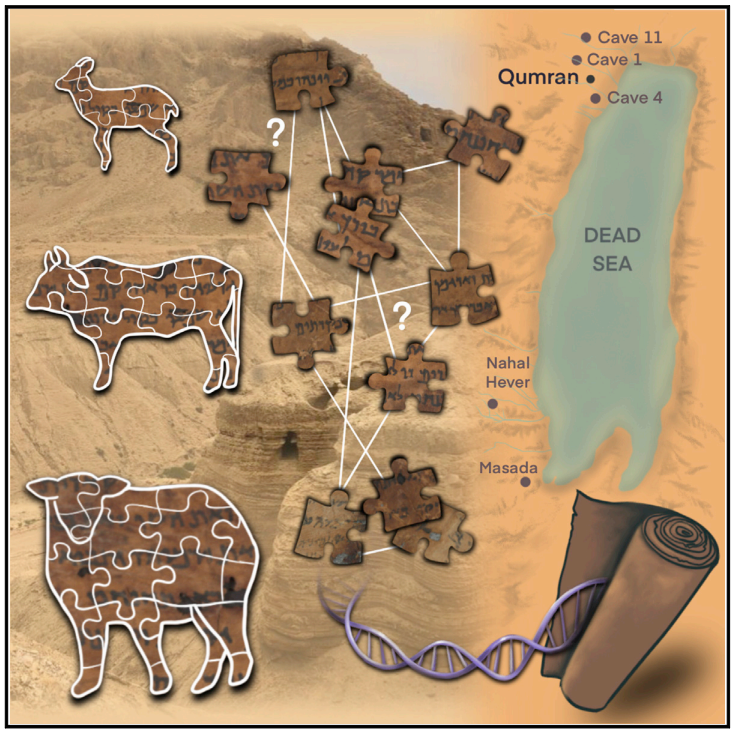
Cell

Article

Illuminating Genetic Mysteries of the Dead Sea Scrolls

Graphical Abstract



Highlights

- Dead Sea Scrolls fragments made of animal skins are classified based on ancient DNA
- Disambiguating the debated relationship between fragments reveals new insights
- Paleogenomics finds that some scriptural scrolls were brought from outside Qumran
- Qumran and Masada scrolls represent the broad cultural matrix of Greco-Roman Judea

Authors

Sarit Anava, Moran Neuhof, Hila Gingold, ..., Noam Mizrahi, Mattias Jakobsson, Oded Rechavi

Correspondence

noam.m.mizrahi@gmail.com (N.M.), mattias.jakobsson@ebc.uu.se (M.J.), odedrechavi@gmail.com (O.R.)

In Brief

Ancient DNA extraction from Dead Sea Scroll fragments made of animal skin is used to determine fragment origin and show unification of scrolls from distinct geological locations, highlighting the potential for genetics to illuminate the history of archaeological objects.

Anava et al., 2020, Cell *181*, 1–14
June 11, 2020
https://doi.org/10.1016/j.cell.2020.04.046

CellPress

Please cite this article in press as: Anava et al., Illuminating Genetic Mysteries of the Dead Sea Scrolls, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.046

Cell

CellPress

Article

Illuminating Genetic Mysteries of the Dead Sea Scrolls

Sarit Anava,[1,10] Moran Neuhof,[1,10] Hila Gingold,[1,10] Or Sagy,[1,10] Arielle Munters,[2] Emma M. Svensson,[2] Ebrahim Afshinnekoo,[3,4,5] David Danko,[3,4,5] Jonathan Foox,[3,4,5] Pnina Shor,[6] Beatriz Riestra,[6] Dorothée Huchon,[7,8] Christopher E. Mason,[3,4,5] Noam Mizrahi,[9,11,*] Mattias Jakobsson,[2,11,*] and Oded Rechavi[1,11,12,*]

[1]Department of Neurobiology, George Wise Faculty of Life Sciences, Tel-Aviv University, Tel-Aviv 6997801, Israel
[2]Human Evolution, Department of Organismal Biology and SciLife Lab, Uppsala University, 751 05 Uppsala, Sweden
[3]Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY 10065, USA
[4]The HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine, Weill Cornell Medicine, New York, NY 10065, USA
[5]The WorldQuant Initiative for Quantitate Prediction, New York, NY 10065, USA
[6]Dead Sea Scroll Projects, Israel Antiquities Authority, Jerusalem 91710, Israel
[7]Steinhardt Museum of Natural History and Israel National Center for Biodiversity Studies, Tel-Aviv University, Tel-Aviv 6997801, Israel
[8]Department of Zoology, George S. Wise Faculty of Life Sciences, Tel-Aviv University, Tel-Aviv 6997801, Israel
[9]Department of Biblical Studies, The Lester and Sally Entin Faculty of Humanities, Tel Aviv University, Tel-Aviv 6997801, Israel
[10]These authors contributed equally
[11]These authors contributed equally
[12]Lead Contact
*Correspondence: noam.m.mizrahi@gmail.com (N.M.), mattias.jakobsson@ebc.uu.se (M.J.), odedrechavi@gmail.com (O.R.)
https://doi.org/10.1016/j.cell.2020.04.046

SUMMARY

The discovery of the 2,000-year-old Dead Sea Scrolls had an incomparable impact on the historical understanding of Judaism and Christianity. "Piecing together" scroll fragments is like solving jigsaw puzzles with an unknown number of missing parts. We used the fact that most scrolls are made from animal skins to "fingerprint" pieces based on DNA sequences. Genetic sorting of the scrolls illuminates their textual relationship and historical significance. Disambiguating the contested relationship between Jeremiah fragments supplies evidence that some scrolls were brought to the Qumran caves from elsewhere; significantly, they demonstrate that divergent versions of Jeremiah circulated in parallel throughout Israel (ancient Judea). Similarly, patterns discovered in non-biblical scrolls, particularly the *Songs of the Sabbath Sacrifice*, suggest that the Qumran scrolls represent the broader cultural milieu of the period. Finally, genetic analysis divorces debated fragments from the Qumran scrolls. Our study demonstrates that interdisciplinary approaches enrich the scholar's toolkit.

INTRODUCTION

The Dead Sea Scrolls (DSS) are a corpus of over 25,000 fragments comprising the remains of about 1,000 ancient manuscripts found in various sites in the Judean Desert mostly between 1947 and the 1960s (Figure 1A) (Collins, 2012; Davies et al., 2002). The largest collection comes from 11 caves in the vicinity of Qumran, along the northwest shore of the Dead Sea. Additional scrolls were found in Masada, Naḥal Ḥever, Naḥal Ṣe'elim, Wadi Murabba'at, and a few other sites (Figure 1A). The Qumran fragments are dated to the Hellenistic-Roman period (or late Second Temple times), i.e., third century BCE to first century CE.

The understanding of the content of any DSS manuscript could drastically change with the addition, removal, or reordering of fragments assigned to it (e.g., Ratzon and Ben-Dov, 2017; Yuditsky, 2017). Thus, much collective work has been invested in sorting out the fragments based primarily on visual resemblance and paleographic analysis of their scribal hands. However, the fragmentary nature and variable state of preservation of the various fragments sometimes result in uncertain identification, joining, and reconstruction of numerous pieces (see example in Figure 1B).

Another set of challenges posed by the DSS is their historical, social, and cultural contextualization. Many scholars agree that some of the texts found in the Qumran caves reflect the ideology and social structure of a sectarian movement of Jews, a group of whom resided in Qumran during the late Second Temple period (e.g., Collins, 2010; Ullmann-Margalit, 2006). However, it is similarly clear that not all Qumran scrolls were written or copied by sectarian scribes (e.g., Dimant, 2014). Distinguishing between "sectarian" and "non-sectarian" texts or between manuscripts potentially copied "locally" or brought from elsewhere has far-reaching implications for proper understanding of the findings.

Figure 1. DSS Fragments Were Found in Different Locations in Different Sizes and Shapes

(A) Map of the Judean Desert. The locations where the DSS were discovered are shown, as well as those of four of the fragments analyzed in this study. (B) Shown are the suggested placings of the different fragments attributed to DSS 4Q404. The three fragments analyzed in this paper, frgs. 5, 8, and 9, are marked. Images are courtesy of the Leon Levy Dead Sea Scrolls Digital Library. Israel Antiquities Authority. Photos: Shai Halevi. See also Table S1.

Were the literary works and ideologies reflected in the scrolls confined to a small and secluded group? Or, alternatively, were these ideas influential on a broader scale, throughout the land? Importantly, the late Second Temple period was the formative era in the history of classical Judaism and nascent Christianity (Brooke, 2005; VanderKam and Flint, 2005). Thus, classification and sub-grouping of texts found in the Qumran caves sheds light on a whole range of historical developments.

In particular, the scriptural scrolls found in the Qumran caves are the oldest known copies of books of the Hebrew Bible and thus offer unprecedented insight into the very process of their formation, early dissemination, and reception. For modern biblical criticism, in contrast to traditional religious tenets, the Hebrew Bible represents the cumulative output of many generations of scribes who authored, transmitted, and, at the same time, constantly revised and interpreted the various texts. The conflicting views can ideally be decided on the basis of empirical evidence, such as manuscripts of biblical books that reflect different stages in their literary history, thereby proving that such works were formed gradually and incrementally rather than by a one-time revelation. Although the DSS contribute greatly to the unveiling of the textual history of many biblical books (Tov, 2012), perhaps the most striking opportunity for testing this question is furnished by the book of Jeremiah, which is extant in different versions corresponding to the long Masoretic text (the Jewish received version) and the short Septuagint (an ancient translation into Greek, eventually adopted by Christians).

The Qumran scrolls also include copies of numerous non-scriptural works. Some of them were previously known through other channels of transmission (the Apocrypha and Pseudepigraph of the Old Testament), whereas others became known for the first time in two millennia thanks to the discovery of the DSS. These works, too, present a whole host of historical riddles. In particular, we have focused on the *Songs of the Sabbath Sacrifice*, a liturgical composition found in multiple copies, indicating its importance in ancient times. This composition is also unique in being the only previously unknown, non-scriptural work whose copies were found outside of Qumran; in addition to eight copies from Qumran Cave 4 and one from Cave 11, there is also a tenth copy found in Masada. The *Songs of the Sabbath Sacrifice* (*Songs*) is a highly intriguing literary piece whose (quasi)mystical content and radical poetic style foreshadow revolutionary developments in Hebrew creativity that fully materialized only many centuries later (Abusch, 2003; Alexander, 2010; Mizrahi, 2013; Schäfer, 2006), but its historical contextualization is still much of a mystery. The socio-religious background of the *Songs* is much contested in scholarship, and the discovery of the Masada copy complicated the debate even further (e.g., Newsom, 1990). Did the Qumran sectarians reach Masada? Or was the work popular among circles that transcend the community that forms the social background of the texts deposited in the Qumran caves?

Finally, most DSS fragments were not excavated orderly *in situ* but rather acquired *post factum* via antiquity dealers. As a result, the original place of finding of some fragments remains unknown or uncertain, thus affecting their historical comprehension. For instance, most of the Qumran scrolls are literary in nature, but

Please cite this article in press as: Anava et al., Illuminating Genetic Mysteries of the Dead Sea Scrolls, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.046

a subset of fragments allegedly found in Cave 4 differs markedly in being legal documents, the like of which were found in other sites throughout the Judean Desert. Is there a scientific method for establishing the original whereabouts of such fragments?

Here we show that a paleogenomic approach can shed new light on a range of issues pertaining to DSS history, including the compositional history of the biblical book of Jeremiah, the dissemination of the *Songs*, and the provenience of the scrolls at large. Because most of the scrolls were written on processed animal skin, we used DNA sequencing to "fingerprint" (or "genotype") fragments based on their genetic signature. The genetic evidence suggests that the manuscripts found in the Qumran caves reflect the scribal and textual diversity that prevailed in this period with regard to transmission of scriptural and non-scriptural literature.

RESULTS

The original place of finding of some fragments remains unknown or uncertain, including their relationship to the site of Qumran. Moreover, the DSS have disintegrated into thousands of fragments whose mutual relationship is often ambiguous or even contested. Thus, scholarship faces two fundamental questions. Which fragment belongs to which scroll? What are the origins of the various scrolls?

To advance the classification of the scrolls and scroll fragments, we sampled and analyzed tiny fragments from scrolls discovered at various sites (Table S1). The selection of fragments was constrained by several factors. As a world cultural heritage, the DSS are now carefully preserved by the Israel Antiquities Authority (IAA), and their physical condition, often extremely fragile, is extensively monitored. Therefore, to minimize damage to these priceless fragments, the IAA conservators performed sampling with extreme caution, considering which fragments could be sampled at all. Tiny amounts of scroll materials were obtained from scroll pieces or, when more sensitive fragments were concerned, from crumbs that fell off or were scraped off the uninscribed back side of the fragments (STAR Methods; Figure 2A). To investigate features typical to scrolls unearthed in different locations, we also examined other leather objects, such as remains of sandals, waterskins, and a garment, that were discovered at various sites. In addition to sampling ancient specimens, we were allowed access to modern parchments prepared for the IAA. These parchments, used as controls, were produced by a professional scribe and treated according to traditional practices for IAA conservation and preservation research purposes (Figure 2B).

Despite the antiquity and degraded state of most scrolls, it was demonstrated 20 years ago that it is possible to amplify animal DNA from the DSS using PCR (Kahila Bar-Gal, 2000). Following recent studies (Teasdale et al., 2015), we used deep sequencing and obtained workable levels of DNA from 286 libraries (with 0.01×/0.006× mean/median coverage of the target animal genomes, respectively; STAR Methods; Table S2; Figures 2A and S1A), enabling study of 39 specimens of minimal size, including 26 DSS fragments. Sequencing the highly degraded endogenous ancient DNA (aDNA) from the DSS is further complicated by high levels of contamination. In addition to endogenous aDNA, not surprisingly, we found ample modern DNA of human origin because humans have handled the DSS since ancient times and most intensively in the last 70 years. The human sequences we detected did not display typical damage patterns and had a longer fragment length compared with that expected of authentic aDNA (Dabney et al., 2013a; Sawyer et al., 2012; Skoglund et al., 2014a). In contrast, given enough sequence data, the reads from ancient samples mapping to animal DNA display DNA fragmentation and higher cytosine deamination at fragment termini that extend into the fragment while showing gradual decline (Figures S1B, S1C, and S2A–S2C; Table S3). The degradation patterns fit those expected from aDNA (Dabney et al., 2013a; Sawyer et al., 2012; Skoglund et al., 2014a) and are like the patterns we obtained from aDNA simulations (using the gargammel software; Figure S2D; STAR Methods).

Species Identification Shedding Light on Scriptural Scrolls

We hypothesized that if skins from diverse types of animals were used for the preparation of different scrolls, then species identification could allow primary sorting of related fragments. Among the ancient material sampled (scrolls, sandals, waterskins, and a garment), we identified skins of three species: domestic sheep, cow, and goat. Upon mapping the sequenced reads to the mitochondrial (mt) genomes of domestic animal species common in the Near East, all identified DSS samples but two (4Q70 and 4Q72b) were found to be made of sheep skin (highest coverage when mapping to *Ovis* mitochondria; Figures 3, 4A–4C, and 4F; Table S4). Note that the DSS are designated by abbreviations that mark their place of finding (e.g., 4Q = Qumran cave 4, 11Q = Qumran cave 11, Mas = Masada) and inventory number or content (e.g., 4Q70 = 4QJer^{a} = first copy of the scriptural book of Jeremiah from Qumran cave 4).

We performed multiple independent validation procedures, such as BLAST searches, FastQScreen, extensive molecular phylogenetic analyses (Hassanin et al., 2012), and the standard pipeline for human aDNA, which confirmed our results (see Table S4 and STAR Methods for details). aDNA simulation achieved similar results (Figures 4D and 4E; STAR Methods).

We initially identified the two Jeremiah specimens that did not originate in sheep, 4Q70 and 4Q72b, as made of cow skin. These fragments represent two of four (of a total of six) copies of the biblical book of Jeremiah that were found in the Qumran caves: 4Q70 (4QJer^{a}), 4Q71 (4QJer^{b}), 4Q72a (4QJer^{d}), and 4Q72b (4QJer^{e}). (Two additional copies, 2Q13 [2QJer] and 4Q72 [4QJer^{c}], could not be sampled at that time.) The original relationship between some of these fragments has been unclear. 4Q71, 4Q72a, and 4Q72b had originally been assigned to the same scroll, but it was later shown that they differ textually (Tov, 1989), and they were eventually published as distinct manuscripts (Tov, 1997). We further investigated the animal origin of 4Q70, 4Q72a, and 4Q72b by sequencing enriched mt fragments (STAR Methods; Table S5; Figure S3A). The analysis based on mt enrichment confirmed the results of the analysis based on shotgun sequencing and revealed that 4Q72b was indeed made of *Bos taurus* (cow), whereas 4Q71 and 4Q72a (thought to come from the same scroll as 4Q72b) are more likely to originate in *Ovis aries* (sheep).

Please cite this article in press as: Anava et al., Illuminating Genetic Mysteries of the Dead Sea Scrolls, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.046

Figure 2. Illustration of Various Samples Used in This Study, Including Miniscule Amounts of Material Used for Sequencing of the DSS, IAA Parchments, and Waterskins

(A) Illustration of the miniscule amount of material used for sequencing of the DSS. (i) Shown is frg. 4Q404 frg. 5v, containing only powder scraped off the scroll. To avoid harming the rare fragment, extraction was carefully done on almost invisible fragments. (ii) Shown is frg. 4Q405, sampled in this research. The fragment consists of tiny scroll pieces with a size and quality similar to a few other DSS fragments used in this study.

(B) Two of the IAA parchments sampled in our analysis. (i) Parchment made of cow skin. (ii) Parchment made of sheep skin.

(C) The two waterskins sampled in our analysis. (i) 1996-9257, notated as WS7. (ii) 1996-9104, notated as WS4. Specimen WS(4or7) was successfully matched to WS7.

Waterskin images in (C) are courtesy of The Israel Museum, Jerusalem. Photos by Elie Posner.

Since scrolls were made of leather sheets of similar material features, such as size and visible properties (e.g. 1QIsa[a]; see below), our working hypothesis is that they usually comprised sheets made of only one animal species. Indeed, our own analysis of fragments issuing from different sheets of the same scroll (4Q37 [4QDeut[j]]; Discussion) confirms that they originate in the same species. Therefore, 4Q72b is unlikely to come from the same scroll as 4Q71 and 4Q72a. The latter two are made of sheep, but their relatedness is relatively low (bootstrap value 40/100) and, hence, inconclusive. Our results also suggest that 4Q70—which is clearly a different manuscript—is made of cow skin. However, this specimen is very poorly preserved and therefore more challenging to study.

The Judean Desert has not been well suited for cow husbandry for at least the last two millennia, and there is no evidence for the processing of cow skins in Qumran (Bouchnick, 2016). Therefore, 4Q70 and 4Q72b most probably originate in another part of the country. This conclusion is corroborated by the fact that, paleographically, 4Q70 dates to the late third or early second century BCE (Yardeni, 1990), long before the deposition of the scrolls in the Qumran caves. Moreover, 4Q70 is textually distinct because it preserves an intermediate compositional stage. Its original text was corrected by a later scribe, and although the nature of these scribal interventions has been debated, the most conspicuous case likely reflects textual adaptation of an originally independent text into the longer version known from the Masoretic text; a short text (4Q70, fragment [frg.] 4i [= column III]) was expanded by an insertion in the margin. The additional paragraph (Jeremiah 7:30–8:3)

Please cite this article in press as: Anava et al., Illuminating Genetic Mysteries of the Dead Sea Scrolls, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.046

CellPress

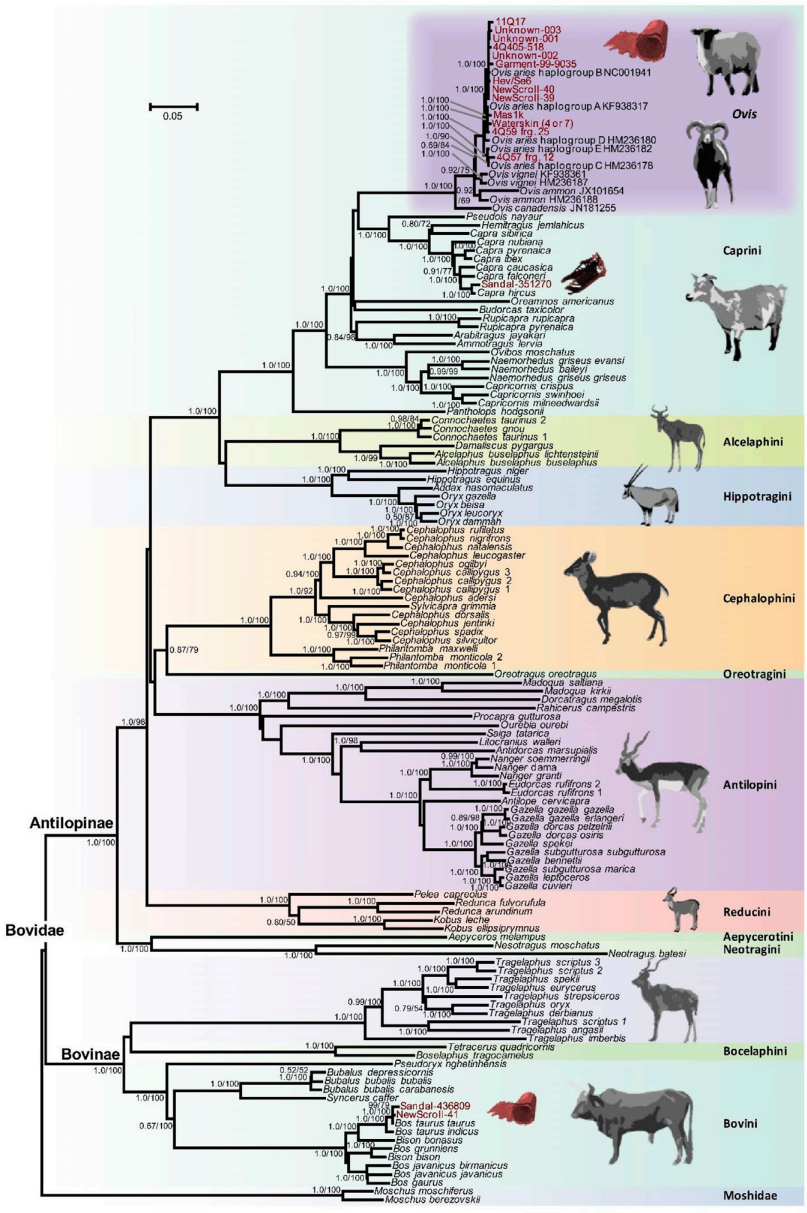


Figure 3. DSS mt Sequences Are Nested with Maximal Support within the Species *Ovis aries*, and Additional Specimens Are Assigned to Other Species

Phylogenetic tree generated based on the Hassanin et al. (2012) dataset from a matrix of 14,867 nucleotide positions and 140 taxa using Bayesian inference under the CAT-GTR model (as implemented in Phylobayes version 4.1). Bayesian posterior probabilities above 0.5/maximum likelihood (ML) bootstrap supports above 50% (using 100 rapid-bootstrap computations) are given near the corresponding node. ML and Bayesian analyses gave similar topologies. Only high-coverage specimens (see STAR Methods for details) were included in this analysis. See also Figure S3 and Tables S4 and S5.

amplifies the prophetic admonition and has become part of the longer, received version (Davis, 2018). Tellingly, long before the discovery of the DSS, critical scholars have argued, on literary grounds, that this paragraph is not well integrated into its present context (e.g., Duhm, 1901, page 84).

The genetic distinction between the Jeremiah fragments has important and broader historical implications. It is revealing that the genetic analyses disassociate fragments that reflect different versions of the book (Discussion). Moreover, at least two of the Jeremiah scrolls (4Q70 and 4Q72b) were probably brought to the Qumran caves from the outside, suggesting that the textual pluriformity of the Jeremiah scrolls applies to Judea in general. Although scholars have previously entertained the theoretical possibility of such "external" manuscripts, our findings not only supply an independent argument in favor of this reasoning but also allow us to pinpoint precisely which manuscripts are the most likely candidates to be identified as scrolls copied locally vs. those brought from elsewhere.

Genetic Differentiation among Sheep-Made Scrolls

Most of the ancient scroll specimens we were able to identify with high confidence were made of sheep (Figures 4A–4C). To distinguish between them, we examined the diversity of haplogroups in the different samples by performing phylogenetic analyses examining mt genomes of 126 sheep breeds (Figure S3B). *Ovis aries* has five different haplogroups, three of which (A, B, and C) are known to populate the area of the Near East (Lv et al., 2015; Meadows et al., 2007). Only these three haplogroups were discovered among our samples, and their distribution proves helpful for reconstructing the provenance of individual scrolls. We sampled four (of ten) copies of the non-biblical *Songs*: 4Q404, 4Q405, 11Q17, and Mas1k (for some of these, we were allowed to sample more than one fragment). Mas1k is the only copy of the work discovered outside of the Qumran caves, in Masada. Tellingly, it also differs from the others in terms of its haplogroup; unlike other sheep samples that belong to haplogroup B, Mas1k belongs to haplogroup A (see Figures 3 and S3B; as well as the "Generation of a sheep phylogenetic tree" subsection of the "Method 2 for identifying species" section in STAR Methods).

Within a given species, mt sequences contain only a very limited number of single-nucleotide polymorphisms (SNPs). We therefore also analyzed sheep sequence variation in the nuclear

Please cite this article in press as: Anava et al., Illuminating Genetic Mysteries of the Dead Sea Scrolls, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.046

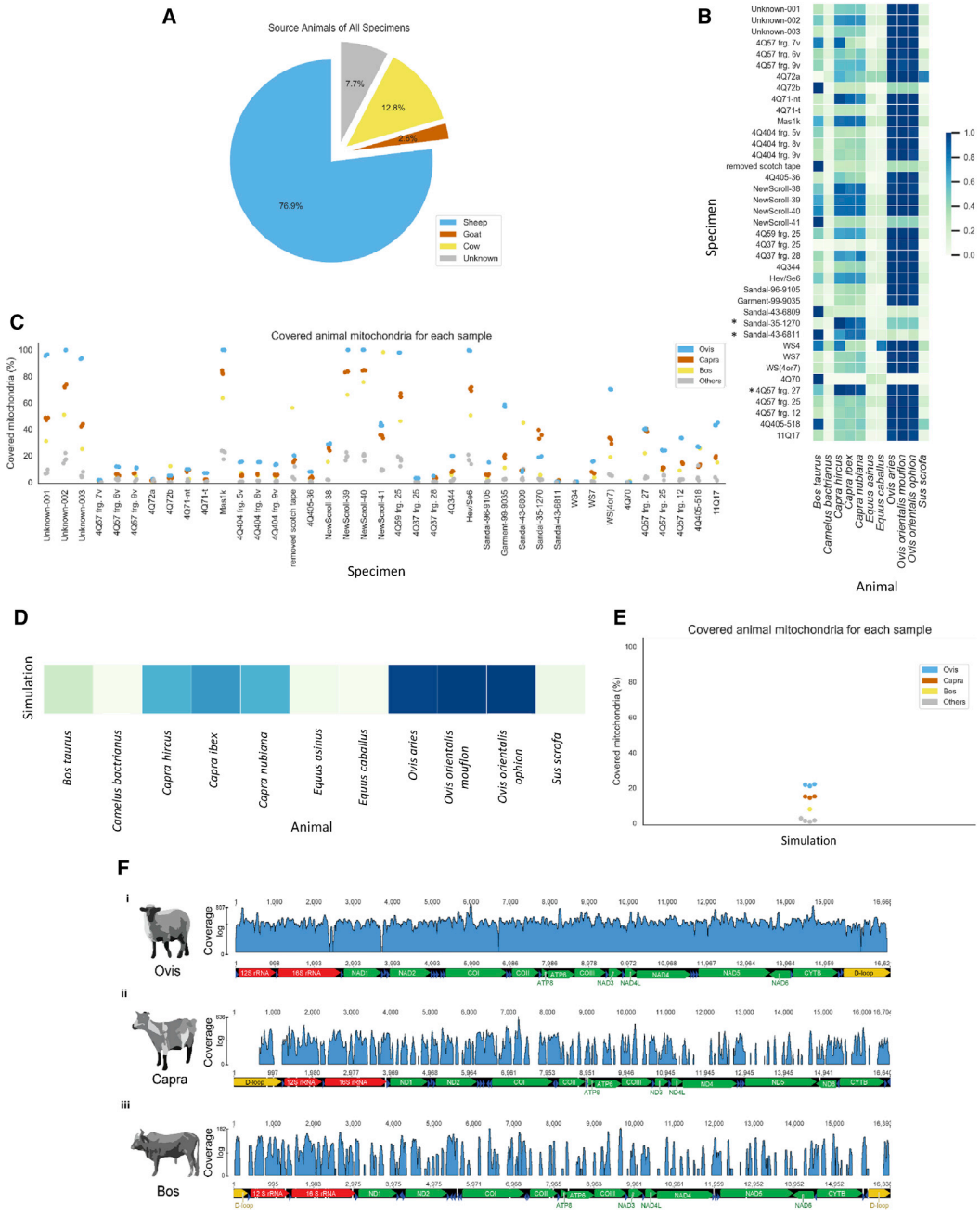


Figure 4. Identification of the Animals from which the Specimens Were Made

(A) Pie chart denoting the distribution of source animals in the different specimens examined.

(B) Heatmap depicting the percentage of the mt reference genome of 11 animals that is covered by the mtDNA in each specimen. Percentage is normalized to the species with the highest coverage. Asterisks indicate specimens whose source animal could not be identified with high confidence.

(C) Strip plot depicting the percentage of the mt reference genome of 11 animals that is covered by the mtDNA in each specimen. The three main animal groups (*Ovis*, *Capra*, and *Bos*) are highlighted.

(D) Similar to (B), based on aDNA simulations.

(E) Similar to (C), based on aDNA simulations.

The simulated reads in (D) and (E) show patterns similar to those witnessed in our aDNA data.

(legend continued on next page)

Please cite this article in press as: Anava et al., Illuminating Genetic Mysteries of the Dead Sea Scrolls, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.046

genome. Obtaining enough DNA for this purpose is very challenging because the extremely delicate state of the DSS only allows, at best, sampling of tiny scroll pieces or crumbs (Figure 2A). To retain a greater fraction of the nuclear DNA sequences and to allow matches to genomes of ancient sheep breeds (while potentially tolerating some level of non-endogenous DNA), we designed high-yield pipelines (STAR Methods; Figures S4A and S5A; Table S3). The processed data behaved similarly to simulated aDNA data (Figures 4D, 4E, and S2D; STAR Methods).

We determined sets of informative positions in the nuclear sheep genome that display variation between the sequences of the different fragments, captured by variant calling procedures (while taking into account the effects of cytosine deamination; see STAR Methods for details). We then performed unsupervised clustering of the fragments using three independent and complementary methods: (1) construction of phylogenetic trees using the maximum likelihood criterion under the GTR-GAMMA model, (2) measurement of similarity based on pairwise comparisons, and (3) measurement of similarity based on pairwise comparisons, which takes into account linkage disequilibrium between variable positions (STAR Methods).

The set of fragments we tested using our new heuristic approach contained positive and negative controls (associated and disparate pieces). Independent of the specific clustering method and parameters used, we uncovered certain consistent and robust patterns of relatedness. In contrast, we did not observe similar relations when using reads suspected to be from human or bacterial contamination and ruled out additional potential biases associated with the variation in coverage and specimen quality (see Figures 5, 6, 7, and S5–S7; Table S6; and the details in "The genetic clustering based on sheep DNA is robust to external factors" section in STAR Methods):

(1) Parchments prepared for the IAA for studying the physical, chemical, and biological processes such materials undergo (Figure 2C; referred to as "IAA parchments") were clustered tightly together (Figure 5), as could be expected from the fact that, as a group, they are distinct from the ancient samples. Additionally, a distantly related breed represented by a modern Tibetan sheep was also always distinguishable from the ancient sheep samples (Figures 6A, 6B, and S6C [i and ii]).

(2) We found that genetic similarities were detected despite especially unfavorable conditions of preservation. For example, 4Q71 (4QJer^{b}) was damaged by adhesive tape used in the 1950s to attach the fragments to each other. Two different parts of the fragment, one affected by the tape and the other not, clustered most closely with one another (bootstrap value 100/100; Figure 5) despite the damaging interventions. Two highly degraded samples from 4Q405, a copy the *Songs*, clustered together (with a moderate bootstrap value of 51/100; Figure 5), although one of the samples (4Q405-36) consisted only of scroll "dust" (removed from the adhesive tape; see example in Figure 2A).

(3) To test the resolution of the analysis, we sampled two ancient waterskins (WS4 and WS7), each made from the skin of a whole sheep, and an isolated fragment that was known to be detached (in modern times) from one of these waterskins, but its specific origin was unknown (Figure 2C). We found that the fragment clusters with WS7 (bootstrap value 100/100; Figure 5). Thus, in specific cases, it is possible to identify individual animals even when the objects were discovered within the same location and dated to the same period.

(4) Similarly, frgs. 8 and 9 of 4Q404, which is another copy of the *Songs*, are too small for secure identification of their place in the manuscript (Figure 1B). According to the genetic analysis, frgs. 8 and 9 cluster closely together (bootstrap value 100/100; Figure 5). This is in line with their resemblance in scribal hands, script, and content (Newsom, 1997).

In addition, genetic sorting based on variations in the nuclear genomes contributes to characterizing ambiguous fragments. Because many scroll fragments were purchased by the archaeological authorities only secondarily, their original location can be dubious. The data support the suspicions, previously suggested on independent grounds, that two scroll fragments, officially assigned to Qumran cave 4, originate elsewhere (Figure 5).

4Q344

This scroll, a document acknowledging a financial debt of one person to another, did not cluster tightly with most other Qumran fragments. Although officially registered as stemming from Qumran Cave 4, scholars have long suspected that this fragment was found elsewhere (Doudna, 1998; Yardeni, 1997). It belongs to a group of scrolls (4Q342–348, 4Q351–354, and 4Q356–361) that are legal documents, whereas virtually all other Qumran texts represent religious literature. Their script sets them apart from the Qumran scrolls, and ^{14}C measurements of some of them (including 4Q344) date them to the late first through early second century CE (Jull et al., 1995; Yardeni, 1997), unlike the earlier Qumran scrolls. Moreover, a few of these fragments (but not 4Q344) can be directly or indirectly joined with documents from the Wadi Seiyal collection (Doudna, 1998; Yardeni, 1997). The genetic evidence supports the philological hypothesis that 4Q344 does not originate from the Qumran caves.

4Q59 (4QIsa^{e}), frg. 25

This fragment did not distinctly cluster with other Qumran scrolls. Although most fragments grouped under the siglum "4Q59" come from the beginning of the book of Isaiah (chapters 2–14), frg. 25 is oddly the only fragment that comes from the end of the book

(F) Mapping of the mt reads of sample Unknown-002 shows clear preference to *O. aries*. (i) Mapping to the mt sheep sequence (*Ovis aries* GenBank: HE577850; mean coverage, 76.2; SD, 64.0). (ii) Mapping to the mt goat sequence (*Capra hircus* GenBank: KP662715; mean coverage, 24.4; SD, 41.3). (iii) Mapping to the mt cow sequence (*Bos taurus* GenBank: V00654; mean coverage, 14.2; SD, 22.5). The coverage is represented in the blue bar graph using a log scale. The reference sequence used is indicated below, with rRNA, tRNA, and protein coding genes indicated in red, blue, and green, respectively. In each case, BLAST searches indicated that the consensus sequence was a sheep sequence.
See also Table S4.

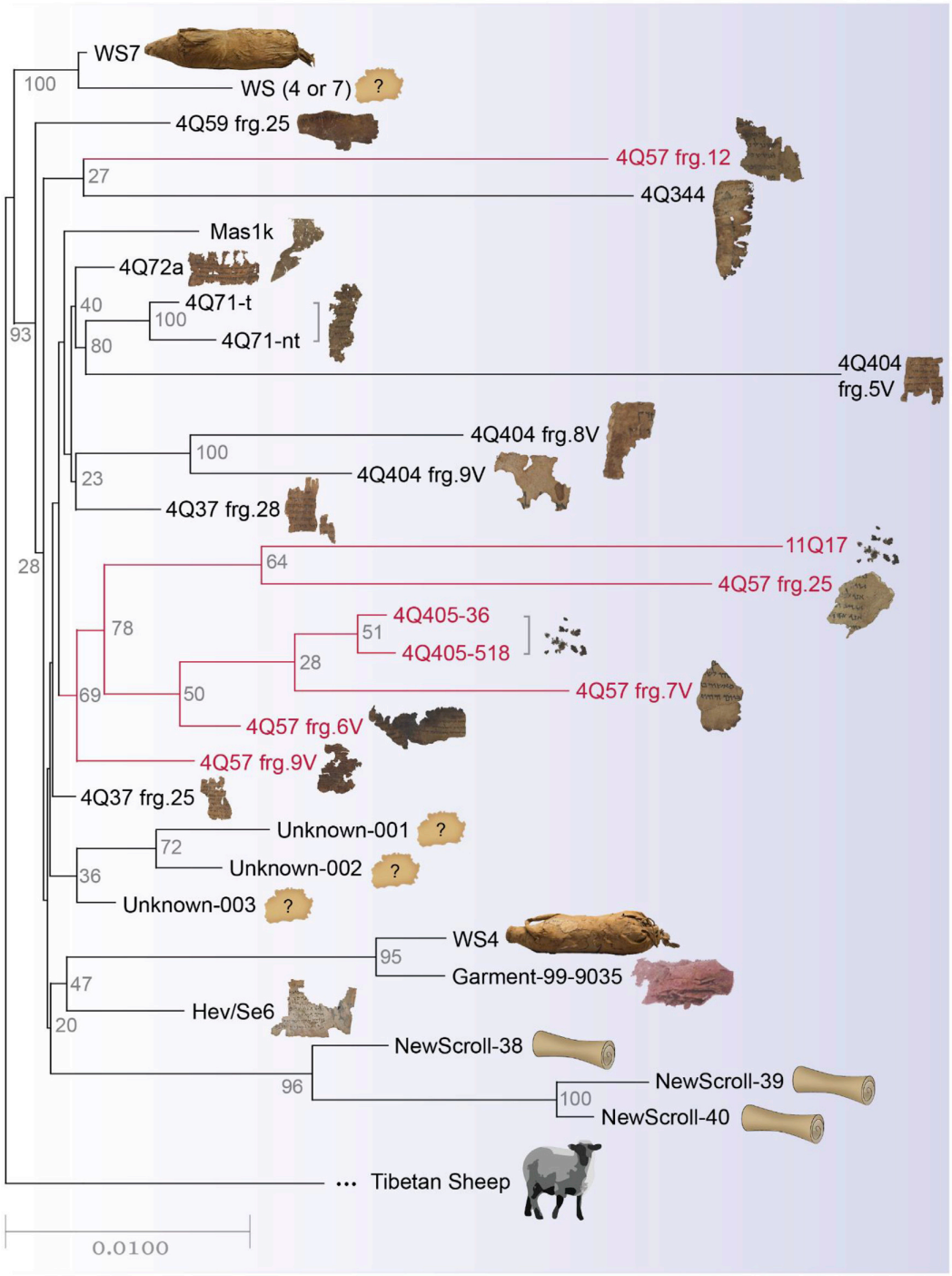


Figure 5. Clustering of the Sheep Specimens Based on Variation in Their Genomic Sequences

Shown is clustering of the specimens based on their nucleotide identity in 4,189 informative positions along the genome; each is covered by at least 3 reads in at least 50% of the examined specimens and outgroup. The nucleotide identity of at least one specimen differs from that existing in the *Ovis aries* reference genome. The tree was generated using the maximum likelihood criterion under the GTR-GAMMA model. Maximum likelihood bootstrap support values based on 500 bootstrap computations are indicated near the corresponding node (bootstrap values lower than 10 are not shown). The QSP cluster is marked in red. Where

(legend continued on next page)

Please cite this article in press as: Anava et al., Illuminating Genetic Mysteries of the Dead Sea Scrolls, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.046

(chapter 59). The editors originally placed it in this scroll only because "it fits nowhere better among the 4QIsaiah fragments" (Skehan and Ulrich, 1997). The genetic evidence suggests that this fragment belongs to a manuscript originating elsewhere.

Genetic Analyses Shedding Light on Scribally Distinct Scrolls

The different animal source of the Jeremiah scrolls (with 4Q71 and 4Q72a deriving from sheep and 4Q70 and 4Q72b from cow; Figure S3A; Table S5) was shown above to stand in relation to their textual diversity (as 4Q71 and 4Q72a match the short text of the Septuagint, 4Q72b matches the long text of the Masoretic tradition, and 4Q70 reflects an originally independent textual tradition brought closer to the long text). This leads to the conclusion that they represent not only a secluded sect but, rather, the broader cultural milieu of Judea in the Second Temple period. Our analyses of variations in the nuclear genome suggest that a similar pattern of textual plurality applies generally to other scrolls as well because some of them can be identified as brought from elsewhere.

Philologists have identified a specific scribal tradition, known as the Qumran scribal practice (QSP), that typifies only a subset of the Qumran scrolls (for the typology and classification of all manuscripts for which there is sufficient textual evidence, see Tov, 2004). QSP is positively defined on the basis of several indicators culled from the realms of orthography (i.e., telltale spellings of common words or linguistic forms), linguistic morphology (the preference for certain alloforms of pronouns, verbs, and adverbs), script (e.g., resorting to the Paleo-Hebrew script for distinguishing God's name from the rest of the text, usually written in the Jewish or Square script), and a range of other scribal practices (e.g., use of scribal marks, particular procedures for deletion and correction, etc.). It should be borne in mind, though, that QSP seems to be a function of a particular social group rather than of a geographical location. In other words, if scrolls were copied locally, then they may be expected to be classified as QSP; at the same time, other QSP manuscripts could theoretically have been penned by sectarian scribes in other places (notably, some QSP manuscripts antedate the Qumran settlement). On the other hand, non-QSP manuscripts are likely to be brought from the outside, being the product of other social and ideological circles. Although the distinction between QSP and non-QSP is still debated (e.g., Tigchelaar, 2010), a QSP profile of a text potentially indicates that a particular manuscript was produced by a sectarian scribe (perhaps but not necessarily at the site of Qumran itself). Significantly, the genetic analyses show that seven of eight samples drawn from QSP manuscripts cluster together (moderate bootstrap value 69/100; Figures 5 [QSP scrolls marked in red], 6C, and S6C [iii] [QSP scrolls marked in gray]; STAR Methods). Tellingly, this cluster includes only manuscripts that reflect QSP.

In contrast, other clusters of manuscripts found in the Qumran caves also consist of manuscripts that are clearly non-QSP, such as the cluster that contains 4Q71, whose paleographic dating is the early second century BCE (Tov, 1989, 1997), probably antedating the deposition of the scrolls, suggesting that these fragments were likely brought to the Qumran caves from the outside. The fragments grouped under this branch are not expected to be tightly related to one another because they were presumably brought from a number of other sites and, hence, reflect the literature available for readers throughout the land.

Thus, the genetic distinction between different branches of the phylogenetic tree supports the philological characterization of diverse scribal traditions (i.e., QSP and non-QSP) and offers an independent tool for classifying ambiguous cases.

In conclusion, our study exemplifies that genome sequencing is an independent scientific method for studying ancient artifacts that should benefit researchers from different fields because it can be used for testing existing hypotheses and for generating new ones.

DISCUSSION

The discovery of the DSS is commonly regarded as one of the most important archeological discoveries of the 20th century, but the full scope of their historical importance depends, to a significant extent, on the question of how representative they are for the broader society at the time. How reflective are they of the cultural history of the religious literature subsequently inherited by Judaism and Christianity? The complexity of the ancient evidence as furnished by the DSS calls for an integrative approach that takes advantage of diverse research strategies.

It should be stated at the outset that genetic information can only reveal a part of the picture. A given scroll could consist of several sheets not necessarily made from the same individual animal; for example, 1QIsa^{a}, which is a fully preserved copy of Isaiah, is 7.34 m long and made of 17 different sheets that cannot all be made of the skin of a single animal. In addition, the production of leather includes use of organic materials that might have influenced the genomic analysis (Campana et al., 2010). Likewise, we also found goat DNA in a fragment that had been classified as part of 4Q57 (4Q57 frg. 27) and were thus unable to confidently identify variants in its sheep DNA. The goat DNA (in addition to the sheep DNA) found in this sample could be authentic because it is possible that materials from various organic sources were used in the process of scroll preparation (e.g., tanning) (Campana et al., 2010; Poole and Reed, 1962).

The challenge is exemplified by the results obtained from two fragments of 4Q37. This scroll contains a scriptural text and was published as a copy of the book of Deuteronomy (4QDeut^{j}), but some of its fragments contain passages from the book of Exodus. The textual evidence thus allows two main possibilities: either both groups of fragments represent two different scrolls—copies of Exodus and Deuteronomy, respectively—written by the same scribe, or both groups stem from a single manuscript whose text represents a conflation of passages from the two biblical books. The official editor favored the second option, hypothesizing that 4Q37 was a collection of excerpted scriptural

available, images of the sampled specimens are shown; the rest of the images in the figure are schematic drawings and do not necessarily represent any specific specimen. Scroll images are courtesy of the Leon Levy Dead Sea Scrolls Digital Library, Israel Antiquities Authority; photos: Shai Halevi. Waterskin images are courtesy of the Israel Museum, Jerusalem; photos: Elie Posner. See also Figures S5–S7 and Tables S5 and S6.

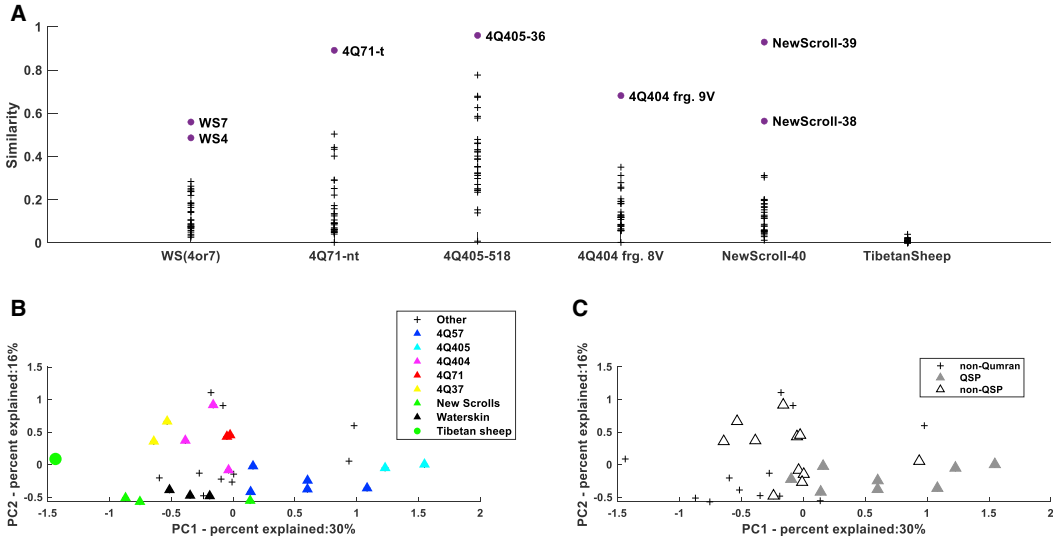


Figure 6. Similarity Analysis of the Various Specimens Using Pairwise Comparison of Shared Genomic Positions

(A) The extent of similarity between selected specimens and all other examined specimens. The outgroup is different from all other specimens. All compared positions are covered by at least 3 reads.

(B) Principal-component analysis (PCA) projection of the sheep specimens based on a pairwise similarity matrix. Different specimen groups are marked in colors (see legend).

(C) PCA projection of the sheep specimens based on a pairwise similarity matrix. QSP and non-QSP fragments are marked by gray and white triangles, respectively (see legend on the right). The QSP fragments are marked in red in Figure 5.

See also Figures S5–S7.

passages (Duncan, 1991, 1995, 1997). Indeed, the contents of the work (Deuteronomy 5–6, 8, 11, and 32; Exodus 12–13) and its presumed order are matched by liturgical compilations of scriptural passages (some of the Qumran phylacteries) (Tov, 2019). The relatively small size of the scroll is also suggestive of its special—possibly liturgical—function (cf. 4QDeut^{n} and 4QDeut^{q}).

We sampled specimens from each group of fragments; frg. 25 comes from column VIII (with frg. 26) and contains the text of Deuteronomy 11, whereas frg. 28 comes from column X (with frgs. 29–33) and contains the text of Exodus 12–13. We found that frgs. 25 and 28 do not cluster together (Figure 5). These genetic results suggest that the fragments originate from different individual animals of the same species. Thus, if they belong to the same scroll, they probably stem from two different sheets. This conclusion supplies information for future attempts for material reconstruction of 4Q37 and literary evaluation of its contents.

Despite the theoretical, methodological, and practical difficulties any study of ancient artifacts is bound to tackle, our investigations have obtained many significant results that could not have been reached by philologists or scientists working in isolation of each other. Some of our novel insights contribute to a better understanding of long-debated problems, whereas others reveal new areas of investigation.

For example, the non-scriptural scrolls can be classified to subgroups that correspond to the scribal distinction between QSP (potentially locally copied) and non-QSP manuscripts (probably copied elsewhere), confirming that the collection exemplifies the cultural plurality that typified Judea in the Hellenistic-Roman period and served as the matrix for the eventual development of Judaism (and Christianity). In addition, the ancient diversity revealed by genetics can shed new light on the provenance of the various copies of the *Songs*. The discovery of Mas1k complicated the debate about the socio-religious background of the *Songs*: should it be taken as a piece of evidence that sectarians fled to Masada after the fall of Qumran (68 CE) or Jerusalem (70 CE) to the Romans and joined the Zealots who captured this fortress? Or does it indicate that the work was popular among circles that transcend the Qumran caves? The fact that only Mas1k belongs to haplogroup A (see Results and STAR Methods) whereas the Qumran copies of the work belong to haplogroup B suggests that the work was known beyond Qumran caves.

The scriptural scrolls, too, as represented by the Jeremiah fragments, include manuscripts deriving from various places and reflecting highly divergent versions of the book, which presumably testify to different stages in the compositional history of Jeremiah. As acknowledged by biblical scholarship, the textual evidence of the Second Temple period demonstrates that different versions of biblical books circulated in parallel; for contemporaries, the holiness of the biblical book did not extend to its precise wording. In contrast, the thousands of surviving manuscripts of the Hebrew Bible produced in the Middle Ages—and later printed editions through the 21st century—witness a virtually identical text, almost down to the letter, exemplifying the loss of textual diversity over the years. The DSS take us back to an earlier stage, typified by textual pluriformity.

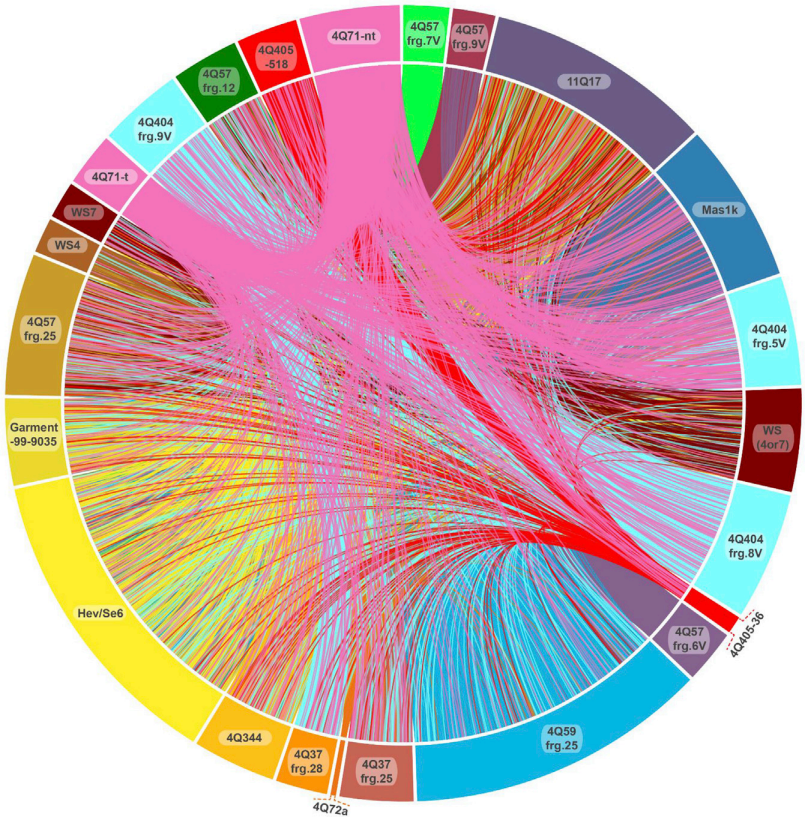


Figure 7. Circos Display of Identical Nucleotides Shared by the Various Specimens

Each line represents 200 positions shared by two specimens, where both have identical nucleotides different from the *Ovis aries* reference genome. The size of each specimen is proportional to the number of its alternative nucleotides compared with the reference genome.

By analogy, the revolution in the study of aDNA in recent years has shown that, as a result of bottleneck events that reduce population size, human genetic diversity decreases in native populations as a function of the distance from Africa (Ramachandran et al., 2005), the cradle of the human species. Similarly, the genetic evidence presented here suggests that the Jeremiah scrolls from the Qumran caves, typified by their textual and genetic diversity, allow a glimpse into the "Africa" of later versions of the book.

Finally, the genetic evidence indicates that two fragments (4Q344 and 4Q59 frg. 25) appear to be different from all other samples taken from the Qumran scrolls, and additional, philological evidence supports the conclusion that they are indeed unrelated to the scrolls found in the Qumran caves. To be sure, these fragments are authentic but were probably unearthed in other caves in the Judean Desert. The use of aDNA methods, such as the methods used and developed in this study, to compare specimens with other ancient specimens, can distinguish between manuscripts that were discovered in the same areas, originate in the same animal source, or were prepared using similar techniques as for the DSS.

This research has approached the study of the DSS from a multidisciplinary perspective, combining genetic analyses of the aDNA extracted from the DSS fragments with philological and historical examination of the ancient literary and religious works copied therein. Such an integrative approach could be applied to analyses of many additional scrolls and diverse manuscripts. We are hopeful that it will encourage others to collaborate and combine their expertise to solve the mysteries of the past.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- METHOD DETAILS
  - Sampling of the specimens
  - Ancient DNA methods
  - DNA Extraction
  - Library Preparation
  - Mitochondrial Capture
  - Quality control, trimming and merging of NGS data
  - Mapping NGS data based on established ancient human pipeline against whole genomes

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cell.2020.04.046.

ACKNOWLEDGMENTS

We would like to thank the IAA for generously providing access to the DSS and for many helpful discussions and consultations. In particular, special thanks are due to Helena Libman, Tanya Treiger, and Tanya Bitler, DSS Senior Conservators, for sampling the scrolls and to Dr. Illit Cohen-Ofri, Head of the DSS Conservation Laboratory, for providing valuable input. We gratefully acknowledge the contributions of Dr. Orit Shamir and Dr. Na'ama Sukenik for allowing us to sample related artefacts. We would like to thank David Mevorach, Senior Curator of the Hellenistic, Roman, and Byzantine Periods at the Israel Museum, for generously providing access to ancient specimens. We would like to thank Hannah Edlund for her technical assistance and Noah Alexander and Jorge Gandara for protocol testing and optimization. We would like to thank Dror Cohen for his assistance in the graphical design of many of the figures shown in the paper, Noa Alon for the design of the graphical abstract, and Viktor Koen for designing the cover suggestion. We would like to thank Prof. Elisha Gootwine and Prof. Eyal Seroussi for sharing their knowledge regarding ancient sheep populations in the area. We would like to thank the Epigenomics Core Facility at Weill Cornell Medicine as well as the Bert L. and N. Kuggie Vallee Foundation, the WorldQuant Foundation, NASA (NNX14AH50G, 15-15Omni2-0063), the Alfred P. Sloan Foundation (G-2015-13964), the Knut and Alice Wallenberg Foundation, and the Göran Gustavsson Foundation. Sequencing was performed at SciLifeLab's National Genomics Infrastructure (NGI), Uppsala, Sweden. Computations were performed at the Swedish National Infrastructure for Computing (SNIC-UPPMAX) and with help and support from the HPC team at the TAU Computational Center. E.M.S. is supported by Formas (2011-1138). Research regarding the *Songs* is supported by the Israel Science Foundation (723/13). O.R. is supported by the Adelis Foundation (01430001000). This work was funded by the generosity of Eric and Wendy Schmidt by recommendation of the Schmidt Futures program (590772).

AUTHOR CONTRIBUTIONS

Conceptualization, O.R., M.J., N.M., S.A., M.N., and H.G.; Methodology, M.N., H.G., S.A., D.H., O.S., A.M., D.D., E.A., J.F., C.E.M., O.R., M.J., and N.M.; Software, M.N., H.G., O.S., A.M., and D.D.; Formal Analysis, M.N., H.G., D.H., O.S., and D.D.; Investigation, S.A. and E.M.S.; Resources, S.A., E.M.S., B.R., and P.S.; Analysis of Ancient Animal DNA, M.N., H.G., S.A., O.S., and A.M.; Microorganism Metagenomics Analysis, D.D., E.A., and J.F; Writing – Original Draft, S.A., M.N., H.G., M.J., N.M., and O.R.; Writing – Review & Editing, P.S. and B.R; Visualization, M.N., H.G., D.H., O.S., and D.D.; Data Curation, M.N.; Supervision, O.R., M.J., and N.M.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 11, 2020
Revised: April 1, 2020
Accepted: April 23, 2020
Published: June 2, 2020

REFERENCES

Achilli, A., Olivieri, A., Pellecchia, M., Uboldi, C., Colli, L., Al-Zahery, N., Accetturo, M., Pala, M., Hooshiar Kashani, B., Perego, U.A., et al. (2008). Mitochondrial genomes of extinct aurochs survive in domestic cattle. Curr. Biol. *18*, R157–R158.

Abusch, R. (2003). Sevenfold hymns in the Songs of the Sabbath Sacrifice and the Hekhalot literature: Formalism, hierarchy and the limits of human participation. In The Dead Sea Scrolls as bBackground to pPostbiblical Judaism and eEarly Christianity, J.R. Davila, ed. (Leiden: Brill), pp. 220–247.

Achilli, A., Bonfiglio, S., Olivieri, A., Malusà, A., Pala, M., Hooshiar Kashani, B., Perego, U.A., Ajmone-Marsan, P., Liotta, L., Semino, O., et al. (2009). The multifaceted origin of taurine cattle reflected by the mitochondrial genome. PLoS ONE *4*, e5753.

Ahmed, S., Grobler, P., Madisha, T., and Kotze, A. (2017). Mitochondrial D-loop sequences reveal a mixture of endemism and immigration in Egyptian goat populations. Mitochondrial DNA A. DNA Mapp. Seq. Anal. *28*, 711–716.

Alexander, P.S. (2010). Qumran and the genealogy of western mysticism. In New prspectives on old texts, E.G. Chazon, B. Halpern-Amaru, and R.A. Clements, eds. (Leiden: Brill), pp. 215–235.

Allentoft, M.E., Sikora, M., Sjögren, K.-G., Rasmussen, S., Rasmussen, M., Stenderup, J., Damgaard, P.B., Schroeder, H., Ahlström, T., Vinner, L., et al. (2015). Population genomics of Bronze Age Eurasia. Nature *522*, 167–172.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

Andrews, S. (2010). FastQC: A Quality Control tool for High Throughput Sequence Data (Babraham Bioinformatics).

Barnett, D.W., Garrison, E.K., Quinlan, A.R., Strömberg, M.P., and Marth, G.T. (2011). BamTools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics *27*, 1691–1692.

Bouchnick, R. (2016). Meat consumption patterns as an ethnic marker in the late Second Temple period: Comparing the Jerusalem city dump and Qumran assemblages. In Bones and identity, N. Marom, R. Yeshurun, L. Weissbrod, and G. Bar-Oz, eds. (Oxford: Oxbow Books), pp. 303–321.

Briggs, A.W., Stenzel, U., Johnson, P.L.F., Green, R.E., Kelso, J., Prüfer, K., Meyer, M., Krause, J., Ronan, M.T., Lachmann, M., and Pääbo, S. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. Proc. Natl. Acad. Sci. USA *104*, 14616–14621.

Břinda, K., Boeva, V., and Kucherov, G. (2016). Dynamic read mapping and online consensus calling for better variant detection. ArXiv, arXiv:1605.09070. https://arxiv.org/abs/1605.09070.

Brooke, G.J. (2005). The Dead Sea Scrolls and the New Testament (Fortress Press).

Campana, M.G., Bower, M.A., Bailey, M.J., Stock, F., O'Connell, T.C., Edwards, C.J., Checkley-Scott, C., Knight, B., Spencer, M., and Howe, C.J. (2010). A flock of sheep, goats and cattle: ancient DNA analysis reveals complexities of historical parchment manufacture. J. Archaeol. Sci. *37*, 1317–1325.

Chang, E.S., Neuhof, M., Rubinstein, N.D., Diamant, A., Philippe, H., Huchon, D., and Cartwright, P. (2015). Genomic insights into the evolutionary origin of Myxozoa within Cnidaria. Proc. Natl. Acad. Sci. USA *112*, 14912–14917.

Chevreux, B., Wetter, T., and Suhai, S. (1999). Genome sequence assembly using trace signals and additional sequence information. German Conference on Bioinformatics (Heidelberg), pp. 45–56.

Cock, P.J.A., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., and de Hoon, M.J. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics *25*, 1422–1423.

Colli, L., Lancioni, H., Cardinali, I., Olivieri, A., Capodiferro, M.R., Pellecchia, M., Rzepus, M., Zamani, W., Naderi, S., Gandini, F., et al. (2015). Whole mitochondrial genomes unveil the impact of domestication on goat matrilineal variability. BMC Genomics *16*, 1115.

Collins, J.J. (2010). Beyond the Qumran community: The sectarian movement of the Dead Sea Scrolls (Grand Rapids, Michigan: Wm. B. Eerdmans Publishing).

Collins, J.J. (2012). The Dead Sea Scrolls: A biography (Princeton, NJ: Princeton University Press).

Dabney, J., Meyer, M., and Pääbo, S. (2013a). Ancient DNA damage. Cold Spring Harb. Perspect. Biol. *5*, a012567.

Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., and Meyer, M. (2013b). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc. Natl. Acad. Sci. USA *110*, 15758–15763.

Davies, P.R., Brooke, G.J., and Callaway, P.R. (2002). The complete world of the Dead Sea Scrolls (London: Thames & Hudson).

Davis, K. (2018). Margins as media: The long insertion in 4QJer[a] (4Q70). In Bible as notepad, L.I. Lied and M. Maniaci, eds. (Berlin, Boston: De Gruyter), pp. 39–53.

Dimant, D. (2014). History, ideology and Bible interpretation in the Dead Sea Scrolls: Collected studies (Mohr Siebeck).

Doudna, G.L. (1998). Dating the scrolls on the basis of radiocarbon analysis. In Dead Sea Scrolls after fifty years, P. Flint and J. VanderKam, eds. (Leiden: Brill), pp. 430–471.

Duhm, B. (1901). Das Buch Jeremia (Tübingen: Mohr Siebeck).

Duncan, J. (1991). Considerations of 4QDt[j] in light of the All Soul's Deuteronomy and Cave 4 phylactery texts. In The Madrid Qumran congress, J.T. Barrera and L. Vegas Montaner, eds. (Leiden: Brill), pp. 199–215.

Duncan, J.A. (1995). 4QDeut[j]. Qumran Cave 4, IX: Deuteronomy, Joshua, Judges, Kings (DJD 14) (Oxford: Clarendon Press), pp. 75–91.

Duncan, J.A. (1997). Excerpted texts of Deuteronomy at Qumran. Rev. Qumran *18*, 43–62.

E, G.X., Zhao, Y.-J., and Huang, Y.-F. (2018). Sheep mitochondrial heteroplasmy arises from tandem motifs and unspecific PCR amplification. Mitochondrial DNA A. DNA Mapp. Seq. Anal. *29*, 91–95.

Gilbert, M.T.P., Wilson, A.S., Bunce, M., Hansen, A.J., Willerslev, E., Shapiro, B., Higham, T.F.G., Richards, M.P., O'Connell, T.C., Tobin, D.J., et al. (2004). Ancient mitochondrial DNA from hair. Curr. Biol. *14*, R463–R464.

Gootwine, E. (2015). The physical appearance of sheep in antiquity, based upon archeological findings in Israel and neighboring lands. Qadmoniot *150*, 125–132.

Gorkhali, N.A., Jiang, L., Shrestha, B.S., He, X.-H., Junzhao, Q., Han, J.-L., and Ma, Y.-H. (2016). High occurrence of mitochondrial heteroplasmy in nepalese indigenous sheep (Ovis aries) compared to chinese sheep. Mitochondrial DNA A. DNA Mapp. Seq. Anal. *27*, 2645–2647.

Günther, T., Valdiosera, C., Malmström, H., Ureña, I., Rodriguez-Varela, R., Sverrisdóttir, Ó.O., Daskalaki, E.A., Skoglund, P., Naidoo, T., Svensson, E.M., et al. (2015). Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. Proc. Natl. Acad. Sci. USA *112*, 11917–11922.

Hahn, C., Bachmann, L., and Chevreux, B. (2013). Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads–a baiting and iterative mapping approach. Nucleic Acids Res. *41*, e129.

Haldane, J.B.S. (1919). The combination of linkage values and the calculation of distances between the loci of linked factors. J. Genet. *8*, 299–309.

Hansen, A., Willerslev, E., Wiuf, C., Mourier, T., and Arctander, P. (2001). Statistical evidence for miscoding lesions in ancient DNA templates. Mol. Biol. Evol. *18*, 262–265.

Haran, M. (1985). Bible scrolls in eastern and western Jewish communities from Qumran to the high middle ages. Hebr. Union Coll. Annu. *56*, 21–62.

Hassanin, A., Ropiquet, A., Couloux, A., and Cruaud, C. (2009). Evolution of the mitochondrial genome in mammals living at high altitude: new insights from a study of the tribe Caprini (Bovidae, Antilopinae). J. Mol. Evol. *68*, 293–310.

Hassanin, A., Bonillo, C., Nguyen, B., and Cruaud, C. (2010). Comparisons between mitochondrial genomes of domestic goat (Capra hircus) reveal the presence of numts and multiple sequencing errors. Mitochondrial DNA *21* (3–4), 68–76. https://doi.org/10.3109/19401736.2010.490583.

Hassanin, A., Delsuc, F., Ropiquet, A., Hammer, C., Jansen van Vuuren, B., Matthee, C., Ruiz-Garcia, M., Catzeflis, F., Areskoug, V., Nguyen, T.T., and Couloux, A. (2012). Pattern and timing of diversification of Cetartiodactyla (Mammalia, Laurasiatheria), as revealed by a comprehensive analysis of mitochondrial genomes. C. R. Biol. *335*, 32–50.

Hofreiter, M., Jaenicke, V., Serre, D., von Haeseler, A., and Pääbo, S. (2001). DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. Nucleic Acids Res. *29*, 4793–4799.

Jull, A.J.T., Donahue, D.J., Broshi, M., and Tov, E. (1995). Radiocarbon Dating of Scrolls and Linen Fragments from the Judean Desert. Radiocarbon *37*, 11–19. https://doi.org/10.1017/S0033822200014740.

Kahila Bar-Gal, G. (2000). Genetic changes in Capra species of Southern Levant over the past 10,000 years as studied by DNA analysis of ancient and modern populations. PhD thesis (Hebrew University of Jerusalem).

Kang, D.D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., and Wang, Z. (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ *7*, e7359.

Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780.

Kircher, M. (2012). Analysis of High-Throughput Ancient DNA Sequencing Data. In Humana Press), D.N.A. Ancient, B. Shapiro, and M. Hofreiter, eds., pp. 197–228.

Kuhn, J.M.M., Jakobsson, M., and Günther, T. (2018). Estimating genetic kin relationships in prehistoric populations. PLoS ONE. Published online April 23, 2018. https://doi.org/10.1371/journal.pone.0195491.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. *33*, 1870–1874.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Lartillot, N., Lepage, T., and Blanquart, S. (2009). PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics *25*, 2286–2288.

Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., Sudmant, P.H., Schraiber, J.G., Castellano, S., Lipson, M., et al. (2014). Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature *513*, 409–413.

Li, H. (2016). seqtk (Elixir bio.tools).

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Li, D., Liu, C.-M., Luo, R., Sadakane, K., and Lam, T.-W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics *31*, 1674–1676.

Lv, F.-H., Peng, W.-F., Yang, J., Zhao, Y.-X., Li, W.-R., Liu, M.-J., Ma, Y.-H., Zhao, Q.-J., Yang, G.-L., Wang, F., et al. (2015). Mitogenomic Meta-Analysis Identifies Two Phases of Migration in the History of Eastern Eurasian Sheep. Mol. Biol. Evol. *32*, 2515–2533.

Malmström, H., Svensson, E.M., Gilbert, M.T.P., Willerslev, E., Götherström, A., and Holmlund, G. (2007). More on contamination: the use of asymmetric molecular behavior to identify authentic ancient human DNA. Mol. Biol. Evol. *24*, 998–1004.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. *17*, 1.

Meadows, J.R.S., Cemal, I., Karaca, O., Gootwine, E., and Kijas, J.W. (2007). Five ovine mitochondrial lineages identified from sheep breeds of the near East. Genetics *175*, 1371–1379.

Please cite this article in press as: Anava et al., Illuminating Genetic Mysteries of the Dead Sea Scrolls, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.046

Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. *2010*, t5448.

Meyer, M., Kircher, M., Gansauge, M.-T., Li, H., Racimo, F., Mallick, S., Schraiber, J.G., Jay, F., Prüfer, K., de Filippo, C., et al. (2012). A High-Coverage Genome Sequence from an Archaic Denisovan Individual. Science *338*, 222–226.

Mizrahi, N. (2013). Aspects of poetic stylization in Second Temple Hebrew: A comparison of the Songs of the Sabbath Sacrifice with ancient Piyyut. In Hebrew in the Second Temple period, M. Bar-Asher, S.E. Fassberg, and R.A. Clements, eds. (Leiden: Brill), pp. 147–163.

Newsom, C.A. (1990). 'Sectually explicit' literature from Qumran. In The Hebrew Bible and its interpreters, W.H. Propp, B. Halpern, and D.N. Freedman, eds. (Tübingen: Mohr Siebeck), pp. 167–187.

Newsom, C.A. (1997). Shirot Olat Hashabbat. Qumran Cave 4, VI: Poetical and liturgical texts, Part 1 (DJD 11) (Oxford: Clarendon Press), pp. 173–402.

Olivieri, A., Gandini, F., Achilli, A., Fichera, A., Rizzi, E., Bonfiglio, S., Battaglia, V., Brandini, S., De Gaetano, A., El-Beltagi, A., et al. (2015). Mitogenomes from Egyptian Cattle Breeds: New Clues on the Origin of Haplogroup Q and the Early Spread of Bos taurus from the Near East. PLoS ONE *10*, e0141170.

Olm, M.R., Brown, C.T., Brooks, B., and Banfield, J.F. (2017). dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. ISME J. *11*, 2864–2868.

Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. *25*, 1043–1055.

Parma, P., Feligini, M., Greeppi, G., and Enne, G. (2003). The complete nucleotide sequence of goat (Capra hircus) mitochondrial genome. Goat mitochondrial genome. DNA Seq. *14*, 199–203.

Poole, J.B., and Reed, R. (1962). The preparation of leather and parchment by the Dead Sea Scrolls community. Technol. Cult. *3*, 1–26.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Ramachandran, S., Deshpande, O., Roseman, C.C., Rosenberg, N.A., Feldman, M.W., and Cavalli-Sforza, L.L. (2005). Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. Proc. Natl. Acad. Sci. USA *102*, 15942–15947.

Ratzon, E., and Ben-Dov, J. (2017). A newly reconstructed calendrical scroll from Qumran in cryptic script. J. Biblic. Lit. *136*, 905.

Reed, R., and Poole, J. (1964). A study of some Dead Sea Scroll and leather fragments from cave 4 at Qumran. Part I - Physical examination. Proc. Leeds Philos. Lit. Soc. Sci. Sect. *9*, 1–13.

Renaud, G., Hanghøj, K., Willerslev, E., and Orlando, L. (2017). gargammel: a sequence simulator for ancient DNA. Bioinformatics *33*, 577–579.

Ryder, M.L. (1958). Follicle arrangement in skin from wild sheep, primitive domestic sheep and in parchment. Nature *182*, 781–783.

Sawyer, S., Krause, J., Guschanski, K., Savolainen, V., and Pääbo, S. (2012). Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. PLoS ONE *7*, e34131.

Schäfer, P. (2006). Communion with the angels: Qumran and the origins of Jewish mysticism. In Wege mystischer Gotteserfahrung; Judentum, Christentum und Islam, C. Judaism, P.S. Islam, and E. Müller-Luckner, eds. (München: Oldenbourg), pp. 37–66.

Simpson, J.T., and Durbin, R. (2012). Efficient de novo assembly of large genomes using compressed data structures. Genome Res. *22*, 549–556.

Skehan, P.W., and Ulrich, E. (1997). Isaiah. In Qumran Cave 4, X: The Prophets (DJD 15), E. Ulrich, ed. (Oxford: Clarendon Press), pp. 7–144.

Skoglund, P., Northoff, B.H., Shunkov, M.V., Derevianko, A.P., Pääbo, S., Krause, J., and Jakobsson, M. (2014a). Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. Proc. Natl. Acad. Sci. USA *111*, 2229–2234.

Skoglund, P., Malmström, H., Omrak, A., Raghavan, M., Valdiosera, C., Günther, T., Hall, P., Tambets, K., Parik, J., Sjögren, K.-G., et al. (2014b). Genomic Diversity and Admixture Differs for Stone-Age Scandinavian Foragers and Farmers. Science *344*, 747–750.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312–1313.

Supreme, J. (2013). Determination of the genetic cause of an internationally unique, naturally occurring muscular dystrophy in Western Australian Merino sheep (Murdoch University).

Teasdale, M.D., van Doorn, N.L., Fiddyment, S., Webb, C.C., O'Connor, T., Hofreiter, M., Collins, M.J., and Bradley, D.G. (2015). Paging through history: parchment as a reservoir of ancient DNA for next generation sequencing. Philos. Trans. R. Soc. B Biol. Sci. *370*, 20130379.

Tigchelaar, E.J.C. (2010). Assessing Emanuel Tov's "Qumran Scribal Practice.". The Dead Sea Scrolls (Brill), pp. 173–208.

Tov, E. (1989). The Jeremiah scrolls from Qumran. Rev. Qumran *14*, 189–206.

Tov, E. (1997). Jeremiah. In Qumran Cave 4, X: The Prophets (DJD 15), E. Tov, ed. (Oxford: Clarendon Press), pp. 145–208.

Tov, E. (2004). Scribal practices and approaches reflected in the texts found in the Judean desert (Leiden: Brill).

Tov, E. (2012). Textual criticism of the Hebrew Bible (Fortress).

Tov, E. (2019). The Qumran Tefillin and their possible master copies. In On wings of prayer: Sources of Jewish Worship; Essays in honor of Professor Stefan C. Reif on the occasion of his seventy-fifth birthday, N. Calduch-Benages, M.W. Duggan, and D. Marx, eds. (Berlin: De Gruyter), pp. 135–149.

Ullmann-Margalit, E. (2006). Out of the cave: A philosophical inquiry into the Dead Sea scrolls research (Cambridge, Mass: Harvard University Press).

VanderKam, J., and Flint, P. (2005). The meaning of the Dead Sea scrolls: their significance for understanding the Bible, Judaism, Jesus, and Christianity (A&C Black).

Wingett, S. (2015). FastQ Screen (Babraham Bioinformatics).

Wood, N.J., and Phua, S.H. (1996). Variation in the control region sequence of the sheep mitochondrial genome. Anim. Genet. *27*, 25–33.

Yang, D.Y., Eng, B., Waye, J.S., Dudar, J.C., and Saunders, S.R. (1998). Technical note: improved DNA extraction from ancient bones using silica-based spin columns. Am. J. Phys. Anthropol. *105*, 539–543.

Yardeni, A. (1990). The palaeography of 4QJer[a]: A comparative study. Textus *15*, 233–268.

Yardeni, A. (1997). Documentary texts alleged to be from Qumran Cave 4. In Aramaic, Hebrew, and Greek documentary texts from Naḥal Ḥever and other sites, with an appendix containing alleged Qumran texts (The Seiyâl Collection, II) (DJD 27), H.M. Cotton and A. Yardeni, eds. (Clarendon Press), pp. 283–317.

Yuditsky, A. (2017). 4Q124: A sapiental Scroll mentioning the sin of Adam and his punishment. A new edition (in Hebrew). Leshonenu *79*, 478–488.

Please cite this article in press as: Anava et al., Illuminating Genetic Mysteries of the Dead Sea Scrolls, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.046

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological Samples		
Dead Sea Scroll fragment samples	DSS Conservation Laboratories of the Israel Antiquities Authority	N/A
IAA parchments	Casali Institute, Hebrew University of Jerusalem	N/A
Ancient leather artifacts	Israel Antiquities Authority	N/A
Chemicals, Peptides, and Recombinant Proteins		
RNase AWAY surface decontamination	ThermoFisher Scientific	7002
AmpliTaq Gold DNA polymerase	ThermoFisher Scientific	N8080259
AMPure XP for PCR purification	Beckman Coulter	A63881
KAPA2G Fast HotStart ReadyMix	Merck (Sigma-Aldrich)	KK5601
Critical Commercial Assays		
High Sensitivity D1000 Kit	Agilent Technologies	5067- 5584, 5067- 5585
myBaits® Mito	Arbor Biosciences	303048
High pure viral nucleic acid large volume kit (for Roche columns)	Roche molecular systems	5114403001
Deposited Data		
Raw and analyzed data	This paper	SRA: SRP151790
Human reference genome build 37, hs37d5	Genome Reference Consortium	GenBank: GCF_000001405.17
human reference genome (GRCh38.p6)	Genome Reference Consortium	GenBank: GCF_000001405.32
Ovis aries reference genome (Oar_v4.0)	International Sheep Genome Consortium	GenBank: GCF_000298735.2
Capra hircus reference genome (ARS1)	USDA ARS	GenBank: GCA_001704415.1
Capra hircus reference genome (CHIR_1.0)	International Goat Genome Consortium	GenBank: GCA_000317765.1
Capra hircus reference genome (CHIR_2.0)	Beijing Genomics Institute	GenBank: GCA_000317765.2
Bos Taurus reference genome (UMD3.1)	Center for Bioinformatics and Computational Biology, University of Maryland	GenBank: GCA_000003055.5
Bos Taurus reference genome (Btau_4.6.1)	Cattle Genome Sequencing International Consortium	GenBank: GCF_000003205.5
Camelus bactrianus reference genome (Ca_bactrianus_MBC_1.0)	Inner Mongolia Agricultural University	GenBank: GCF_000767855.1
Equus asinus reference genome (ASM130575v1)	Inner Mongolian Agricultural University	GenBank: GCF_001305755.1
Equus caballus reference genome (EquCab2)	The Genome Assembly Team	GenBank: GCA_000002305.1
Sus scrofa reference genome (Sscrofa9)	The Swine Genome Sequencing Consortium (SGSC)	GenBank: GCA_000003025.1
Sus scrofa reference genome (Sscrofa10)	The Swine Genome Sequencing Consortium (SGSC)	GenBank: GCF_000003025.5
Gallus gallus reference genome (WASHUC2)	Genome Reference Consortium	GenBank: GCA_000002315.5
Human mitochondrial genome (rCRS)	GenBank	GenBank: J01415.2
Ammotragus lervia mitochondrial genome	RefSeq	GenBank: NC_009510.1
Bos taurus mitochondrial genome	GenBank	GenBank: V00654.1
Bos taurus isolate RHV529 mitochondrial genome	GenBank	GenBank: KJ709685.1

(*Continued on next page*)

Please cite this article in press as: Anava et al., Illuminating Genetic Mysteries of the Dead Sea Scrolls, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.046

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Equus asinus mitochondrial genome	RegSeq	GenBank: NC_001788.1, GenBank: X97337.1
Equus caballus mitochondrial genome	RefSeq	GenBank: NC_001640.1
Equus caballus mitochondrial genome, strain: trotter	GenBank	GenBank: AB859014.1
Oryx dammah mitochondrial genome	Hassanin et al., 2012	GenBank: JN632677.1
Capra hircus isolate 64_Ch65 mitochondrial genome	GenBank	GenBank: KR059208.1
Capra hircus isolate 985 mitochondrial genome	GenBank	GenBank: KJ192236.1
Capra hircus mitochondrial genome	RefSeq	GenBank: NC_005044.2
Capra hircus mitochondrial genome	GenBank	GenBank: KP662715.1
Capra ibex mitochondrial genome	Hassanin et al., 2009	GenBank: NC_020623.1
Capra nubiana mitochondrial genome	Hassanin et al., 2009	GenBank: FJ207527.1
Ovis aries mitochondrial genome	GenBank	GenBank: AF010406.1
Ovis aries mitochondrial genome, Assaf breed	GenBank	GenBank: HE577850.1
Ovis orientalis mouflon mitochondrial genome	Lv et al., 2015	GenBank: NC_026063.1
Ovis orientalis ophion mitochondrial genome	GenBank	GenBank: KF312238.2
Sus scrofa mitochondrial genome	GenBank	GenBank: AJ002189.1
Sus scrofa mitochondrial genome	GenBank	GenBank: AF034253.1
Camelus bactrianus mitochondrial genome	GenBank	GenBank: AP003423.1
Eptesicus fuscus mitochondrial genome	GenBank	GenBank: MF143474.1
Pipistrellus kuhlii mitochondrial genome	GenBank	GenBank: KU058655.1
Rousettus aegyptiacus mitochondrial genome	RefSeq	GenBank: NC_007393.1
Rhinolophus Monoceros mitochondrial genome	GenBank	GenBank: AF406806.1
Megaderma lyra mitochondrial genome	GenBank	GenBank: MG586969.1
Hipposideros armiger mitochondrial genome	GenBank	GenBank: JN980966.1
Rhinopoma hardwickei mitochondrial genome	GenBank	GenBank: AF263231.1
Tibetan high-altitude sheep dataset	Sichuan Agricultural University	SRA: SRR5149616
Software and Algorithms		
Analysis code	This paper	https://github.com/neuhofmo/Dead_Sea_Scrolls
MergeReadsFastQ_cc.py	Kircher, 2012	https://bioinf.eva.mpg.de/fastqProcessing/
FilterUniqSAMCons_cc.py	Kircher, 2012	https://bioinf.eva.mpg.de/fastqProcessing/
bwa	Li and Durbin, 2009	http://bio-bwa.sourceforge.net/
cutadapt	Martin, 2011	https://cutadapt.readthedocs.io/en/stable/
Bowtie	Langmead et al., 2009	http://bowtie-bio.sourceforge.net/index.shtml
Bowtie2	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
Fastqc	Andrews, 2010	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
FastQ Screen	Wingett, 2015	https://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/

(*Continued on next page*)

Please cite this article in press as: Anava et al., Illuminating Genetic Mysteries of the Dead Sea Scrolls, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.046

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
SGA	Simpson and Durbin, 2012	https://github.com/jts/sga
bamtools	Barnett et al., 2011	https://github.com/pezmaster31/bamtools
Seqtk	Li, 2016	https://github.com/lh3/seqtk
Mira	Chevreux et al., 1999	http://www.chevreux.org/projects_mira.html
MITObim	Hahn et al., 2013	https://github.com/chrishah/MITObim
SAMtools	Li et al., 2009	http://www.htslib.org/
bedtools	Quinlan and Hall, 2010	https://bedtools.readthedocs.io/en/latest/
Geneious	Biomatters Ltd.	https://www.geneious.com/
BLAST	Altschul et al., 1990	https://blast.ncbi.nlm.nih.gov/Blast.cgi
MAFFT	Katoh and Standley, 2013	https://mafft.cbrc.jp/alignment/software/
RAxML	Stamatakis, 2014	https://cme.h-its.org/exelixis/web/software/raxml/index.html
Phylobayes	Lartillot et al., 2009	http://www.atgc-montpellier.fr/phylobayes/
Biopython	Cock et al., 2009	https://biopython.org/
gargammel	Renaud et al., 2017	https://github.com/grenaud/gargammel
ococo	Brinda et al., 2016	https://github.com/karel-brinda/ococo
Mega7	Kumar et al., 2016	https://www.megasoftware.net/
MATLAB R2017b	MathWorks	https://www.mathworks.com/products/new_products/release2017b.html
Circa	OmGenomics	http://omgenomics.com/circa/
MEGAHIT	Li et al., 2015	https://github.com/voutcn/megahit
MetaBAT2	Kang et al., 2019	https://kbase.us/applist/apps/metabat/run_metabat/release
dRep	Olm et al., 2017	https://drep.readthedocs.io/en/latest/
CheckM	Parks et al., 2015	https://ecogenomics.github.io/CheckM/
Remap	NCBI	https://www.ncbi.nlm.nih.gov/genome/tools/remap
pysam	Pysam developers	https://github.com/pysam-developers/pysam

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Oded Rechavi (odedrechavi@gmail.com).

Materials Availability

This study did not generate new unique reagents.

Data and Code Availability

The accession number for the sequencing data reported in this paper is SRA: SRP151790. The code used for the processing and analysis is available in the project's GitHub repository, at https://github.com/neuhofmo/Dead_Sea_Scrolls.

METHOD DETAILS

Sampling of the specimens

The scrolls were sampled by the team of the DSS Conservation Laboratories of the IAA. While sampling the scrolls, the conservators used gloves and a delicate scalpel to cut small pieces ($\leq 0.5 cm^2$) from the edge of the scroll fragments where there was no writing. In cases where the scrolls were in a poor state of preservation, crumbs were scraped off from the back side of the scroll using a scalpel.

Please cite this article in press as: Anava et al., Illuminating Genetic Mysteries of the Dead Sea Scrolls, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.046

The samples were collected into clean Eppendorf tubes for further analysis in a lab dedicated to aDNA. Overall, 39 specimens were sampled (for the list of sampled specimens and produced libraries see Tables S1 and S2). Furthermore, four pieces of IAA parchment prepared from sheep and cattle skin, part of an IAA funded research, were sampled under similar conditions, as a positive control of our species identification methods. The IAA parchments were sampled at the Casali Institute, Hebrew University of Jerusalem. The leather objects from other sites in the Judean Desert were sampled by the Organic curators and conservators of the IAA.

Ancient DNA methods

We generated genomic sequence data from 35 ancient and four modern specimens, as well as mitochondrial capture data from 3 pieces of ancient scroll fragments. All samples were prepared in a facility dedicated to aDNA work at the Evolutionary Biology Centre in Uppsala, Sweden. The work in this facility is separated from work on modern and amplified DNA. The lab has an airlock, positive air pressure, is fitted with UV-lights, and is thoroughly cleaned with bleach and RNase AWAY surface decontamination (ThermoFisher Scientific, #7002, eliminate RNase and DNA from laboratory surfaces) to remove contaminating DNA. To further minimize the risk of contamination all work is conducted in a laminar flow hood by specially trained staff wearing protective clothing (face mask, full body suit, and triple gloves).

DNA Extraction

Multiple extracts and libraries were generated from most of the samples (Table S2); these were processed in a random order to avoid batch effects that could influence the results. Due to the delicate nature of the material (see Figure 2A), DNA was initially extracted using several silica-based methods to determine the most suitable method for these degraded samples. The testing phase was done on fragments: Unknown-001, Unknown-002, Unknown-003, 4Q405-518 and 11Q17. The testing phase included: i) Extraction buffer as in (Malmström et al., 2007) combined with Roche columns (Roche, 5114403001) according to Meyer et al. (2012) and Dabney et al. (2013b). ii) Silica based method as in Yang et al. (1998) with modifications as in Günther et al. (2015) and Sawyer et al. (2012) (AY/AY*; Table S2) or iii) Silica based method as in Allentoft et al. (2015), Dabney et al. (2013b), and Meyer et al. (2012) (MPBR; Table S2). All extractions were eluted in 44-110 μl of EB or TET buffer. To remove contamination a 30 min washing step with 1ml EDTA or EDTA with 1% Triton X-100 was included. DNA was extracted also from this fraction (Malmström et al., 2007) (WAY for EDTA or WAY* EDTA with 1% Triton X-100; Table S2). All pieces were extracted overnight. The MPBR method has previously been shown to capture more short fragments compared to other silica-methods (Allentoft et al., 2015; Dabney et al., 2013b; Meyer et al., 2012), and our own experiments confirm this observation.

A second phase silica-based extraction was done as in Dabney et al. (2013b) and Meyer et al. (2012) (GR; Table S2) followed as described above. A 30 min washing step with 1ml EDTA was employed here as well, followed by DNA extraction from this fraction (Malmström et al., 2007) (WGR; Table S2). In this phase we processed the following fragments: 4Q57 frg. 27, 4Q57 frg. 6v, 4Q57 frg. 9v, 4Q71-nt, Mas1k, 4Q404 frg. 5v, 11Q17, 4Q405-36 and NewScroll-41.

For leather specimens a third phase silica-based extraction was done as in Dabney et al. (2013b) and Meyer et al. (2012) with some modifications designed for hair and leather (Gilbert et al., 2004) (GRL; Table S2) followed as described above. In this phase we processed the following specimens: Sandal-96-9105, Garment-99-9035, Sandal-43-6809, Sandal-35-1270, Sandal-43-6811, WS4, WS7, WS(4or7). Each round of extraction included 2 blank tubes which were processed according to the different phase of extractions.

Library Preparation

DNA libraries were prepared using 20μl of extract, with blunt-end ligation coupled with P5 and P7 adapters and single indexes described in Meyer and Kircher (2010) with modifications as in Günther et al. (2015). From each extract one to five double strand libraries were built. Since aDNA is already fragmented the shearing step was omitted from the protocol. Library blank controls including water as well as extraction blanks were carried along during every step of library preparation. Each blunt-end library was amplified in 4 to 12 replicates with one negative PCR control per index-PCR. The amplification reactions had a total volume of 25 μl, with 3 μL of DNA library, and the following in final concentrations; 1X AmpliTaq Gold Buffer, 2.5mM $MgCl_2$, 250nM of each dNTP, 2.5U AmpliTaq Gold (Thermofisher, #N8080259), and 200nM each of the IS4 primer and P7-index primer (Meyer and Kircher, 2010). The optimal number of cycles was determined by qPCR. The negative controls were also run on qPCR in order to obtain amplification. Libraries were amplified between 14-22 cycles using conditions as in Meyer and Kircher (2010). For each library four amplifications with the same indexing primer were pooled and purified with AMPure XP beads (Beckman Coulter, A63881). Libraries were quantified on Tapestation using the High Sensitivity Kit (Agilent Technologies, #5067-5584, #5067-5585). In total, 15 extraction blanks were passed through the entire protocol. In order to get anything to amplify in these blanks, we needed to run 27-30 PCR cycles (compared to between 14 and 22 cycles for the reported samples). Three of the 15 blanks showed no amplification, not even adapter-dimers. 12 of the 15 blanks showed amplification and presence of DNA (as determined from qPCR, PCR and Tapestation), with a distinct adapter-dimer peak (around 115-125 bp in size). Of these 12, five showed a peak with an insert (around 200 bp), which had a lower intensity compared to the adapter-dimer peak. Four of these blanks with an insert were sequenced to determine the origin of the insert DNA. The sequence data of these four blanks were mapped against hg19 (human), sheep, cow, and goat, and the genome coverage was < 0.0004x for all blanks for all four species (compared to a mean/median genome coverage of the target animal of 0.01x/0.006x across 286 sample libraries), and the clonality was $\geq$ 0.95. To further examine any potential

Please cite this article in press as: Anava et al., Illuminating Genetic Mysteries of the Dead Sea Scrolls, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.046

contaminant, perhaps from reagents, the number of sequence reads aligning to the mt of hg19 (human), sheep, cow, and goat were tallied. For Blank1, we obtained 5,278,436 sequence reads, and 6 sequence reads mapped to hg19 mt, none to the other animals' mts. For Blank2, we obtained 58,062,897 sequence reads, and 13 sequence reads mapped to hg19 mt, 5 to the sheep mt, 1 to the goat mt and 2 to the cow mt. For Blank3, we obtained 617 sequence reads, none of the which aligned to any of the animals mts. For Blank4, we obtained 8,912,671 sequence reads, and none mapped to ht19 mt, 32 to the sheep mt, 30 to the goat mt and 41 to the cow mt, all mapping to the same region. We conclude that the majority of the DNA in the blanks were adapter-dimers and non-mammalian contaminants. For initial screening, 10 libraries were pooled at equimolar concentrations for sequencing on an Illumina HiSeq 2500 using v.4 chemistry and 125 bp paired-end reads and HiSeq X10 using v.X chemistry and 150 bp paired-end reads at the SNP & SEQ Technology Platform at Uppsala University. The average/median coverage for the identified species after filtering reads aligned to human (see Remapping, additional trimming and filtering of human contamination), was 0.01x/0.006x respectively. Differences in sequenced data are derived from the variation in the size and quality of the sampled specimens (see Tables S2 and S3 and Figure S1A).

Mitochondrial Capture

To enrich for mitochondrial DNA, capture was done by using commercial myBaits® Mito (Arbor biosciences, Ann Arbor, MI, USA, #303048) for: Cow (*Bos taurus*), goat (*Capra hircus*) and sheep (*Ovis aries*). Capture experiments were conducted according to the myBaits protocol v.4.01 for degraded DNA libraries with some modifications. Two libraries were captured in each reaction (½ of a bait reaction per library). Hybridization step was done at 60°C for 42 hours. Library amplification was done in 50 μL reaction using KAPA2G Fast HotStart ReadyMix (Merck Sigma-Aldrich, KK5601). Three amplification reactions were made for each library using 5 μL of library per reaction. Libraries were pooled at equimolar concentrations and sequenced as described above. The captured mitochondria were then only used for species identification (see *Jeremiah fragments species identification using mitocapture*).

Quality control, trimming and merging of NGS data

Only sequences with the correct index were included in the processing of data to safeguard against contamination, possibly occurring outside the clean room facilities. Paired-end reads were trimmed and merged using MergeReadsFastQ_cc.py (Kircher, 2012) if an overlap of at least 11 base pairs was found.

Mapping NGS data based on established ancient human pipeline against whole genomes

We first processed the sequence data using standard and well-established parameter settings for working with ancient human DNA sequences (e.g., Günther et al., 2015 and references therin). Merged reads were mapped single ended with bwa aln 0.7.13 (Li and Durbin, 2009) to the human reference genome (build 37, hs37d5) using the following non-default parameters: seeds disabled -l 16500 -n 0.01 -o 2 (Lazaridis et al., 2014; Skoglund et al., 2014b) as well as to the sheep *Ovis aries* (Oar_v4.0, GenBank: GCF_000298735.2), the goat *Capra hircus* (ASM170441v1, GenBank: GCA_001704415.1), and the cow *Bos taurus* (UMD3.1, GenBank: GCA_000003055.5). Each sample was then merged at library level using SAMtools (v 0.1.19) (Li et al., 2009) before removal of PCR duplicates to ensure maximum retention of reads. To remove PCR duplicates, reads with identical start and end positions were collapsed using a modified version of FilterUniqSAMCons_cc.py (Kircher, 2012). Only sequence reads longer than 35 bp, with a mapping quality of at least 30, and with a minimum of 90% agreement with the reference were retained.

The majority (60.6%) of ancient specimens examined show one or more features characteristic of degraded DNA (Briggs et al., 2007; Hansen et al., 2001; Hofreiter et al., 2001; Sawyer et al., 2012): (i) short fragments (50-60 bp) and (ii) cytosine deamination consistently higher at fragment termini (Hansen et al., 2001; Sawyer et al., 2012). The process of making parchments includes chemical treatments that likely affect DNA deamination. The ancient parchments display typical aDNA deamination which shows a gradual decline in deamination while moving from the terminal base further into the fragment. In some samples, with remarkably low amounts of data, we did not detect smooth damage curves, but the signal of greater levels of deamination toward the fragment ends was still present. The fragments with the largest amount of data (Unknown-002 and Mas1k) have standard postmortem damage patterns. In contrast, in the IAA parchments, we observe a clear deamination signal for the terminal base at the DNA fragments. However, this deamination does not appear to extend into the fragment. This pattern may represent DNA lesions which are the result of the harsh treatment applied to these parchments in the process of their preparation (see Figure S2).

The reads mapped to sheep, cow, and goat, depending on the animal of origin, were generally shorter than the reads mapped to human (Figures S1B and S1C). See Species identification and Alignment of WGS DNA to organisms for more details regarding the reads processing and alignment.

As the amount of retained sequence reads was low for many samples, we explored various possibilities of retaining more sequence data from the ancient samples. The low levels of retained sequence information were not unexpected; the standard criteria for ancient human sequence have been set relatively strict, the quality of the reference sheep genome and reference goat genome are much lower than the human reference genome, and the levels of diversity and differentiation among Ovis and Capra species a few thousand years ago are unknown. All these factors will result in loss of sequence data when using stringent mapping conditions.

Please cite this article in press as: Anava et al., Illuminating Genetic Mysteries of the Dead Sea Scrolls, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.046

Remapping, additional trimming and filtering of human contamination

The fastq files generated in Quality control, trimming and merging of NGS data were processed with an additional stage of adapter trimming and quality control using cutadapt version 1.1 (Martin, 2011) with the parameters -n 6 -q 26. We removed additional repetitive sequences (such as TTTTTTTTTTT; AAAAAAAAAAA; CCCCCCCCCCC; GGGGGGGGGGG) from both ends to avoid low quality reads, contaminants and reads with low information content. To remove human contamination (see Figures S1B and S1C), the data was then aligned to the human reference genome (GRCh38.p6, GenBank: GCF_000001405.32) using bowtie2 v2.2.8 (Langmead and Salzberg, 2012) with the –very-sensitive flag (Chang et al., 2015). The quality of the samples was assessed using Fastqc v0.11.5 (Andrews, 2010).

To illustrate the effect of the human contamination on the read lengths, we compared the mean length of the following groups: reads mapped to human ("Aligned to human"), reads mapped to sheep after removing human contamination ("Filtered"), and reads mapped to sheep without removing human contamination ("Unfiltered"). The mean length was calculated based on the .sam files produced by bowtie2 (Figure S1B).

Species identification

Species identification was done in the following parallel methods.

Method 1 for identifying species

Alignment to various nuclear and mitochondrial reference genomes. The data from Mapping NGS data based on established ancient human pipeline against whole genomes were mapped to the whole genomes mentioned in that section as well as the following mitochondrial genomes: *Bos taurus* (GenBank: V00654.1), *Bos taurus* (GenBank: KJ709685.1), *Capra hircus (*GenBank: KR059208.1*)*, *Capra hircus* (GenBank: KJ192236.1), *Capra hircus* (GenBank: NC_005044.2), *Equus asinus* (GenBank: NC_001788.1), *Equus caballus* (GenBank: NC_001640.1), *Oryx dammah* (GenBank: JN632677.1), *Ovis aries* (GenBank: AF010406.1), and *Sus scrofa* (GenBank: AJ002189.1) in order to determine the species of origin. The resulting mitochondrial and whole genome coverages were then compared across species to infer the species for each sample (see Tables S4B and S4C).

FastQ Screen. In order to increase the resolution and be able to include additional species we used FastQ Screen (Wingett, 2015) which has been previously used in identifying species used for making parchments (Teasdale et al., 2015). It allows screening of a library of sequences against a user-determined sequence database, thus allowing for easy detection of the sequence composition in a library.

Merged fastq files (Quality control, trimming and merging of NGS data) were pooled per fragment. Since all merged reads were used, the reads were pre-processed using SGA (v.0.10.14) (Simpson and Durbin, 2012). All reads shorter than 35 nt were removed, and the SGA's dust threshold was set to the default 4. Using FastQ Screen (v.0.9.2) (Wingett, 2015) together with Bowtie (v.1.1.2) (Langmead et al., 2009) and bwa mem (v.0.7.13) (Li and Durbin, 2009). The following genomes were included in the sequence whole genome database: Human (build 37), *Ovis aries (*Oar v4.0, GenBank: GCF_000298735.2), *Capra hircus* (CHIR_1.0, GenBank: GCA_000317765.1), *Bos taurus* (UMD3.1, GenBank: GCA_000003055.5), *Equus Caballus* (EquCab2, GenBank: GCA_000002305.1), *Sus scrofa* (Sscrofa9, GenBank: GCA_000003025.1), *Gallus gallus* (WASHUC2, GenBank: GCA_000002315.5). Performing FastQ Screen on whole genomes yielded inconclusive results, due to high portions of human contamination, hence, we focused on the mt genomes consistent with the standard procedures of species identification. The mitochondrial genome database included: human (rCRS, GenBank: J01415.2), *Ammotragus lervia* (GenBank: NC_009510.1), *Bos Taurus* (GenBank: V00654.1), *Camelus bactrianus* (GenBank: AP003423.1), *Capra hircus* (GenBank: KJ192236.1), *Equus asinus* (GenBank: NC_001788.1), *Oryx dammah* (GenBank: JN632677.1) and *Ovis aries* (GenBank: AF010406.1). The results can be found in Table S4B.

De novo assembly of the mitochondria. To minimize the biases introduced when mapping reads to a reference genome, we wanted to confirm our results through a *de novo* assembly approach. The analysis was restricted to reads that mapped to either *Ovis aries* (GenBank: AF010406.1), *Capra hircus* (GenBank: KJ192236.1, GenBank: KR059208.1 or GenBank: NC_005044.2) or *Bos taurus* (GenBank: KJ709685.1) mitochondria, using the methods described in Mapping NGS data based on established ancient human pipeline against whole genomes. These reads were then converted back into fastq format using bamtools (v.2.3.0) (Barnett et al., 2011), and assembled *de novo* using the method described in Hahn et al. (2013). In the specific case of fragment Unknown-002, in order to assemble the mitochondria, the 244,787 reads which were converted back to fastq format had to be randomly subsampled down to 50,000 reads using Seqtk (v1.0) (Li, 2016) with the s-parameter set to 100. We used Mira 4.0.2 (Chevreux et al., 1999) for the initial seeding step using *Sus scrofa* (GenBank: AJ002189.1), thereafter MITObim (Hahn et al., 2013) was used on top of Mira with kmers set to 25. For those samples with enough reads to perform the *de novo* assembly (namely Unknown-001, Unknown-002, Unknown-003, 4Q57 frg. 7v, Mas1k, NewScroll-39, NewScroll-40, NewScroll-41, 4Q59 frg. 25, Hev/Se6, WS(4or7), 4Q405 and 11Q17), the resulting contigs were blasted (Altschul et al., 1990) against the NCBI nucleotide database (November 2016) using default parameters (see Table S4C).

Method 2 for identifying species

Alignment of mitochondrial DNA to organisms. We pooled together shotgun sequencing libraries produced from the same specimen and removed the first and last nucleotide of every read using a custom script. The pooled libraries were then mapped to the human genome as described in Remapping, additional trimming and filtering of human contamination. Similarly, reads from mitocapture protocols of the same specimen (Mitochondrial Capture) followed the same procedures.

Please cite this article in press as: Anava et al., Illuminating Genetic Mysteries of the Dead Sea Scrolls, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.046

The reads which could not be aligned to the human genome were saved (using the bowtie2–un parameter) and then aligned to 11 different mitochondrial reference genomes of wild and domestic animal species common in the Near East: *Bos taurus*: GenBank: V00654.1; *Camelus bactrianus*: GenBank: AP003423.1; *Capra hircus*: GenBank: KP662715.1; *Capra ibex*: GenBank: NC_020623.1; *Capra nubiana*: GenBank: FJ207527.1 (Hassanin et al., 2009); *Equus asinus*: GenBank: X97337.1; *Equus caballus*: GenBank: AB859014.1; *Ovis aries*: GenBank: HE577850.1; *Ovis orientalis mouflon*: GenBank: NC_026063.1; *Ovis orientalis ophion*: GenBank: KF312238.2; *Sus scrofa*: GenBank: AF034253.1. The alignment was performed using bowtie2 v2.2.8 (Langmead and Salzberg, 2012) with the –very-sensitive flag. As all the specimens tested mapped to either sheep, goat, or cow, we focused on these three species in our downstream analysis.

Mitochondrial DNA coverage calculation. For each mapping analyses, the bowtie2 SAM files produced in *Alignment of mitochondrial DNA to organisms*, were converted to BAM using the view and sort programs of SAMtools v1.3.1 (Li et al., 2009). The per-base coverage of the mitochondrial genome was then calculated from the BAM using bedtools genomecov with the –d flag (Quinlan and Hall, 2010). The coverage was visualized with Geneious 9.0.5 (Biomatters Ltd.), as shown in Figure 4F. The percentage of bases with non-zero coverage out of the reference mitochondrial genome was then calculated with Geneious 9.0.5 and consensus sequences were extracted. Depending on the sample, the best coverage was obtained for either the sheep (most samples), the goat (e.g., Sandal-35-1270), or the cow (e.g., NewScroll-41, Garment-43-6809) reference sequence. Mitochondrial sequences of different Caprini species are highly similar in size and length; nevertheless, in most of the DSS examined, the percent of the sheep (*Ovis aries*) and mouflon (*Ovis orientalis*) mito-genome that was covered by aligning sequences was higher than the coverage of goat (*Capra hircus*), and the rest of the species examined (see Figures 4A–4C and 4F and Table S4A). Our results align with earlier studies from the 1950s and 1960s which suggested, based mostly on microscopic inspection of fibers in the skin, that many DSS derive from sheep (Ryder, 1958), although results were not always conclusive (Poole and Reed, 1962; Reed and Poole, 1964). The different consensus sequences obtained were used as query to perform a BLAST (Altschul et al., 1990) search against the NCBI nucleotide database (March 2017). In some cases, the reference sequence used for the mapping did not influence the species identification and the different consensus sequences gave the same BLAST result (either sheep, cow or goat). In a few cases (namely specimens 4Q404 frg. 8v, 4Q404 frg. 9v, 4Q344 and some of 4Q57 fragments), BLAST searches of consensus sequences obtained from mapping to the goat or the cow mt genome best matched goat or cow (although the coverage was lower than the coverage obtained for sheep). This suggests that DNA contaminants from other animals could be present in some of the scroll samples (see Figures 4A–4C and Tables S4 and S5). When only reads that align to one organism were retained, only one species was detected (see Figure S4B and *Filtration of mitochondrial sequences from contaminants*). The traces from different animal sources in the scrolls may be due to the process of their preparation in antiquity. It is known from literary sources that methods for rendering skins into parchments entailed the use of materials from various organic sources (for example, for tanning) (Campana et al., 2010; Haran, 1985; Poole and Reed, 1962).

Filtration of mitochondrial sequences from contaminants. For each sample produced in Alignment of mitochondrial DNA to organisms, we aligned the consensus sequences, obtained from the sheep, goat and cow mapping, using MAFFT v7.308 (Katoh and Standley, 2013) under the L-ins-i algorithm. In the absence of contaminants, we expect the *p*-distance between pairs of sequences to be null, although the coverage of the mitochondrial genome would be different. As a case in point, a scroll made of sheep, would present a greater number of reads mapping to the sheep reference, less reads mapping to the goat reference (only in regions that are conserved between these two species) and the least amount of reads mapping to the cow reference (since sheep and cow are more distant and share less homologous regions than sheep and goat).

However, as we expect only one mitochondrial genome within each individual, the reads mapping to these different sequences should be the same. Consequently, we infer that the presence of *p*-distances between sequences higher than 0.01 would indicate the presence of contaminants. In this case, an additional filtration was performed by removing all the reads that mapped to the putative contaminant. For example, if the *p*-distance between the cow and sheep consensus obtained is 0.01, all reads that mapped to the cow sequence were removed from the sheep mapping, and vice versa. If the distance was lower than 0.01 the sequences were considered to be identical and the consensus with the lower amount of ambiguous characters was retained (see schematic drawing of the process in Figure S4B).

After filtration, only consensus sequences showing a reference-sequence coverage higher than 20% were included in the phylogenetic analyses (in *Generation of mitochondrial phylogenetic tree*). Since the mtDNA control region of sheep and goat contain repeated elements whose number vary among and within individuals (Gorkhali et al., 2016; E et al., 2018; Parma et al., 2003; Wood and Phua, 1996), an issue which can affect read mapping, the mtDNA control region was excluded when computed the distance between sequences.

Jeremiah fragments species identification using mitocapture. To further investigate the origin of the Jeremiah fragments in this study, we performed a mito-capture procedure, which yielded higher mt DNA content relatively to the shotgun sequencing (see Mitochondrial Capture). For the mito-capture process, we used commercial baits designed for cow, sheep, and goat mitochondrial genomes. We aligned the reads for each one of the three animals and detected only negligible amount of reads that aligned to goat. To control for potential contamination and misalignment (see *Filtration of mitochondrial sequences from contaminants*), we focused only on the reads aligned to either cow or sheep, but not the other. We also removed reads that aligned to bat species, which might have inhabited the Qumran caves and contaminated the samples (*Eptesicus fuscus,* GenBank: MF143474.1; *Pipistrellus kuhlii,* GenBank: KU058655.1; *Rousettus aegyptiacus,* GenBank: NC_007393.1; *Rhinolophus Monoceros,* GenBank: AF406806.1; *Megaderma lyra,*

GenBank: MG586969.1; *Hipposideros armiger,* GenBank: JN980966.1; *Rhinopoma hardwickei,* GenBank: AF263231.1). In order to distinguish between aDNA to potential modern contamination, we further focused on reads that were shorter than 75 nt. In the cases of 4Q70 and 4Q72b, we detected notably higher coverage of the cow genome (see Table S5). In addition, for 4Q72b, BLAST searches show that the consensus sequence of the reads aligned to both cow and sheep resulted in cow hits. As to 4Q72a, the coverage observed for the sheep mt genome was higher than that of the cow mt genome, however, the mild difference suggests a potential presence of amplified mitocapture baits. To investigate the source of these DNA reads, we took advantage of a blank sample which represents the DNA products acquired in the processes of DNA extraction, library preparation, mitocapture and sequencing. The blank sample contains reads that align to cow or sheep; however, while the cow reads were verified as cow using BLAST searches, the consensus sequence made from sheep reads actually resulted in Capra hits (and not Ovis hits). Thus, while the cow reads in 4Q72a might have been the result of spurious amplification, this is unlikely the case of the sheep reads, which are likely to represent the true origin of the fragment. This conclusion is also supported when restricting the analysis to regions supported by at least 3 reads, which yielded a notably higher coverage of the sheep genome. Additionally, a maximum likelihood phylogenetic tree of the three Jeremiah fragments and representatives of the Bovidae diversity confirmed these results (see Figure S3A). The phylogenetic tree is generated as in *Generation of mitochondrial phylogenetic tree* below, with the following additional constraints (to reduce the potential effect of discussed contamination): consensus sequence requires a minimum coverage of 3 reads and agreement of at least 75% of reads for every position.

Generation of mitochondrial phylogenetic tree. To determine from which animal skin the specimens were made, a phylogenetic analysis was performed; the reconstruction is based on complete mitochondrial (mt) sequences, from a representative sample of the Bovidae diversity. Only specimens which achieved at least 20% mtDNA coverage were used for the phylogenetic analysis. The reads used were produced and cleaned as described in *Alignment of mitochondrial DNA to organisms* and *Filtration of mitochondrial sequences from contaminants*.

The analyses were based on the large Cetartiodactyla dataset of Hassanin et al. (2012). To reduce computation time, only Bovidae species and their closest outgroup, the Moshidae, were considered. Additionally, together with the specimen sequences, several representative sequences of the genetic diversity of the genus *Ovis* (sheep and mouflons) (Lv et al., 2015) were added manually, since this genus was only represented by a single sequence in the original dataset. Positions with more than 50% of missing data among the aligned sequences were excluded from the alignment. The Bovidae dataset included 140 sequences and 14,867 nucleotide positions. Phylogenetic analyses were performed both under the maximum likelihood and the Bayesian criteria. Following Hassanin et al. (2012), 36 partitions were considered in the maximum likelihood analysis: the tRNA and non-coding regions, the 12S rRNA, the 16S rRNA, and the three codon positions of 11 protein partitions. All mitochondrial protein coding gene were analyzed separately except the overlapping ATP8 + ATP6 and ND4L + ND4, which were jointly analyzed. The variable control region was excluded from this dataset. The maximum likelihood analysis was performed using RAxML version 8.2.9 (Stamatakis, 2014) under the GTR+CAT model. Branch supports were computed using 100 rapid-bootstrap computations. The Bayesian analysis was performed using Phylobayes version 4.1 (Lartillot et al., 2009) under the CAT-GTR + Gamma 4 mixture model of evolution. Four independent MCMC chains were ran for 23,000 trees. After a visual inspection of the likelihood evolution of the chains the first 13,000 trees of each chain were excluded from the analysis (burn-in). The remaining trees were then sampled every 10 trees. The Bayesian consensus was then computed from the 4,000 (1000x4) sampled trees. Convergence was assessed with the bpcomp and tracecomp programs which are part of Phylobayes. The maximum and average differences in clade posterior probability among the chains were 0.296 and 0.004. Additionally, the effsize and rel_diff parameters were always higher than 150 and lower than 0.3, respectively, which indicates a correct convergence of the chain (see Figure 3).

Generation of a sheep phylogenetic tree. To better decipher the phylogenetic position of the specimen sequences, a phylogenetic analysis was performed including only members of the genus Ovis. Specifically, all Ovis complete mt sequences available in the NCBI nucleotide database (on 03 November 2016) were downloaded. Sequences were aligned using the L-ins-i algorithm of MAFFT (Katoh and Standley, 2013) as implemented in Geneious 9.0.5 (Biomatters Ltd). Control regions, which are known to include repeated elements in sheep (Gorkhali et al., 2016; E et al., 2018; Wood and Phua, 1996), were excluded from the alignment as well as positions with more than 50% of missing data. The final sheep dataset included 126 sequences and 15,437 positions. The dataset was partitioned and analyzed in the same manner as described for the Bovidae dataset. Phylogenetic analyses were performed both under the maximum likelihood and the Bayesian criteria. The reads used were produced and cleaned as described in *Alignment of mitochondrial DNA to organisms* and *Filtration of mitochondrial sequences from contaminants*.

Four phylobayes chains were ran for 22,000 trees and sampled every 10 trees. The Bayesian consensus was then computed from 3,600 sampled trees, after removing the first 13,000 trees of each chain (burn-in). The maximum and average differences in clade posterior probability among the chains were 0.19 and 0.004. Similarly, the effsize and rel_diff parameters were always higher than 100 and lower than 0.3, respectively, indicating a correct convergence of the chains.

Phylogenetic relationships recovered the monophyly of the main sheep haplogroups (Figure S3B). However, relationships among sheep haplogroups slightly differ from previous studies as haplogroup E appears nested within clade haplogroup C rather than being placed as sister clade of haplogroup C. Because of low support within haplogroup C and given the long branch of 4Q57 frg. 12, it cannot be excluded that 4Q57 frg. 12 could belong to haplogroup E rather than haplogroup C.

Generation of a cow phylogenetic tree. To characterize the mt ancestry of the two samples originating from cow based on the whole genome shotgun sequencing (NewScroll-41 and Garment-43-6809), a phylogenetic analysis was performed including only

members of the genus Bos. Specifically, representative of the different *B. taurus* haplogroups were chosen based on previous works (Achilli et al., 2008, 2009; Olivieri et al., 2015) and the tree was rooted with zebu (*Bos indicus*) sequences. Complete mt sequences were aligned using the L-ins-i algorithm of MAFFT (Katoh and Standley, 2013) as implemented in Geneious 9.0.5 (Biomatters Ltd) and trimmed from positions with more than 50% of missing data. Unlike the sheep and goat analyses, the control region, which does not include long repeated elements in cow were included in the alignment and represented by an additional partition in the ML analysis. The final cow dataset included 148 sequences and 16,340 positions. The dataset was partitioned and analyzed in the same manner described for the Bovidae dataset. Four phylobayes chains were ran for 35,000 trees. However, one of the chains did not converge and was thus excluded. The Bayesian consensus was then computed from 9,000 (3,000x3) sampled trees after removing the first 5,000 trees of each chain. The maximum and average differences in clade posterior probability among the chains were 0.189 and 0.003. Similarly, the effsize and rel_diff parameters were always higher than 50 and lower than 0.3, respectively, which indicates a correct convergence of the analysis.

Phylogenetic relationships recovered the monophyly of the main cow haplogroups P, Q, R and T (Figure S3C). Relationships among the different T haplogroups (T1-4) are, however, not supported. Although relationships among the T1 and T3 haplogroups are unresolved, the two cow samples Garment-43-6809 and IAA parchment NewScroll-41 appear to be nested within the T1 haplogroup (Figure S3C) which is frequent among extant Egyptian specimens (Olivieri et al., 2015).

Generation of goat phylogenetic tree. To characterize the mt ancestry of sample Sandal-35-1270 originating from goat, a phylogenetic analysis was performed including members of the genus *Capra*, *Hemitragus*, and *Peudois*. Specifically, all *C. hircus* and *C. aegagrus* complete mt sequences available in the NCBI nucleotide database (on 30 June 2017) were downloaded. Complete mt sequences were aligned using the L-ins-i algorithm of MAFFT (Katoh and Standley, 2013) as implemented in Geneious 9.0.5 (Biomatters Ltd). Similar to the sheep analysis, the control region, which include repeated elements in goat (Parma et al., 2003) were excluded. The final goat dataset included 137 sequences and 15,429 positions after excluding positions with more than 50% of missing data. The dataset was partitioned and analyzed as indicated above. The four phylobayes chains were ran for 41,000 trees. The Bayesian consensus was then computed from 15,200 sampled trees, after removing the first 3000 trees. The maximum and average differences in clade posterior probability among the chains were 0.27 and 0.004, respectively. Similarly, the effsize and rel_diff parameters were always higher than 50 and lower than 0.3, respectively, except for the rel_diff parameters nmode = 0.630 and allocent = 0.692. To further confirm that the convergence was sufficient to address the position of Sandal-35-1270 we compared the consensus tree obtained with the first 7,600 sampled trees with the consensus obtained with the last 7,600 trees of the four chains. In both cases the support for the monophyly of the different goat haplogroups, and the relationships within these haplogroups remained the same.

Phylogenetic relationships recovered the monophyly of the main goat haplogroups (Figure S3D). While most goats in the world belong to haplogroup A, the sample Sandal-35-1270 was surprisingly found to cluster within haplogroup G (Figure S3D). Members of haplogroup G are found nowadays in Turkey, Iran and Egypt (Ahmed et al., 2017; Colli et al., 2015). Our results thus suggest that haplogroup G was more widespread in the past and that it is not the result of small-scale domestication event as it has been suggested (Colli et al., 2015).

Alignment of WGS DNA to organisms

The reads which could not be aligned to the human genome using the method described in Remapping, additional trimming and filtering of human contamination were saved (using the bowtie2 –un flag) and then aligned to seven different reference genomes of the following organisms: *Bos taurus*, GenBank: GCF_000003205.5; *Camelus bactrianus*, GenBank: GCF_000767855.1; *Capra hircus*, GenBank: GCA_000317765.2; *Equus asinus*, GenBank: GCF_001305755.1; *Equus caballus*, GenBank: GCF_000002305.2; *Ovis aries*, GenBank: GCF_000298735.2; *Sus scrofa*, GenBank: GCF_000003025.5. The alignment was performed using bowtie2 v2.2.8 (Langmead and Salzberg, 2012) with the parameters: –local -D 22 -R 4 -N 1 -L 18 -i S,1,0.50 to enable flexible alignment (see Alignment and filtering protocol for higher read retention for the motivation behind flexible alignment). The reads were processed and PCR duplicates were removed using the script FilterUniqueSAMCons_rand.py (Kircher, 2012) as in Mapping NGS data based on established ancient human pipeline against whole genomes. As controls, the same process was done for unfiltered reads (reads before alignment to the human reference genome) and for the subset of reads which aligned to the human genome (using the bowtie2 –al flag) (see schematic drawing in Figure S4A). Furthermore, for the various analyses, we also applied additional filters on the read lengths and mapping quality. The exact values chosen are specified for each analysis in the appropriate section.

Genomic DNA coverage calculation

The bowtie2 SAM output files, produced after the alignment of the processed reads as described before (see Mapping NGS data based on established ancient human pipeline against whole genomes and Alignment of WGS DNA to organisms) were used for the coverage calculation. Coverage for each base was calculated using the pysam module for Python 3.6.2 (https://github.com/pysam-developers/pysam; based on SAMtools; Li et al., 2009) and Biopython (Cock et al., 2009) and using a designated wrapper script. For each sample, only positions where all the covering reads had the same nucleotide value were considered, to reduce bias caused by contamination.

Processing of sheep outgroup

In order to validate our analyses which are based on genetic variation, we added an outgroup in the form of a WGS data retrieved from a single high-altitude sheep in Tibet (SRA: SRR5149616, Tibetan Sheep), which should be phylogenetically distant both from ancient

Please cite this article in press as: Anava et al., Illuminating Genetic Mysteries of the Dead Sea Scrolls, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.046

and modern sheep in the region. A subset of the reads (50%) was downloaded from NCBI SRA database, passed the same alignment process as the rest of the samples and then used in the clustering described in the following sections.

aDNA data simulations using gargammel

We used gargammel (Renaud et al., 2017) with the following parameters: –comp 0.85,0,0.15 -n 20000000 –loc 4 –scale 0.5 -rl 150, based on the components described below:

a) To determine the composition of each sheep sample, we examined the average number of sheep reads (15%), human reads (17%) and bacteria reads (68%) in all of the sheep specimens studied in our analysis.
b) To determine the bacterial composition within the bacteria fraction, we analyzed the bacterial species using de-novo assembly of bacterial scaffolds and blast again NCBI database, as described in Metagenomic analysis. We then picked the bacterial species common to at least five specimens. We added the bacterial genomes from NCBI and distributed their proportion uniformly within the bacterial fraction.
c) To determine the levels of deamination and its pattern, we used mapDamage with default parameters, based on Mas1k, a scroll fragment with high levels of information. We adjusted the soft-clipping notation in the respective SAM file to properly represent read edges.
d) To determine the read length distribution, we analyzed the read length distribution in the samples and found the approximated log-normal distribution representation (rounded).

Clustering the sheep specimens based on variation in their genomic data sequences

Out of the 39 specimens sequenced, we were able to identify 30 as likely to be made of sheep, using one of the following criteria:

1) Their Ovis mitochondrial coverage was higher than 20% and the phylogenetic analysis (Figure 3; Species identification) placed them within the Ovis clade.
2) Their Ovis mitochondrial coverage was higher than the coverage reached when mapped to any other organism, and their mitochondrial coverage (when normalized to the specie with the highest coverage) is higher by at least 10% than the coverage of any other reference mitochondria (see Figures 4A–4C and 4F; Table S4A).
3) One exception to the above rules is 4Q71-nt, for which the Ovis was preferable to any other species by only 5%; however, since 4Q71-t has reached the abovementioned requirement, and as we know for sure that 4Q71-nt and 4Q71-t belong to the very same specimen, we included both of them in our analysis.

In order to identify variation in genomic sequences between the various samples, we first implemented a variant calling procedure using the ococo v.0.1.2.3 (Břinda et al., 2016) on the BAM produced in Alignment of WGS DNA to organisms (the number of variants identified for each specimen can be seen in Figure S5C). Next, for each individual sample, we recognized a set of positions in which the nucleotide identity that derived from the WGS deviates from that of the *Ovis aries* reference genome (GenBank: GCF_000298735.2). To avoid false positive positions, we restricted the set of positions to contain only those that are covered by at least 3 individual reads. Furthermore, we demand full agreement on the identity of the nucleotide in the position of interest among all the reads of a given sample, as well as full agreement on the nucleotide identity between different samples of the same specimen. Having a set of positions in which at least one specimen shows deviation from the nucleotide identity of the reference genome, we next turned to characterize the extent of information embedded in each such position for all the specimens analyzed. Specifically, for a given fragment, an informative position is defined as a position showing either high-confident deviation of the nucleotide identity from the reference (as described above), or is covered by at least 3 reads, where all of them contain the same nucleotide as in the reference. Then, we looked for positions that were found to be informative for more than 50% out of the 30 examined sheep specimens and the additional outgroup. Based on these positions, we created a vector of the corresponding nucleotides, clustered the fragments and generated a tree using the Maximum Likelihood criterion under the GTR-GAMMA model, using RAxML version 8.2.9 (Stamatakis, 2014), with the command -m ASC_GTRGAMMAX –asc-corr = lewis -p 82331 -f a -x 34839 -N 500. 500 Bootstrap computations were performed in order to assess the reliability of the branches of the constructed tree. In order to reduce non-genetic bias potentially caused by the antiquity of the DSS fragments and other ancient specimens, we excluded genetic variation which might be derived from cytosine deamination and restricted our analysis to transversions only. Tree visualization was done using Mega7 (Kumar et al., 2016; Figure 5). For this analysis, we used 91 libraries that belonged to 30 specimens originated from sheep (as identified in *Species identification* sections, see Table S4 and Figure 4A-C).

In addition, we used the nucleotide vectors (used for generating the tree in Figure 5) to measure the amount of information per specimen (number and percent of informative positions), as displayed in Figures S5B and 5D and Table S6. We also calculated the variants detected when mapping to the sheep genome (as they appear in the VCF file), to measure the effect of the overall deviation from the reference detected for every specimen, as displayed in Figure S5C. The figures were produced using custom python scripts.

Trees based on partial sets of the specimens, as shown in Figure S6B, were generated similarly. For this purpose, we set the minimal coverage threshold to 5, 7 and 10 of the specimens (compared to 16 specimens in Figure 5). As a result, when the minimal coverage threshold is low, the sequence representing each specimen is longer and contains a relatively higher fraction of

non-informative positions. To avoid the potential effects of gapped sequences on the clustering, we further restrict a minimal informative positions per specimen to 40%. Hence, the number of the specimens in the trees presented in Figure S6B is positively correlated with the coverage threshold.

To control for the effects of possible human contamination, we focused on the human-aligned reads found in our samples. In parallel to the generation of the tree presented in Figure 5, we used a similar "high-yield" pipeline on the human reads filtered in Remapping, additional trimming and filtering of human contamination (see Figure S4A). We applied the same criteria described above and generated a 400 nt-long vector of human-aligned reads representing deviation from the sheep reference genome (see Figure S6A [i]) and a 2,165 nt-long vector of human-aligned reads representing deviation from the human reference genome (see Figure S6A [ii]).

Alignment and filtering protocol for higher read retention

The specific sheep breeds that were herded around the Dead Sea in the Roman Period are unknown, but research suggests that they differ from the breed which populates Israel today (Gootwine, 2015). Therefore, we used relaxed alignment criteria when aligning to the sheep genome (on the filtered reads produced in Remapping, additional trimming and filtering of human contamination). We detected damage patterns in sheep reads when removing reads with Mapping Quality lower than 30 (see Figure S2C); We observed a tradeoff between the relaxation of the alignment criteria and the quality of the damage patterns, and found that, as might be expected, allowing additional mismatches (as in the "high yield" protocol) overshadows the striking damage pattern, and does not allow, by definition, to unequivocally rule out the possibility that the clustering is affected by modern sequences. Nevertheless, the correct clusters which emerged were constructed solely based on sheep-aligning sequences (except for the case of WS7, clustering based on human or bacteria-aligning sequences does not reveal meaningful clusters, see Figures S6A and S7C), and modern sheep DNA, unlike modern human contaminant DNA from people handling the scrolls, is unlikely to have contaminated the samples.

Piece-to-piece analysis

In parallel to the analysis described in Clustering the sheep specimens based on variation in their genomic data sequences and using the data from Alignment of WGS DNA to organisms, we also compared each possible pair of specimens (pairwise comparison), hence benefitting from the maximum amount of information shared by each two specimens (termed henceforth as piece-to-piece). For this analysis we used only reads with length > 30nt and MapQ $\geq$ 22 (in addition to the standard removal of PCR duplicates). These parameters were chosen as they allow higher read retention while keeping the general characteristics of the original library (see Figure S5A and stage 3 in Table S3). We verified that the base quality for the positions used was high (BaseQ = 60 for more than 95% of positions).

The piece-to-piece analysis is done in several steps. First, we created a set of informative positions for each specimen as described in Clustering the sheep specimens based on variation in their genomic data sequences. Similarly, to avoid false positive positions, we restricted the set of positions to those that are covered by at least 3 individual reads with full agreement among all the reads of a given library. Next, for the set of informative positions, we created a matrix representing the nucleotide identity (either the reference or alternative) in each fragment, while demanding full agreement on the nucleotide identity between different samples of the same specimen. For example, for specimen 4Q404 frg. 8, we used 6 libraries; we included a given position in our analysis only if there are no conflicts regarding the nucleotide identity among the libraries. Of course, we did not require any agreement between different specimens (for instance, the informative positions of specimen 4Q frg. 8 were determined independently of the rest of the 4Q404 specimens).

Then, for each pair of specimens, we extracted the set of positions for which both specimens have reliable information on the nucleotide identity, and at least one of the two specimens shows deviation from the nucleotide identity of the reference genome. Finally, for each pair of specimens we calculated the fraction of positions in which the two specimens have an identical alternative nucleotide, resulting in n X n similarity matrix, where n = number of specimens. We then ran PCA analysis using MATLAB R2017b on the similarity matrices for the sole purpose of visualization (Figures 6, S6C, and S6D). In addition, we present the level of similarity between selected specimens, which we used for validation, to the rest of the specimens (see Figures 6 and S6C). In this analysis, 72.69% out of the specimen pairs had at least 1,000 shared position (see Figure S5E). As a control, a similar analysis was made on reads that aligned to human (see Figure S7A for comparison), to rule out bias by human contamination.

The genetic clustering based on sheep DNA is robust to external factors

We examined whether the scroll fragments might also cluster based on technical characteristics or DNA contaminations:

First, the scrolls do not cluster together simply based on the amount of sequence information retained from each sample either in terms of global coverage or in terms of the coverage along the shared positions (for details see Figures 7 and S5B–S5E and Table S6). Second, to consider the possibility that some contaminant sequences, which partially align to the sheep genome, contribute to driving the clustering patterns that we observe, we performed several additional control analyses (described above), including clustering based on non-sheep aligning sequences (human and bacterial, see Figures S6A, S6C, and S7). When we attempted to cluster the specimens based on reads that align to the human genome, we did not detect patterns that agree with the relatedness of the fragments in antiquity. This analysis failed to identify the connections and clashes between specimens which are predicted to cluster together or apart based on complementary lines of evidence, such as textual content, scribal character, paleographic profiles, and historical considerations (Figures S6A and S7A). Indeed, the human-aligning sequences and the absence of damage patterns for these sequences, suggest that the contaminant human DNA does not arise from the DNA of the people who made the

Please cite this article in press as: Anava et al., Illuminating Genetic Mysteries of the Dead Sea Scrolls, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.046

scrolls, but rather from people who handled the scrolls in more recent times, during the 70 years that passed since their discovery (Figures S1B, S1C, and S2). The two waterskin samples did however share similar patterns of human contamination, likely since they were held and handled separately from the other samples in modern times.

Moreover, we performed a metagenomics analysis (see below), classifying the different bacteria associated with each fragment. As expected, specimens 4Q71-t and 4Q71-nt, which derived from the very same scroll fragment, share similar bacteria population. However, as to the whole set of specimens, clustering based on the reads that align to bacteria do not agree with the clusters which are based on the variation in the sheep genome, suggesting that our main results are not governed by potential contamination (Figure S7C).

In summary, the DNA of different contaminants, modern or ancient, does not appear to drive and/or bias our results, and focusing on differences in the sheep's nuclear genome sequences can enable sorting of the ancient specimens.

Circos display of shared positions between specimens

In order to generate a circos plot displaying the shared positions between the specimens, we generated two .bed files: The first file determined the order of the specimens in the circos plot and their size; the sizes represent the number of positions in which each specimen has coverage of at least 3 reads and differs from the sheep reference genome. The second file represents the number of shared positions between each pair of sheep specimens, which was calculated based on the number of shared positions and the similarity score described in *Piece-to-piece analysis*. Rows added to the file based on the number of shared positions between the specimens, and each row indicates a relationship between random points on the two specimens. Each line represents 200 shared positions. Both files were produced using a custom python script, and then used in the Circa program to generate a circos plot (see Figure 7).

Metagenomic analysis

We assembled reads into contigs using MEGAHIT (v1.1.3) (Li et al., 2015). MEGAHIT was run using the -kmin-1pass flag and '-m 0.95'. We binned contigs from each sample using MetaBAT2 (v2.12.1) (Kang et al., 2019) which groups samples based on tetranucleotide frequency. MetaBAT2s default minimum contig length of 1500bp was relaxed to 1kbp.

Contig bins isolated from different samples may represent the same microbial specie. We dereplicated all contig bins using dRep (v2.2.3) (Olm et al., 2017). We assessed the quality of dereplicated bins using CheckM (v1.0.18) (Parks et al., 2015). Contig bins with less than 80% genome completeness or more than 10% contamination were discarded. We were left with 17 bins which we refer to as Metagenome Assembled Genomes (MAGs). We also used CheckM to roughly assess the taxonomic lineage of each genome.

We mapped reads from each sample to the MAGs using Bowtie2 (v2.3.0, very sensitive local preset; Langmead and Salzberg, 2012). We assessed read coverage across each genome using SAMtools and a custom script to aggregate SAMtools results (https://github.com/dcdanko/gimmebio/tree/master/gimmebio/assembly). The results of this analysis are displayed in Figure S7C.

Linkage Disequilibrium analysis

The goal of the Linkage Disequilibrium (LD) analysis is to control for the potential effect of dependence between variable positions. We created a "genetic map" of the Oar_v4.0 genome that represents the genetic distance of each position from the beginning of the chromosome. For this purpose, we remapped the Illumina OvineSNP50 Genotyping BeadChip positions from Oar_v1 (Supreme, 2013) to Oar_v4 using NCBI Remap (https://www.ncbi.nlm.nih.gov/genome/tools/remap). SNPs that were found to belong to two different chromosomes in the different genome versions, were removed from the dataset to avoid inaccurate mapping. In addition, nucleotides with genetic distance higher than both of their neighbors were removed, to reduce mapping mistakes. Between SNPs with known genetic locations, linear interpolation was used to map additional positions.

For every fragment, at every genetic location a "consensus" nucleotide was designated if it was covered by at least 4 reads after pooling together the libraries of the same specimen, while demanding full agreement on the nucleotide identity within and among the libraries representing the same specimen. The reads were produced as described in Alignment of WGS DNA to organisms.

For each pair of specimens, we calculated a similarity score by comparing all of the genetic locations that had a "consensus" nucleotide in both fragments, and the nucleotide identity of at least one of them differs from the reference genome. Each of these nucleotide locations are given a weight that corresponds to its genetic distance from the closest nucleotides (on either side) that also fit this condition. The weights allow regions with a relatively high number of covered variable positions to affect the scoring similarly to regions which have a low number of covered variable positions. That way, we minimize the bias caused by highly covered genomic regions on the global score (Kuhn et al., 2018). The cutoff for this weighting (i.e., the distance between the closest fitting nucleotide on one side and the closest one on the other) was set to be 5cM (recombination rate of 0.05, Haldane, 1919) in order to limit the influence of isolated nucleotides (see Figure S6C).

The weight of the i-th location on a chromosome which fits the mentioned condition is calculated as following:

$$d_i = \textit{genetic distance of the } i^{th} \textit{ fitting location from the } 5' \textit{ end of chromosome}$$

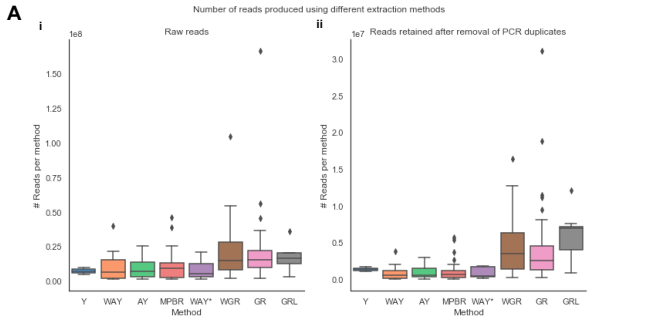
$$w = \min\{d_{i+1} - d_{i-1},\ 5cM\}$$

Please cite this article in press as: Anava et al., Illuminating Genetic Mysteries of the Dead Sea Scrolls, Cell (2020), https://doi.org/10.1016/j.cell.2020.04.046

QUANTIFICATION AND STATISTICAL ANALYSIS

All the details regarding statistical tests used in this article can be found in the appropriate figures and figure legends. In all figures displaying phylogenetic trees (Figures 3, 5, S3, S6A, S6B, and S7B), we provide bootstrap support scores. Bootstrap support scores are calculated by RAxML and PhyloBayes, as described in the figure legends and in the STAR Methods section.

Supplemental Figures



Extraction Methods:
 # Y - Silica based method as in Yang et al. 1998.
 # AY - Silica based method as in Yang et al. 1998 with modifications as in Malmstrom et al. 2007.
 # WAY - a 30 min washing step with 1ml EDTA. DNA extraction from this fraction according to AY.
 # WAY* - a 30 min washing step with 1ml EDTA with 1% Triton X-100. DNA extraction from this fraction according to AY.
 # MPBR - Silica base method as in Meyer et al. 2012, Dabney et al. 2013 and Allentoft et al. 2015.
 # GR - A second phase silica-based method as in Meyer et al. 2012 and Dabney et al. 2013.
 # WGR - A 30 min washing step with 1ml EDTA. DNA extraction from this fraction according to GR.
 # GRL - A second phase silica-based method as in Meyer et al. 2012 and Dabney et al. 2013 with modifications from Gilbert et al. 2004 designed for hair and leather.

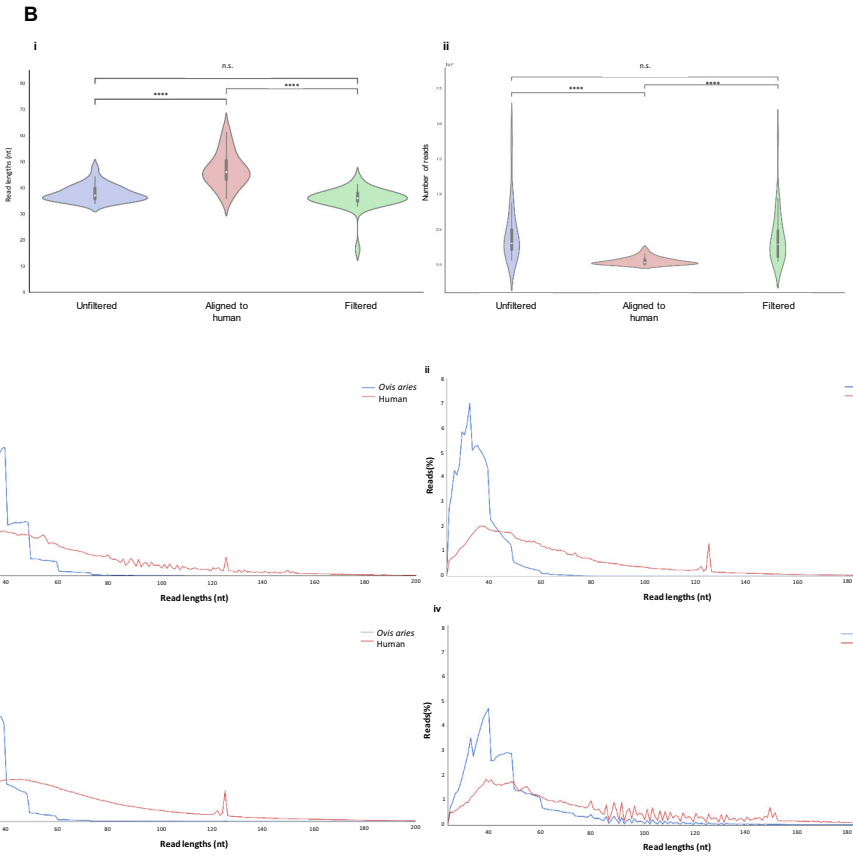
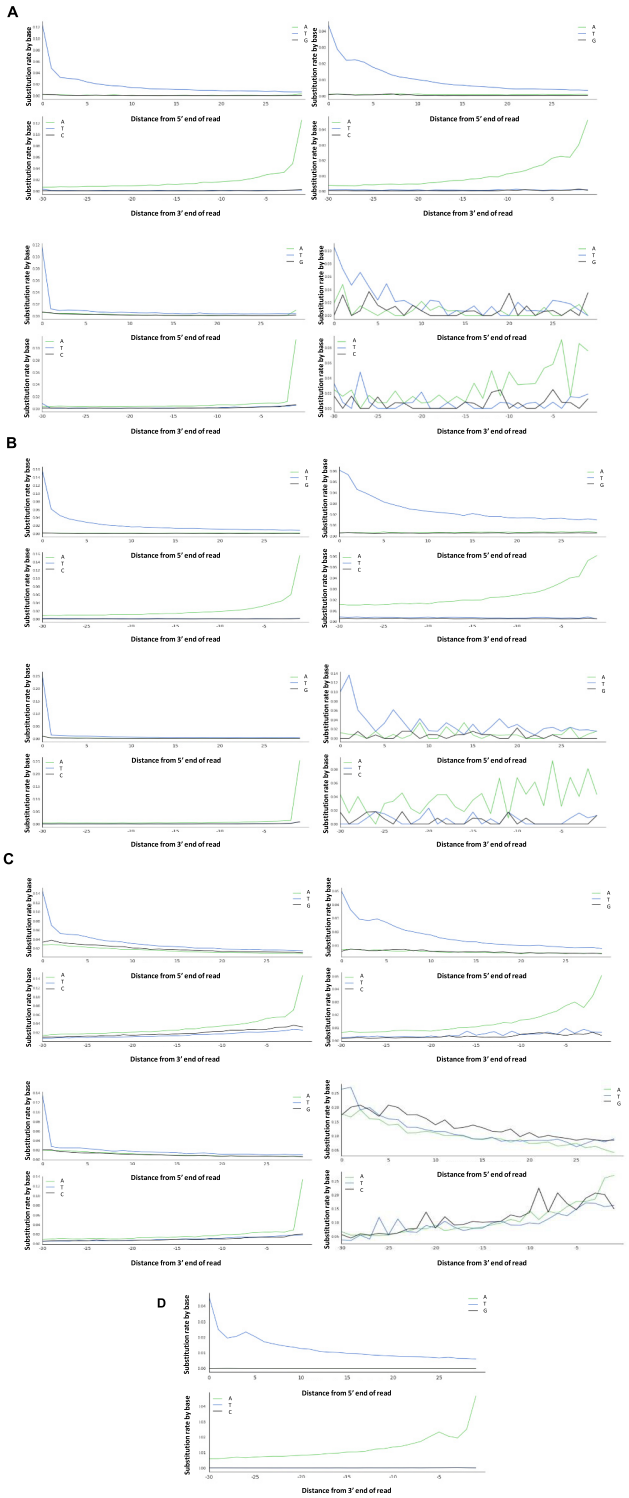


Figure S1. Characteristics of Ancient and Modern DNA Extracted from the Specimens Sampled in the Study, Related to STAR Methods

A) Comparison of the DNA reads sequenced using various DNA extraction methods. Shown are boxplots displaying the number of reads per library produced by each of eight production methods. i) raw reads sequenced for each library. ii) reads retained after the removal of human contaminants, alignment to the sheep

(legend continued on next page)

genome, and removal of PCR duplicates. During our experiments, three methods (namely WGR, GR, GRL) were shown to produce the highest number of reads and were later used in most of the following DNA extractions and thus in most of the samples used in this study. Extraction method names and descriptions can be found in the bottom panel. For more details see Tables S2 and S3. Whiskers mark 1.5 IQRs (Interquantile range: $IQR = Q_3 - Q_1$. B) The effects of removing human DNA contamination from the specimens, demonstrated for reads aligned to sheep. i) The median read length of reads aligned to sheep of the following groups: (a) all reads (unfiltered), (b) reads aligned to human and (c) reads that did not align to human (Mann-Whitney U-test: unfiltered-human: 88.5, human-sheep: 45.5). ii) The number of reads aligned to sheep of the following groups: (a) all reads (unfiltered), (b) reads aligned to human and (c) reads that did not align to human (Mann-Whitney U-test: unfiltered-human: 32, human-sheep: 74). All P-values represent Mann-Whitney test with Bonferroni correction for multiple comparisons. **** corrected p value $< 1x10^{-7}$. Calculations based on all libraries which belong to specimens identified as sheep. C) Most of the human DNA fragments extracted from the specimens are longer, and distinguishable from the ancient DNA of the animals whose skins were used for the preparation of the specimens. Read length plots show the read length distribution of the merged reads in each specimen. Shown here are read length plots for three representative specimens: (i) *Mas1k*, an ancient scroll fragment from Masada, (ii) *Hev/Se6*, an ancient scroll fragment from Nahal Se'elim, (iii) *4Q404 frg. 8v*, an ancient scroll fragment from Qumran Cave 4 and (iv) *NewScroll-041*, an IAA parchment.



(legend on next page)

Figure S2. The Ancient Scrolls, but Not the IAA Parchments, Show a Typical Cytosine Deamination Signature, Related to STAR Methods

Damage plots show estimated nucleotide misincorporation patterns. For each specimen, top panels show substitution rate of C to G/A/T in relation to the distance from the 5′ end of the reads; bottom panels show substitution rate of G to A/T/C in relation to the distance from the reads' 3′ end. A) Shown here are damage plots for four representative specimens: top left: *Mas1k*, an ancient scroll fragment from Masada; top right: *Sandal-35-1270*, an ancient sandal likely made of goat skin; bottom left *NewScroll-039*, IAA parchment, which shows a modern damage pattern with deamination in the first and last base, perhaps as a result of DNA lesions caused by the harsh parchment preparation treatment; bottom right: *Sandal-43-6811*, a sandal made of an unknown animal source, for which damage patterns appear to be noisy due to poor coverage. Interestingly, the preference of C- > T and G- > A substitutions can still be observed. The presented damage plot is based on the alignment to sheep; similar patterns were seen for cow and goat. B) Same specimens as in (A), with damage plots produced using the human standard pipeline described in the STAR Methods section. C) Same specimens as in (A), with damage plots after the removal of reads with MapQ lower than 22 and/or length lower than 30, while keeping reads with alignment length higher than 30 (see STAR Methods). A tradeoff between the relaxation of the alignment criteria and the quality of the damage patterns can be observed. The results correspond to data shown in Table S3. D) Reads simulated in similar setting to ancient samples. Shown here damage plots for a simulated aDNA library produced by gargammel (see STAR Methods). The plot presented was produced similarly to (A) and shows typical cytosine deamination signature.

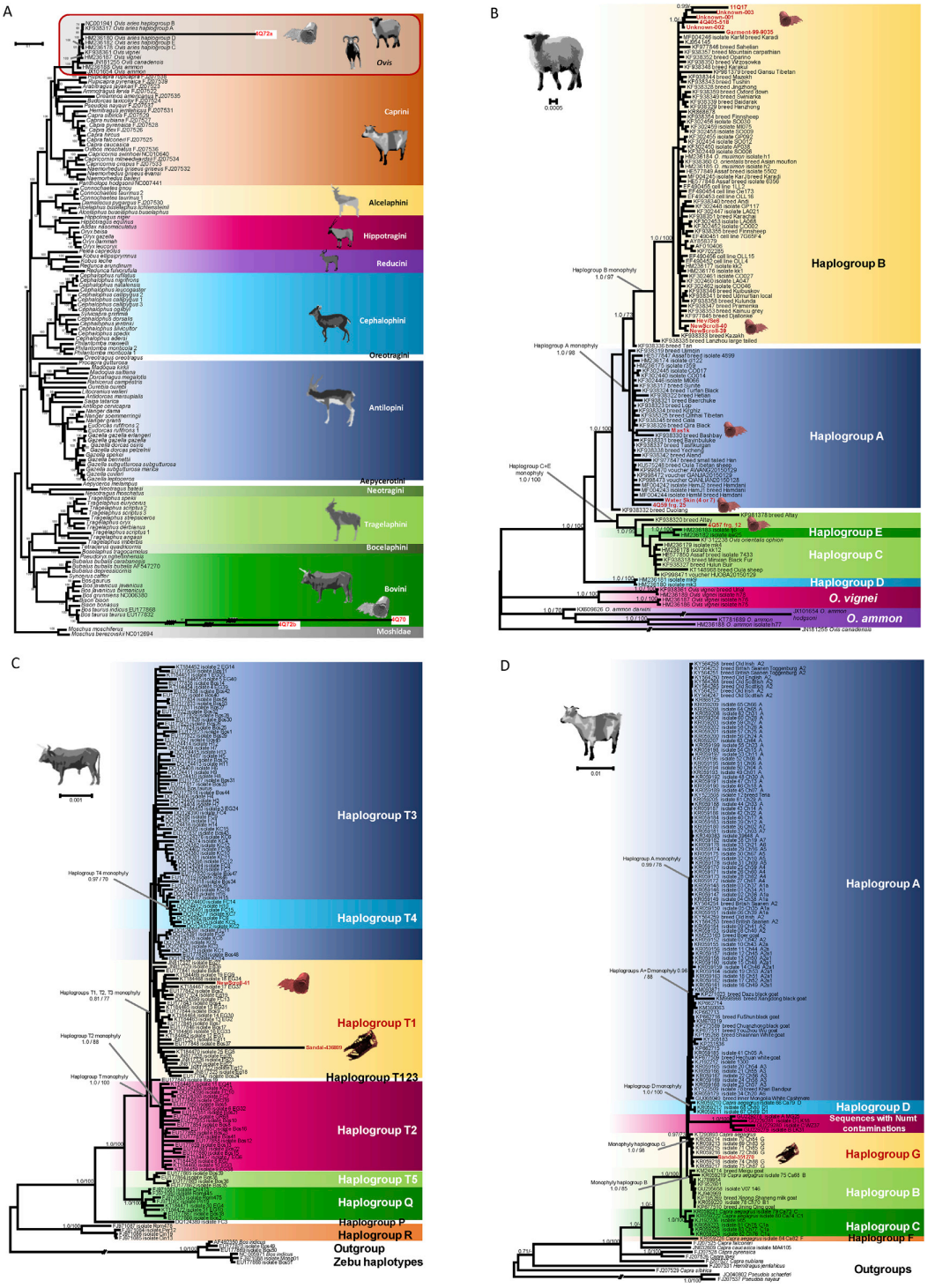


Figure S3. Phylogenetic Position of DSS Samples among Representatives of Sheep, Cow, Goat, and Bovidae Diversity, Related to Figure 3

A) Jeremiah fragments among Bovidae diversity: Phylogenetic tree generated based on Hassanin et al. (2012) dataset from a matrix of 14,867 nucleotide positions and 140 taxa using maximum likelihood criterion under the CAT-GTR model (as implemented in RaxML version 4.1). ML bootstrap supports above 50% (using 100 rapid-bootstrap computations) are given near the corresponding node. The hash marks indicate that the branches leading to 4Q72b and 4Q70 have

(legend continued on next page)

been shortened five times for visualization purposes. B) Sheep: Phylogenetic tree generated from a matrix of 126 individuals and 15,437 mt positions (excluding the control region) using Bayesian inference under the CAT-GTR model as implemented in Phylobayes. Bayesian posterior probabilities/ML bootstrap supports (using 100 rapid-bootstrap computations) are only given for relevant nodes. A minus sign (''−'') indicates that the corresponding node is absent from the ML bootstrap consensus tree. Lineage and haplogroup names follow Lv et al. (2015). The hash mark indicates that the branches leading to the outgroup and to *Ovis ammon hogdsoni* have been shorten to half for visualization purposes. C) Cow: (specimens Sandal-436809 and NewScroll-41) Phylogenetic tree generated from a matrix of 148 individuals and 16,342 mt positions (including the control region) using Bayesian inference under the CAT-GTR model as implemented in Phylobayes. Bayesian posterior probabilities/ML bootstrap supports (using 100 rapid-bootstrap computations) are only given for relevant nodes. Lineage names follow Achilli et al. (2009). The hash mark indicates that the branches leading to the outgroup have been shorten to half for visualization purposes. D) Goat: (specimen Sandal-351270) Phylogenetic tree generated from a matrix of 137 individuals and 15,429 mt positions (excluding the control region) using Bayesian inference under the CAT-GTR model as implemented in Phylobayes. Bayesian posterior probabilities/ML bootstrap supports (using 100 rapid-bootstrap computations) are only given for relevant nodes. A minus sign (''−'') indicates that the corresponding node is absent from the ML bootstrap consensus tree. Lineage names follow Colli et al. (2015). The presence of numt contaminants in some sequence was identified by Hassanin et al., 2010. The hash mark indicates that the branches leading to the outgroup have been shorten to half for visualization purposes.

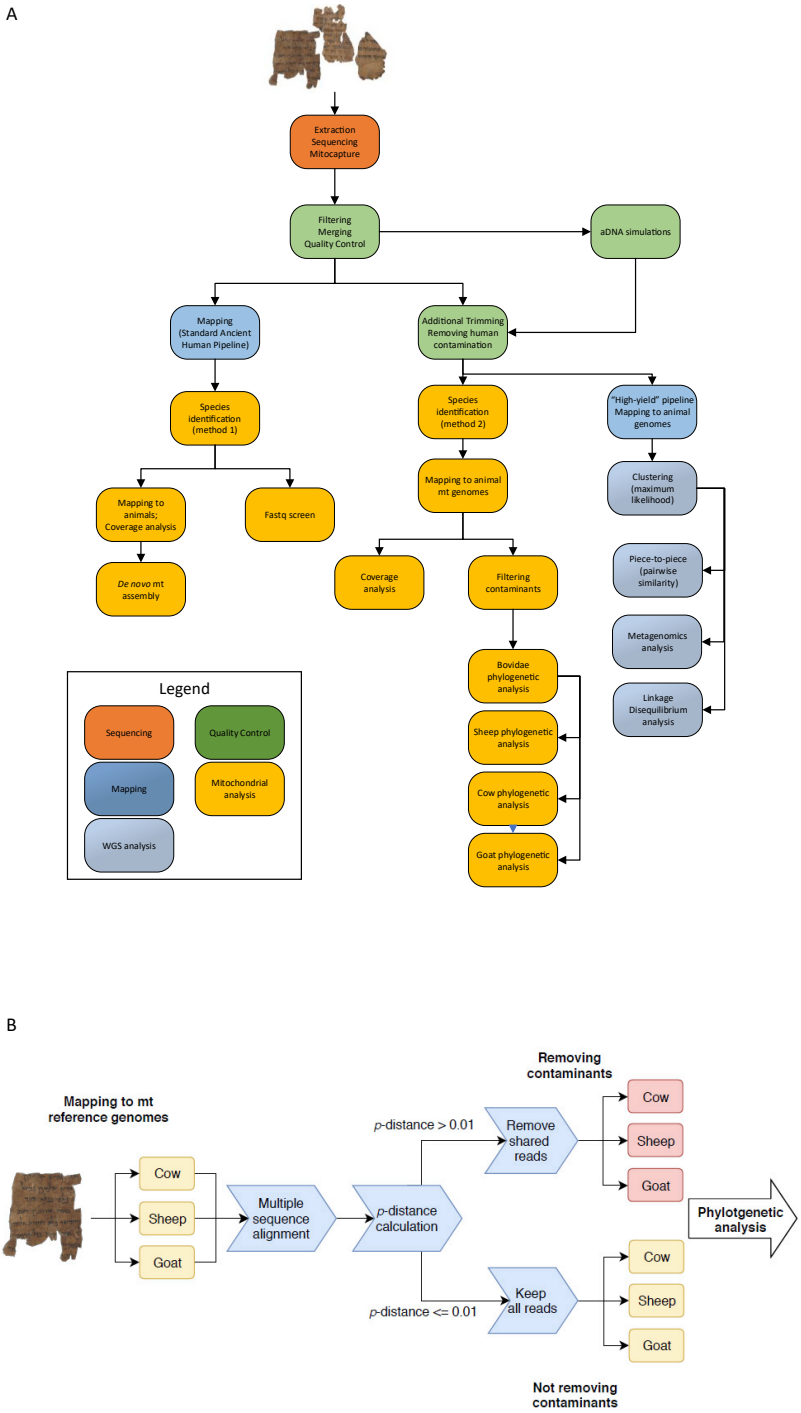
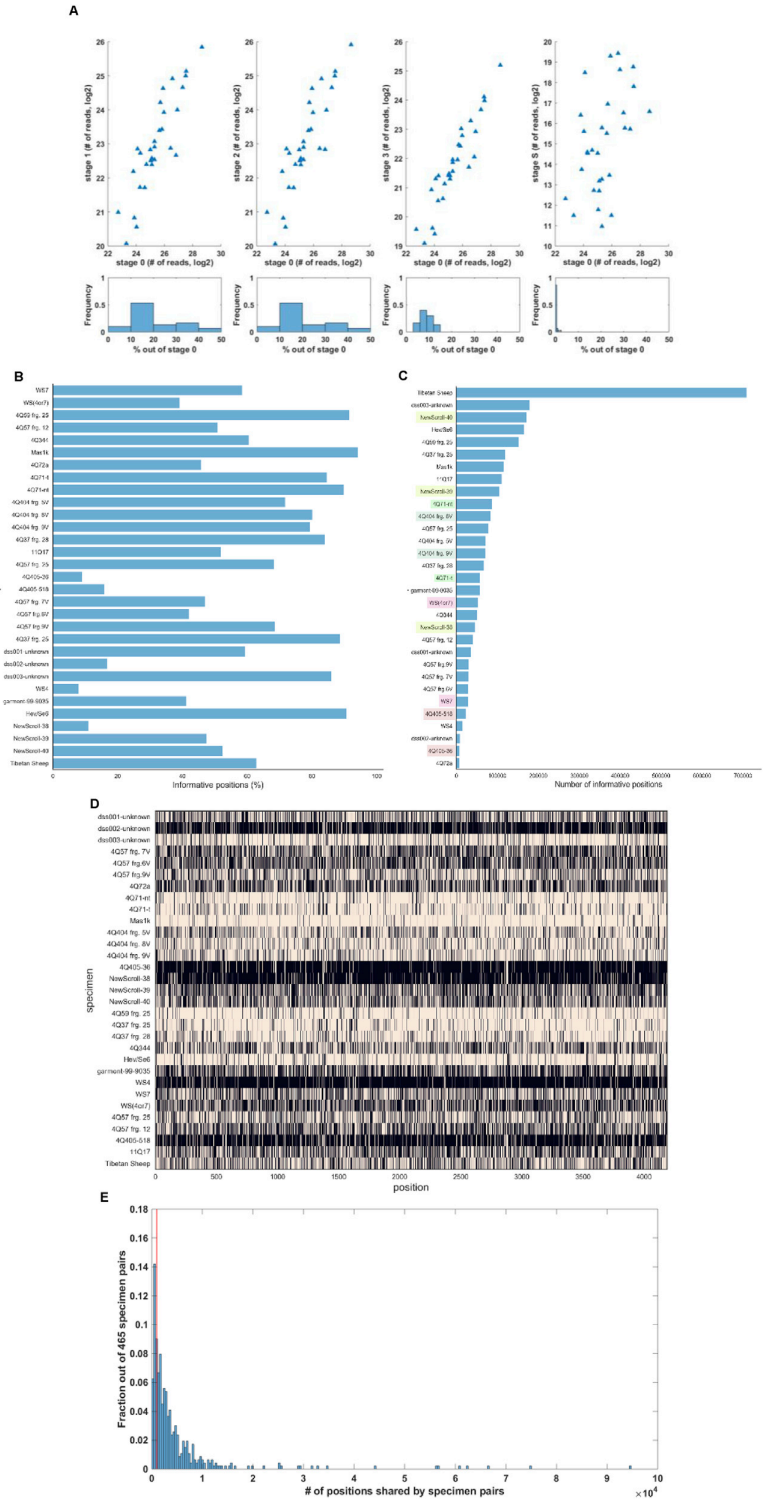


Figure S4. Schematic Drawings of Analyses Described in This Study, Related to STAR Methods

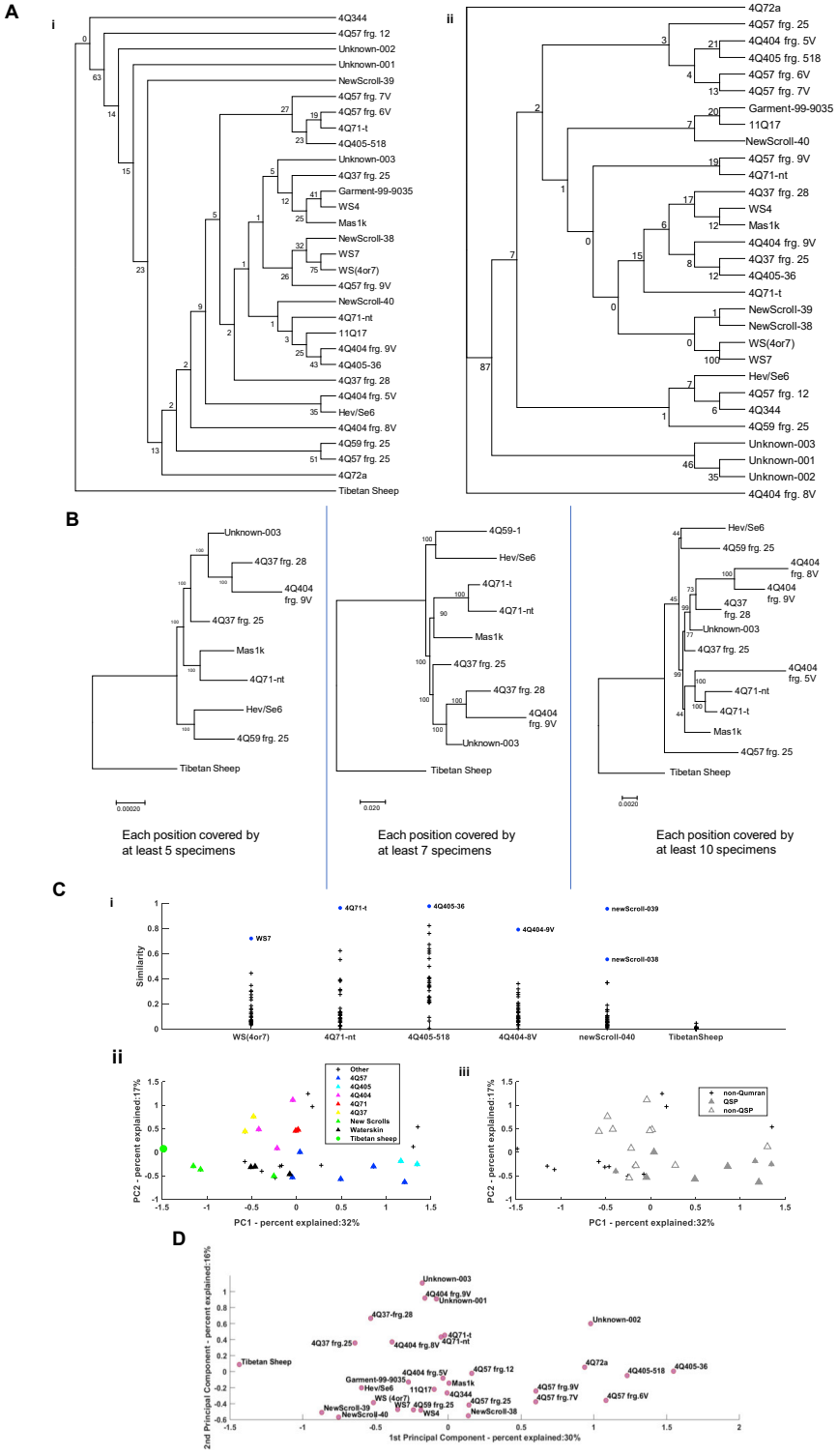
A) Schematic drawing of the major steps of the analysis pipelines performed on the specimens. Shown are the major steps used for processing and analyzing the specimens as described in the STAR Methods. B) Schematic drawing of the mitochondrial reads filtering process. The drawing describes the process of removing reads when consensus sequences, based on different mitochondrial reference genomes, show *p*-distance higher than 0.01 (as described in the STAR Methods). Only consensus sequences with coverage higher than 20% are used in the downstream phylogenetic analysis.



(legend on next page)

Figure S5. The Clustering of the Specimens Is Not Biased Because of the Amount of Information in the Specimens or the Mapping Quality of the Reads, Related to STAR Methods and Figures 5 and 6

A) High positive correlation between the amount of raw reads and the amount of processed reads along the high-yield pipeline and in the standard human pipeline. The three left panels correspond to different stages along the high-yield pipeline (the results correspond to data shown in Table S3 after summing and grouping according to each of the sheep specimens). As can be seen in the upper panel, there is a significantly high correlation between the amount of raw reads in the various specimens to the amount of the reads in each specimen after applying filtering and cleaning procedures (described in detail in the STAR Methods). Spearman correlation between the stages 1-3 to stage 0 (from left to right): 0.885, p: 3.868e-7; 0.899, p: 2.717e-7; 0.95, $p < 2.225e-100$. In addition, we compared the number of reads produced by the standard human pipeline (stage S) to the raw reads (Spearman: 0.484, p: 0.007, right panel) and the reads produced in stage 3 (Spearman: 0.503, p: 0.005). The lower panels show histograms of the percentage of reads remained after each of the processing stages compared to stage 0. As can be seen, despite the drastic decay in the number of reads induced by the standard human pipeline, we detect significant positive correlation between stage S and stages 0-3. Stage 0: raw reads; Stage 1: removal of human reads and alignment to sheep; Stage 2: removal of PCR duplicates; Stage 3: read length $\geq$ 30 nt, MapQ $\geq$ 22. Panels B-D demonstrate that the phylogenetic analysis of the specimens is not biased due to the amount of information in the specimens. B) Shown is a bar plot depicting the percentage of informative positions out of the 4,189 positions used for each fragment to generate the phylogenetic tree in Figure 5. The specimens are presented here in accordance with their location along the phylogenetic tree (see Figure 5). As can be seen, the amount of information does not direct the relationship between the different specimens. C) Shown is the number of positions with nucleotides different from the reference genome (Oar_v4.0) for each specimen. Marked are subsets of specimens which their association was demonstrated in Figure 6A (for instance, the three IAA parchments are marked in a distinctive color). As can be seen, the coupling between different specimens is not biased by the extent of deviation from the reference genome (i.e., in order to be clustered together, the two specimens must show similar variations in their covered nucleotide content). D) Shown is a visual representation of the varying amount of informative positions in the different fragments used to generate the tree in Figure 5. Positions covered by each specimen are marked in white. E) Histogram of the number of informative positions shared by specimen pairs. Shown is the distribution of the number of informative positions, shared by each of the different 465 specimen pairs analyzed in the pairwise comparison (i.e., all possible pair combinations for n = 31 specimens; see Figure 6). As can be seen, the majority of the comparisons are based on thousands of informative positions. 72.69% of the specimen pairs share more than 1000 positions. The red line marks number of positions = 1000.



(legend on next page)

Figure S6. Clustering of the Sheep Specimens Is Not Driven by Human DNA Contamination or Sparse Genomic Coverage, Related to Figures 5 and 6

A) Clustering of the sheep specimens based on variation in their genomic sequences of the human contamination. i) Clustering of human contamination reads that aligned to the sheep genome. Shown is a clustering of the specimens based on their nucleotide identity in 400 informative positions along the genome, each is covered by at least 3 reads in at least 50% of the 30 examined specimens and outgroup. ii) Clustering of human contamination reads that aligned to the human genome. Shown is a clustering of the specimens based on their nucleotide identity in 2,165 informative positions along the genome, each is covered by at least 3 reads in at least 12 out the 30 examined specimens. The nucleotide identity of at least one specimen differs from that existing in the *Ovis aries* (in (i)) or *Homo sapiens* (in (ii)) reference genome. In (ii), we looked at the variation in the human contamination compared to the *Homo sapiens* genome, hence the sheep outgroup is not represented. The tree was generated using the maximum likelihood criterion under the GTR-GAMMA model. Maximum likelihood bootstrap support values (based on 500 bootstrap computations) are indicated near the corresponding node. For the purpose of clarity, branch lengths are not reflected in the figure. B) Clustering of the specimens with the highest amount of information. Shown are phylogenetic trees (as in Figure 5), containing fragments with at least 40% informative positions out of the positions used for generating the tree (see STAR Methods for further details). As can be seen, even after strict filtering procedures, our results are consistent with the results achieved and displayed in Figure 5. The three panels differ from each other by the minimal information required per position. Here shown trees based on positions for which at least 5, 7 or 10 specimens have sufficient information. C) Similarity analysis of the various specimens using pairwise comparison of shared genomic positions while taking into account the genetic distance between adjacent positions. As can be seen, the shown results are similar to the results shown in Figure 6, suggesting that the extent of similarity between specimens is not affected by highly covered genomic regions. i) The extent of similarity between selected specimens and all other examined specimens. As can be seen, the outgroup is strikingly different from all other specimens. All compared positions covered by at least 4 reads. ii) PCA projection of the sheep specimens based on pairwise similarity matrix. Different specimen groups are marked in colors (see legend). iii) PCA projection of the sheep specimens based on pairwise similarity matrix. QSP and non-QSP fragments are marked in gray and white triangles, respectively (note legend on the right). The QSP fragments are marked in red in Figure 5. D) Similarity analysis of the various specimens using pairwise comparison of shared genomic positions. Shown is a PCA projection of the sheep specimens based on pairwise similarity matrix. The plot shown corresponds to the results presented in Figure 6. All specimens are annotated.

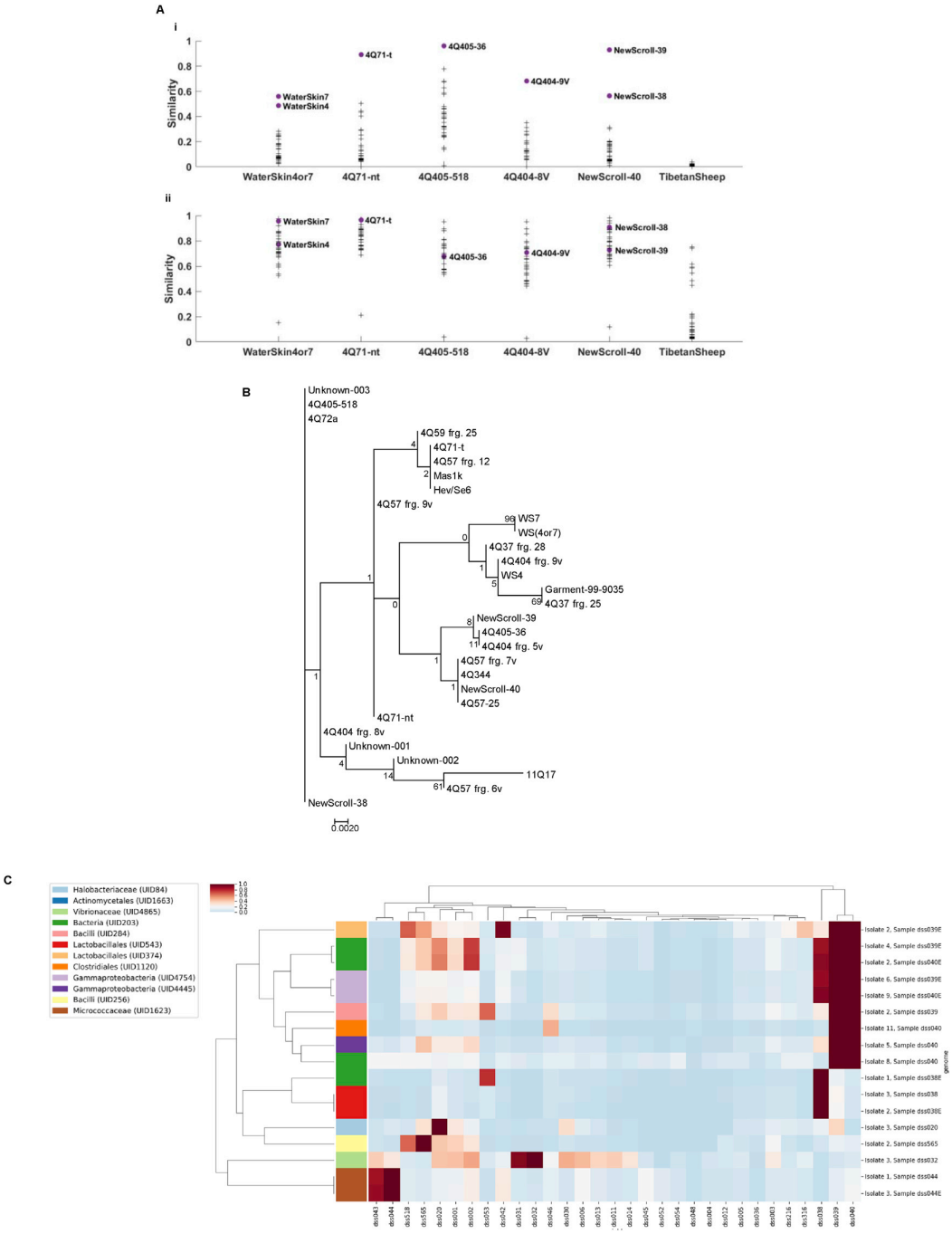


Figure S7. Human and Microbial DNA Does Not Bias the Results of Our Analyses, Related to Figures 5 and 6

A) Similarity analysis of the various specimens using pairwise comparison is not biased by contamination. i) The extent of similarity between selected specimens and all other examined specimens. As can be seen, the outgroup is different from all other specimens. The comparison is based on genomic positions covered by at least 3 filtered reads. This figure is identical to Figure 6A. ii) Same as (i), where the comparison is based on genomic positions covered by reads that are aligned to human. As can be seen, the similarity pattern governed by human contamination dissociate the examined specimens from their natural related specimens. B) Clustering of the sheep specimens based on the standard human pipeline. Shown is a clustering of the specimens based on their nucleotide identity in 161 informative positions along the genome, each is covered by at least 2 reads in at least 3 of the examined specimens. The Maximum Likelihood tree was generated as in Figure 5. As the standard pipeline does not explicitly filter contamination, some of the results may be affected, to some extent, by human or environmental

(legend continued on next page)

contamination (see The Genetic Clustering Based on Sheep DNA Is Robust to External Factors in the STAR Methods). The numbers on the nodes represent bootstrap values. C) Clustering of the specimens based on the genome coverage of Metagenome Assembled Genomes (MAGs). This heatmap shows the coverage of 17 deduplicated high quality MAGs (y axis) in each sample (x axis; for full sample names see Table S2). MAGs are named based on the sample where they were computationally isolated. The colored bar indicates the rough lineage of a given genome (see STAR Methods). As can be seen, the resulting clusters do not agree with the clusters which are based on the variation in the sheep genome, suggesting that our main results are not governed by potential contamination.