Science
AAAS

Supplementary Materials for

A genetic probe into the ancient and medieval history of Southern Europe and West Asia

Iosif Lazaridis†, Songül Alpaslan-Roodenberg†, Ayşe Acar, Ayşen Açıkkol, Anagnostis Agelarakis, Levon Aghikyan, Uğur Akyüz, Desislava Andreeva, Gojko Andrijasevic, Dragana Antonović, Ian Armit, Alper Atmaca, Pavel Avetisyan, Ahmet İhsan Aytek, Krum Bacvarov, Ruben Badalyan, Stefan Bakardzhiev, Jacqueline Balen, Lorenc Bejko, Rebecca Bernardos, Andreas Bertsatos, Hanifi Biber, Ahmet Bilir, Mario Bodružić, Michelle Bonogofsky, Clive Bonsall, Dušan Borić, Nikola Borovinić, Guillermo Bravo Morante, Katharina Buttinger, Kim Callan, Francesca Candilio, Mario Carić, Olivia Cheronet, Stefan Chohadzhiev, Maria-Eleni Chovalopoulou, Stella Chryssoulaki, Ion Ciobanu, Natalija Čondić, Mihai Constantinescu, Emanuela Cristiani, Brendan J. Culleton, Elizabeth Curtis, Jack Davis, Tatiana I. Demcenco, Valentin Dergachev, Zafer Derin, Sylvia Deskaj, Seda Devejyan, Ruben Davtyan, Vojislav Djordjević, Kellie Sara Duffett Carlson, Laurie R. Eccles, Nedko Elenski, Atilla Engin, Nihat Erdoğan, Sabiha Erir-Pazarcı, Daniel M. Fernandes, Matthew Ferry, Suzanne Freilich, Alin Frînculeasa, Michael L. Galaty, Beatriz Gamarra, Boris Gasparyan, Bisserka Gaydarska, Elif Genç, Timur Gültekin, Serkan Gündüz, Tamás Hajdu, Volker Heyd, Suren Hobosyan, Nelli Hovhannisyan, Iliya Iliev, Lora Iliev, Stanislav Iliev, İlkay İvgin, Ivor Janković, Lence Jovanova, Panagiotis Karkanas, Berna Kavaz-Kındığılı, Esra Hilal Kaya, Denise Keating, Douglas Kennett, Seda Deniz Kesici, Anahit Khudaverdyan, Krisztián Kiss, Sinan Kılıç, Paul Klostermann, Sinem Kostak Boca Negra Valdes, Saša Kovačević, Marta Krenz-Niedbała, Maja Krznarić Škrivanko, Rovena Kurti, Pasko Kuzman, Ann Marie Lawson, Catalin Lazar, Krassimir Leshtakov, Thomas E. Levy, Ioannis Liritzis, Kirsi O. Lorentz, Sylwia Łukasik, Matthew Mah, Swapan Mallick, Kirsten Mandl, Kristine Martirosyan-Olshansky, Roger Matthews, Wendy Matthews, Kathleen McSweeney, Varduhi Melikyan, Adam Micco, Megan Michel, Lidjia Milasinovic, Alissa Mittnik, Janet M. Monge, Georgi Nekhrizov, Rebecca Nicholls, Alexey G. Nikitin, Vassil Nikolov, Mario Novak, Iñigo Olalde, Jonas Oppenheimer, Anna Osterholtz, Celal Özdemir, Kadir Toykan Özdoğan, Nurettin Öztürk, Nikos Papadimitriou, Niki Papakonstantinou, Anastasia Papathanasiou, Lujana Paraman, Evgeny G. Paskary, Nick Patterson, Ilian Petrakiev, Levon Petrosyan, Vanya Petrova, Anna Philippa-Touchais, Ashot Piliposyan, Nada Pocuca Kuzman, Hrvoje Potrebica, Bianca Preda-Bălănică, Zrinka Premužić, T. Douglas Price, Lijun Qiu, Siniša Radović, Kamal Raeuf Aziz, Petra Rajić Šikanjić, Kamal Rasheed Raheem, Sergei Razumov, Amy Richardson, Jacob Roodenberg, Rudenc Ruka, Victoria Russeva, Mustafa Şahin, Ayşegül Şarbak, Emre Savaş, Constanze Schattke, Lynne Schepartz, Tayfun Selçuk, Ayla Sevim-Erol Michel Shamoon-Pour, Henry M. Shephard, Athanasios Sideris, Angela Simalcsik, Hakob Simonyan, Vitalij Sinika, Kendra Sirak, Ghenadie Sirbu, Mario Šlaus, Andrei Soficaru,

Bilal Söğüt, Arkadiusz Sołtysiak, Çilem Sönmez-Sözer, Maria Stathi, Martin Steskal, Kristin Stewardson, Sharon Stocker, Fadime Suata-Alpaslan, Alexander Suvorov, Anna Szécsényi-Nagy, Tamás Szeniczey, Nikolai Telnov, Strahil Temov, Nadezhda Todorova, Ulsi Tota, Gilles Touchais, Sevi Triantaphyllou, Atila Türker, Marina Ugarković, Todor Valchev, Fanica Veljanovska, Zlatko Videvski, Cristian Virag, Anna Wagner, Sam Walsh, Piotr Włodarczak, J. Noah Workman, Aram Yardumian, Evgenii Yarovoy, Alper Yener Yavuz, Hakan Yılmaz, Fatma Zalzala, Anna Zettl, Zhao Zhang, Rafet Çavuşoğlu, Nadin Rohland, Ron Pinhasi, David Reich

† Co-lead authors: Iosif Lazaridis and Songül Alpaslan-Roodenberg

Correspondence to: E-mail: Iosif Lazaridis (lazaridis@genetics.med.harvard.edu), Songül Alpaslan-Roodenberg (MuyesserSongul_Alpaslan@hms.harvard.edu), Ron Pinhasi (ron.pinhasi@univie.ac.at), David Reich (reich@genetics.med.harvard.edu)

This PDF file includes:

Materials and Methods

The materials and methods of this paper were shared across it and two other studies (*1, 2*) and to avoid duplication are described uniquely in the Supplementary Information of (*1*).

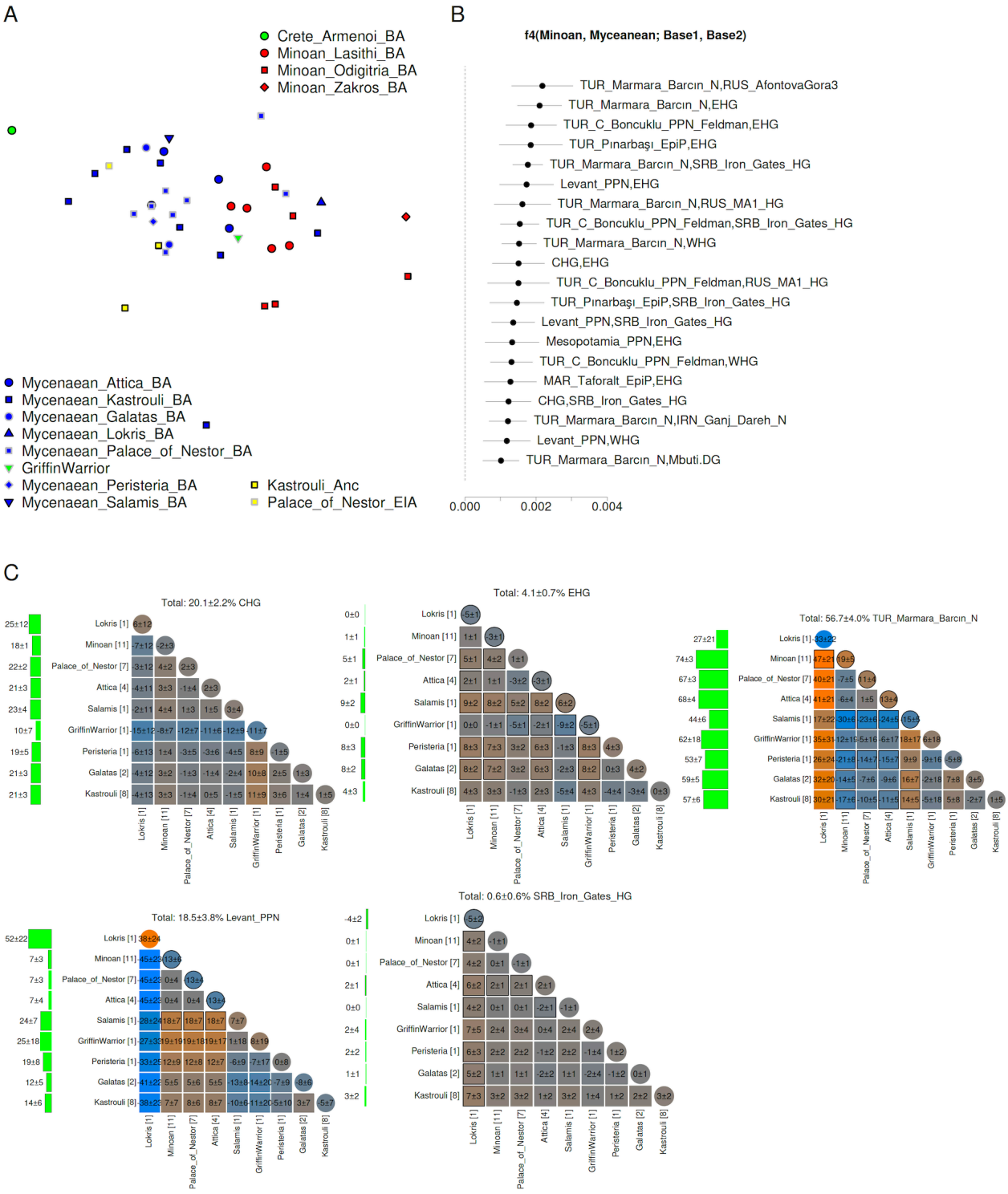
A
Crete_Armenoi_BA
Minoan_Lasithi_BA
Minoan_Odigitria_BA
Minoan_Zakros_BA
Mycenaean_Attica_BA
Mycenaean_Kastrouli_BA
Mycenaean_Galatas_BA
Mycenaean_Lokris_BA
Mycenaean_Palace_of_Nestor_BA
GriffinWarrior
Mycenaean_Peristeria_BA
Mycenaean_Salamis_BA
Kastrouli_Anc
Palace_of_Nestor_EIA
B
f4(Minoan, Myceanean; Base1, Base2)
TUR_Marmara_Barcın_N,RUS_AfontovaGora3
TUR_Marmara_Barcın_N,EHG
TUR_C_Boncuklu_PPN_Feldman,EHG
TUR_Pınarbaşı_EpiP,EHG
TUR_Marmara_Barcın_N,SRB_Iron_Gates_HG
Levant_PPN,EHG
TUR_Marmara_Barcın_N,RUS_MA1_HG
TUR_C_Boncuklu_PPN_Feldman,SRB_Iron_Gates_HG
TUR_Marmara_Barcın_N,WHG
CHG,EHG
TUR_C_Boncuklu_PPN_Feldman,RUS_MA1_HG
TUR_Pınarbaşı_EpiP,SRB_Iron_Gates_HG
Levant_PPN,SRB_Iron_Gates_HG
Mesopotamia_PPN,EHG
TUR_C_Boncuklu_PPN_Feldman,WHG
MAR_Taforalt_EpiP,EHG
CHG,SRB_Iron_Gates_HG
TUR_Marmara_Barcın_N,IRN_Ganj_Dareh_N
Levant_PPN,WHG
TUR_Marmara_Barcın_N,Mbuti.DG
0.000
0.002
0.004
C
Total: 20.1±2.2% CHG
Total: 4.1±0.7% EHG
Total: 56.7±4.0% TUR_Marmara_Barcin_N
Total: 18.5±3.8% Levant_PPN
Total: 0.6±0.6% SRB_Iron_Gates_HG
Lokris [1]
Minoan [11]
Palace_of_Nestor [7]
Attica [4]
Salamis [1]
GriffinWarrior [1]
Peristeria [1]
Galatas [2]
Kastrouli [8]

Fig. S 1 Aegean Bronze Age. (A) Mycenaean individuals from mainland Greece differ along PC1 (horizontal) of Fig. 1 from Minoan individuals from Crete, but their ancestry distributions overlap. The Griffin Warrior is within the Bronze Age variation, between the Minoan and Mycenaean distributions. (B) Highly significant f_4(Minoan, Mycenaean; Base1, Base2) statistics ($Z>5$; ±3 s.e. plotted) show that Mycenaeans share more genetic drift with Ancient North Eurasians such as Afontova Gora3(*48*) and Eastern European hunter-gatherers and Minoans more with Anatolian, Mesopotamian, and Levantine farmers. (C) The Griffin Warrior has no detectible EHG ancestry and significantly less (-6±1%) than the rest of the population of the Palace of Nestor in Pylos. Elite Mycenaeans (the Griffin Warrior and the sample from Peristeria(*4*)) do not systematically differ from other Mycenaean populations within the resolution of our 5-way model.

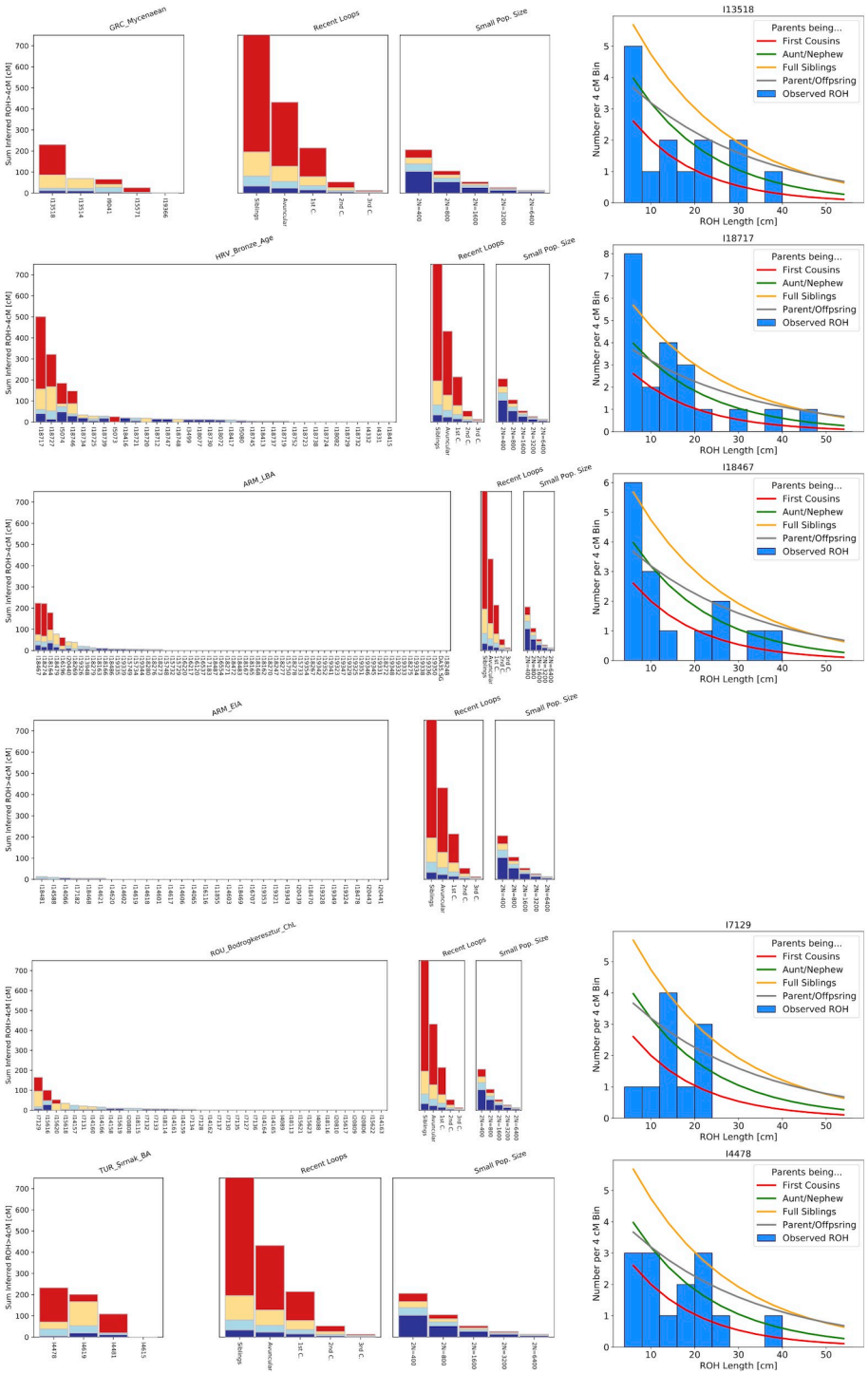

Fig. S 2 Runs of homozygosity (ROH) in the Southern Arc. We show a subset of populations for which individuals with closely-related parents are detected; on the left is the sum of ROH (blue: 4-8cM; cyan: 8-12cM; yellow: 12-20cM; red: 20-300cM) for different individuals within a population with their theoretical expectation for different classes of relatives (full siblings, avuncular, 1st-3rd cousin) and demographies. On the right, we show a histogram of ROH segments for the most inbred individual from each population.

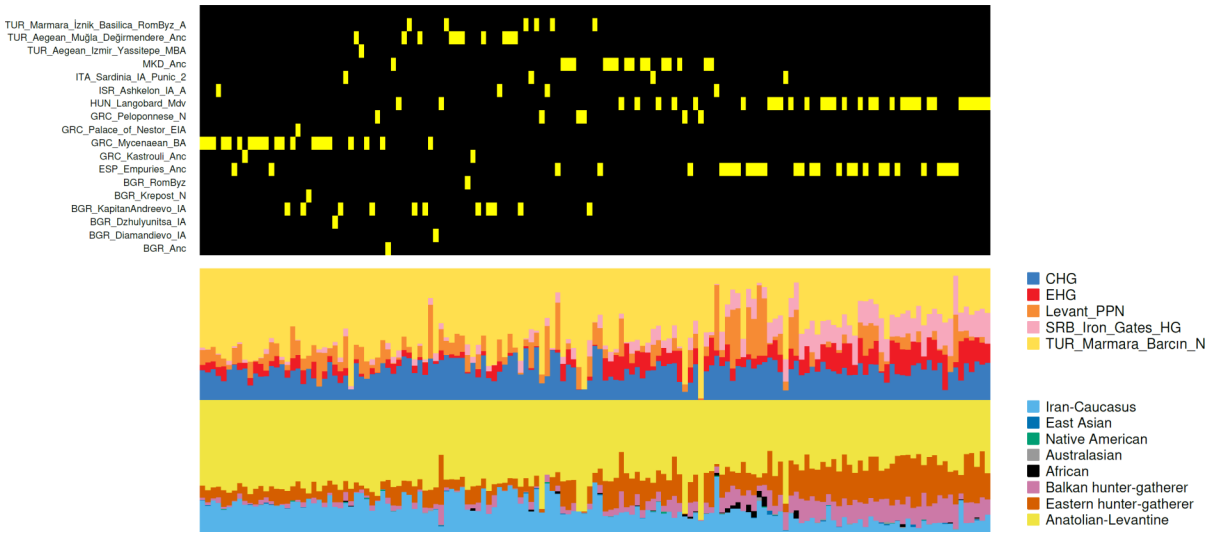

Fig. S 3 Mycenaean-like individuals. In the Iron Age, individuals similar to the Bronze Age Mycenaeans were found in Greece, North Macedonia, and Bulgaria and can be detected from Spain to Lebanon to Hungary as outliers. Individuals are ordered left-right by their rank of similarity to Mycenaeans with respect to their proportions of their five inferred ancestry components. Individuals (yellow marks on black background) are shown in conjunction with F4admix and ADMIXTURE proportions.

S1: Migrations into and out of the Southern Arc

The broad patterns of ancestry in the Southern Arc in terms of a 5-source model of admixture are explored in (*1*). This model was developed in the Supplementary Text of (*1*) using all individuals and trying to assess their ancestry as a whole and then evaluating different models in the individuals themselves. In this section we investigate some ancestry outliers of likely external origin (within the Southern Arc), and trace possible migrations from the Southern Arc beyond its borders. Our analysis here pertains to the entire dataset of (*1*) but is included in the present Supplementary Text as most of the detected outliers belong to the recent period (from the Late Bronze Age onwards) covered by it.

Non-West Eurasian ancestry in the Southern Arc

We first examined the ADMIXTURE output (*1*) to identify a set of individuals with at least 10% of their ancestry not from the four West Eurasian-related components which dominate the ancestry of Southern Arc individuals. Low levels of such admixture might be due to noise for the ancient DNA samples, and our 10% threshold is intended to identify a smaller subset for further investigation. We also excluded low coverage (marked _lc) samples.

A total of 47 Southern Arc individuals were identified according to this criterion, and we show their ADMIXTURE proportions in Fig. S 4. The African-maximized "black" component is found in Levantine individuals as early as the Natufians and should thus not be interpreted as evidence of recent African influence in West Eurasia. A likely explanation is the partial derivation of the Natufians from Paleolithic Iberomaurusian(*49*) North African-related ancestors as suggested in (*50*) Indeed, the average proportion of this component in all Natufian individuals (including those for which it is less than the detection threshold of 10%) is 9.1%, while in Taforalt from Morocco it is 41.4%, thus suggesting ~22% of North African influence, similar to the ~27% inferred using an admixture graph framework in (*50*)

The remaining outliers are driven primarily by eastern non-African ancestry, the main component of which is the "dark blue" component maximized in present-day East Asians. This includes samples from Moldova and Romania of likely "steppe nomad" derivation from the medieval period, as well as an undated sample from Moldova (I20071) of putative Middle Bronze Age, another one (I20086) of putative Eneolithic to Early Bronze Age, and an undated sample from Kalehöyük(*51*) of the Iron Age (MA2196).

Individuals from Shahr-i Sōkhta from southeastern Iran also show eastern non-African admixture, and are also notable for their possession of some "gray" (Australasian-maximized) component. This component is probably related to neighboring South Asian populations, as it is also found in the ADMIXTURE analysis in virtually all ancient samples from Pakistan.(*52*)

We highlight two other sets of outliers with non-West Eurasian ancestry.

Three outliers from Albania are also outliers of the 5-way admixture model (*1*). Their East Asian-related ancestry suggest that they may have Central Asian ancestry (which could be consistent with them being of Turkic Central Asian derivation, at least in part, given their post-medieval time frame). The city of Korça (just 1.2 km from the village of Barçi) was already invaded by the Ottomans in the 15th century CE and thus the individuals sampled there could very well be descendants of the Ottomans.

A set of samples from Çapalıbağ at Muğla (1300-1650CE; average of C14 dates of 1480CE) are also likely to be of Turkic ancestry as they postdate the establishment of the Seljuq

dynasty in Anatolia and also have substantial East Asian-related ancestry which could have been mediated via Central Asian ancestors. We dated the admixture timing of this population using DATES(*52*) using the Anatolian Byzantine population as one source and a diverse set of samples from Central Asia between 500-1500 years ago as the second source. The obtained date is 12.2±1.4 generations or, assuming a generation length of 28 years,(*53*) a 95% confidence interval of ~267-418 years (mean 342 years) prior to their time which would be consistent with either the admixture taking place prior to or after the arrival of Seljuqs in Anatolia. We also estimated the admixture date for present-day Turkish people from Anatolia genotyped on the Human Origins array(*35*) using the same sources, finding it to be 30.6±1.9 generations or ~755-958 years (mean 856 years). The estimated admixture time using the Çapalıbağ and present-day Turkish data coincides with the early centuries of the 2nd millennium CE, roughly the period in which the Seljuqs and Ottomans gained control of Anatolia from the Romans (Byzantines) before the final capture of the Imperial capital city of Constantinople in 1453CE.

The problem of the sources and mode of arrival of Turkic speakers in Anatolia is complex and is beyond the scope of this paper, but we hope that the data presented here would be useful to future studies of the topic, as they establish the "Roman-Byzantine" baseline population across large parts of Anatolia on which the Turkic population influence could be studied in the future, as well as the first known sample ancient DNA population of plausibly Turkic descendants from Anatolia.

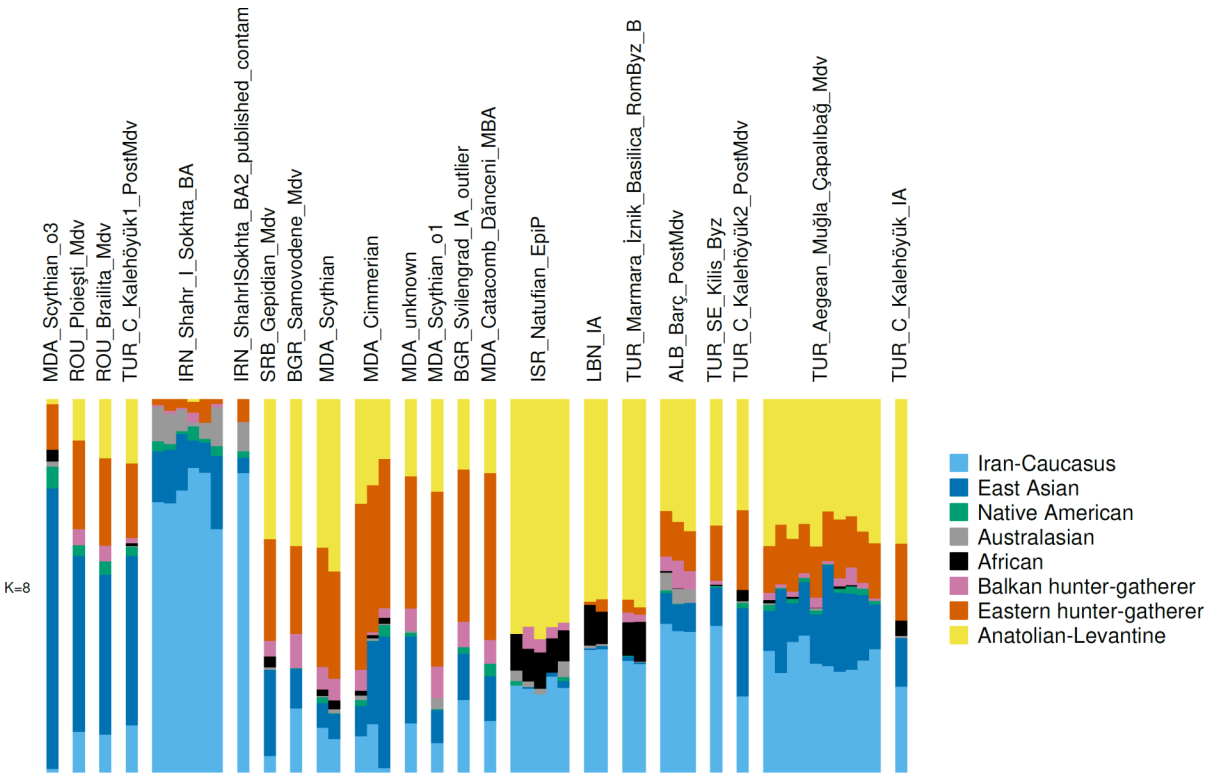

Fig. S 4 Non-West Eurasian outliers

Mycenaean-like ancestry in southeastern Europe and beyond

A recent study identified the presence of individuals resembling the Bronze Age Mycenaean population of Greece(*4*) in a time transect of the Spanish site of Empúries (Greek

Ἐμπόριον).(*21*) The approach adopted there was to identify the outliers in the site and find that they were most similar to the Mycenaean population.

Here we try to perform the converse operation: beginning with the Mycenaean Greeks (to the sampling of which our study adds many new individuals), can we identify other individuals in our total dataset that could be drawn from the same population. Such an operation would hopefully identify the known individuals from Empúries but might also disclose other such individuals in either the new data of our paper or the literature at large.

Our approach to data mine the dataset is as follows. We estimate the Mahalanobis distance of each sample to the total Mycenaean population and order samples according to this distance using a p-value cutoff of 0.01 to identify samples that are not significantly different than the Mycenaeans. To ensure that we identify samples that are genuinely within the Mycenaean genetic variation (to the limits of our ability), we perform this operation in three types of data: ADMIXTURE coefficients, F4admix coefficients (excluding the SRB_Iron_Gates_HG component for numerical stability as this is zero in most individuals), and the first 10 principal components of the West Eurasian PCA (*1*).

In Table S 1 we list the individuals that are indistinguishable from Mycenaeans according to our procedure in all three tests. This does indeed identify two individuals from Empúries (I8215 and I8208) as highly similar to the Mycenaean population. The strong similarity of these two Classical and Hellenistic individuals (4th-3rd century BCE) to the Mycenaeans of a 1,000 years earlier has interesting implications beyond their local Iberian setting and underscores the importance of "Big Picture" studies to produce a framework through which the analysis of local populations can be better interpreted:

The western Mediterranean Greek colonists in this site in Spain were derived from 6th c. BCE Massaliotes (Ancient Μασσαλία, modern Marseilles in France) who themselves were derived from Phocaeans (Ancient Φώκαια, modern Foça in Turkey) who themselves were colonists from Phokis (Φωκίς) in mainland Greece with Ionian kings who traced descent from Codrus (and thus from Attica).[1] Whatever the origin of the specific individuals unearthed at Empúries, their genetic similarity to the Mycenaean population suggests that no major admixture had occurred in their ancestry from the Bronze Age to their own time, e.g., in either Asia Minor (during the founding of Phocaea) or western Europe, which would have introduced ancestry more prevalent in either region (e.g., CHG or WHG) compared to mainland Greece. We do not have all the links in this long chain of transmission of the Aegean ancestry into the western Mediterrenan, yet we do have samples of Myceanean age from the site of Kastrouli near Delphi in Phokis, two Archaic sample from Phokis (I17962; 773-544 calBCE, and I17959; 800-500 BCE) closer to the time of the foundation of Phocaea, and Mycenaean samples from Attica and can thus confirm that the population of the putative ancestors of the Western Mediterranean Greeks were indeed similar to that of the Mycenaeans in general on the basis of I17962 which appears to by Mycenaean-like according to our procedure (I17959 is ranked #76 and is not listed in the Table, but we do not ascribe any importance to this as this is a lower coverage sample with only ~15k SNPs covered).

From Greece itself there is another post-Mycenaean (Proto-Geometric/Early Iron Age) individual (I19368) from the vicinity of the Palace of Nestor at Pylos which is also confirmed by our procedure to be Mycenaean-like and thus similar to the people that lived in Greece a few centuries earlier across the LBA to Iron Age transition.

[1] Paus. 7.3.10: http://www.perseus.tufts.edu/hopper/text?doc=urn:cts:greekLit:tlg0525.tlg001.perseus-eng1:7.3.10

Another sample which resembles Mycenaeans genetically is ASH068 an Iron Age "Philistine" from the Levant, also identified as resembling the Late Bronze Age population of southern Greece in the original publication.(*22*)

Two other samples from the literature were identified:

SZ19 is a Langobard-era sample from Szólád, Hungary from the 5th-6th c. CE. SZ19 was a young female of 17-25 years old who was also a genetic outlier in the group of individuals buried there, had a distinct burial type, and also had a "stylistically distinct (possibly Roman)"(*54*) artifact associated with her burial. Quite possibly she was related to the population of the Aegean and the southern Balkans given the similarity to Mycenaeans detected here.

I20257 is an ancient adolescent female from Değirmendere in Muğla from the Aegean region of Turkey (750-480 BCE). Her similarity to the Mycenaean population is not surprising given the proximity to Greece and her time postdating the colonization of the coast of Anatolia. Two other samples from the same site are more distant (I20229 and I20233). Thus only 3 of 10 samples from this site are similar to Mycenaeans. We cannot speak of a general similarity here, but rather that the "Carian" population at Değirmendere included Mycenaean-like individuals while being generally distinct. Thus, the previously plausible theory that culturally Greek people in the classical period and earlier did not mix with locals—suggested by the patterns at Empúries—is not supported by the data.

I5737, a Middle Bronze Age sample from Yassıtepe (Izmir, Ancient Σμύρνα / Smyrna) is also identified, predating the Mycenaean samples (2033-1920 calBCE). Its EHG ancestry is 2.9±2.6% so we cannot be certain that it was present here as in most Mycenaean samples, but its overall genetic makeup appears to be similar. This individual also had Y-chromosome I-P58 linking him to southeastern Europe. We cannot speak of the population in general here, but this sample provides the earliest direct evidence of human migration from the Balkans to Anatolia, a pattern that recurs more than a millennium later at Değirmendere and provides evidence of a long history of genetic interchange across the Aegean. Two Roman/Byzantine samples from the Basilica at Nicaea are the remaining Mycenaean-like samples from Anatolia (I8366 and I8368).

Overall, however, our procedure only identified a very small number of individuals from Anatolia as being genetically similar to Mycenaeans, which is notable given the colonization of Anatolia by Ionian Greeks and the later incorporation of it to the Hellenistic Kingdoms and Roman Empire which used *koine* Greek as its language in the east. Possible explanations for this are either that our sampling bias—that our dataset has few samples derived from contexts specific to ancient colonists—or that the colonists of Anatolia intermarried with the local population as suggested in ancient times by Herodotus for Ionian colonists from Athens who intermarried with local Carian women (again, different from the pattern seen at Empúries where many in the culturally Greek population retained their genetic affinity to Greece despite a long history of serial colonization.[2] The same could be true for individuals sampled from Samsun (Ancient Ἀμισός / Amisos) and Bodrum (Ancient Ἁλικαρνασσός / Halikarnassos) which were certainly places of ancient settlement and were the colonists may have intermarried with locals which would have modified their ancestry in a more "eastern" direction.

To the west of Greece, 1 sample from Italy, a Punic sample from Sardinia (MSR002) is identified as Mycenaean-like.(*55*) We note that the samples from Italy do not include Sicily and Southern Italy at the time or postdating Greek colonization, but they do include a large set of samples from Imperial Rome which we infer to be mostly of Anatolian rather than Aegean or southeastern European origin.

[2] http://www.perseus.tufts.edu/hopper/text?doc=Hdt.+1.146&fromdoc=Perseus%3Atext%3A1999.01.0126

Many samples from southeastern Europe north of Greece are identified as being genetically similar to the Mycenaean population of southern Greece.

From North Macedonia, a sample (I7233; 897-811 calBCE) is quite early in time. We cannot speak of a general Mycenaean-like population here as the remaining samples from the 1st millennium BCE do not bear this close resemblance to the Mycenaean population.

By far, the greatest number of Mycenaean-like individuals in our dataset outside Greece itself is found in neighboring Bulgaria where 10 such samples (from several sites) are identified. A Neolithic outlier individual from Krepost(*7*) is the earliest. This individual has no EHG ancestry according to our estimation, but is a mixture of mainly Anatolian Neolithic and CHG-related ancestry. Thus, it may somewhat resemble Mycenaeans, but it would be difficult to speak of continuity since its 6th millennium BCE date on its basis, especially as this pattern is not supported by other Neolithic/Chalcolithic era samples from Bulgaria or Greece, some of which post-date the Krepost individual.

More convincing are several 1st millennium BCE individuals from Rozovo (I19500), Diamandievo (I19481), Dzhulyunitsa (I5769), and Kapitan Andreevo (about half of the samples here). As these sites are inland, they should not be attributed to maritime contacts and the foundation of colonies in the Thracian coast by Greek settlers, but may better suggest a similarity of population in the southern Balkans with the Aegean.

Future studies of intermediate regions between southern Greece, North Macedonia, and Bulgaria will be important in further mapping the extent of the Mycenaean-like population and its relationships to those further north in the Balkans.

		ADMIXTURE		PCA		F4admix		
Population	Individual	Distance	P-value	Distance	P-value	Distance	P-value	Sum of Ranks
GRC_Mycenaean_Palace_of_Nestor_BA	I19366	0.848	0.997	0.006	0.940	0.396	0.941	5
GRC_Mycenaean_Kastrouli_BA	I13579	3.445	0.841	0.186	0.666	0.752	0.861	15
GRC_Mycenaean_Attica_BA	I15571	2.932	0.891	0.531	0.466	1.072	0.784	32
ISR_Ashkelon_IA_A	ASH068	4.942	0.667	0.635	0.426	0.355	0.949	42
GRC_Mycenaean_Palace_of_Nestor_BA	I13514	4.392	0.734	0.348	0.555	2.768	0.429	49
ESP_Empuries_Anc	I8208	4.224	0.754	0.441	0.507	2.804	0.423	53
GRC_Mycenaean_Palace_of_Nestor_BA	I13516	4.051	0.774	0.103	0.748	3.410	0.333	53
GRC_Mycenaean_Attica_BA	I15582	5.423	0.608	0.895	0.344	0.282	0.963	55
GRC_Kastrouli_Anc	I17962	1.417	0.985	0.834	0.361	2.631	0.452	55
GRC_Mycenaean_Palace_of_Nestor_BA	I19364	5.937	0.547	0.159	0.691	3.015	0.389	57
GRC_Mycenaean_Palace_of_Nestor_BA	I13518	4.617	0.707	0.829	0.362	1.918	0.590	59
GRC_Mycenaean_Kastrouli_BA	I13578	6.645	0.467	0.622	0.430	1.543	0.672	62
GRC_Mycenaean_Galatas_BA	I9010	9.391	0.226	0.637	0.425	0.973	0.808	66
ESP_Empuries_Anc	I8215	5.625	0.584	0.985	0.321	1.570	0.666	75
GRC_Mycenaean_Galatas_BA	I9041	3.696	0.814	1.226	0.268	2.104	0.551	75
GRC_Mycenaean_Attica_BA	I16709	3.808	0.802	0.345	0.557	4.591	0.204	81
BGR_KapitanAndreevo_IA	I20186	10.901	0.143	0.971	0.324	0.907	0.824	90
GRC_Mycenaean_Salamis_BA	I9006	7.835	0.347	1.329	0.249	3.074	0.380	119
GRC_Palace_of_Nestor_EIA	I19368	7.415	0.387	1.799	0.180	2.661	0.447	120
BGR_KapitanAndreevo_IA	I19493	2.735	0.908	2.463	0.117	3.150	0.369	129
BGR_Krepost_N	I0679_d	12.379	0.089	1.125	0.289	3.699	0.296	145
GRC_Mycenaean_Kastrouli_BA	I13428	8.505	0.290	1.591	0.207	3.796	0.284	148
GRC_Mycenaean_Palace_of_Nestor_BA	I13519_d	7.196	0.409	0.845	0.358	5.937	0.115	152
GRC_Mycenaean_Kastrouli_BA	I13577	3.827	0.800	3.691	0.055	1.986	0.575	160
GRC_Mycenaean_Palace_of_Nestor_BA	I13517_d	5.207	0.635	2.658	0.103	3.490	0.322	161
BGR_Dzhulyunitsa_IA	I5769	4.878	0.675	3.862	0.049	1.526	0.676	164
BGR_KapitanAndreevo_IA	I20180	6.424	0.491	3.136	0.077	3.244	0.355	177
ITA_Sardinia_IA_Punic_2	MSR002	16.201	0.023	1.089	0.297	5.400	0.145	187
GRC_Mycenaean_Attica_BA	I14872	4.725	0.694	0.520	0.471	8.186	0.042	188
TUR_Aegean_Muğla_Değirmendere_Anc	I20257	9.404	0.225	3.957	0.047	2.892	0.409	206
TUR_Aegean_Izmir_Yassitepe_MBA	I5737	9.794	0.201	4.306	0.038	2.269	0.519	210
GRC_Mycenaean_Kastrouli_BA	I13433	3.152	0.871	4.329	0.037	4.176	0.243	215
BGR_KapitanAndreevo_IA	I20184	16.900	0.018	3.992	0.046	2.546	0.467	236
GRC_Peloponnese_N	I3920	12.020	0.100	4.499	0.034	3.101	0.376	243
GRC_Mycenaean_Palace_of_Nestor_BA	I13510	11.117	0.134	4.398	0.036	3.885	0.274	252
BGR_Anc	I19500	10.105	0.183	4.600	0.032	4.447	0.217	260
MKD_Anc	I7233	10.564	0.159	4.617	0.032	5.045	0.169	277
HUN_Langobard_Mdv	SZ19	6.631	0.468	4.663	0.031	6.079	0.108	287
TUR_Marmara_İznik_Basilica_RomByz_A	I8366	13.929	0.052	4.913	0.027	5.831	0.120	315
TUR_Aegean_Muğla_Değirmendere_Anc	I20229	8.040	0.329	5.984	0.014	5.733	0.125	315
BGR_KapitanAndreevo_IA	I20181	11.521	0.117	6.288	0.012	4.441	0.218	316
TUR_Aegean_Muğla_Değirmendere_Anc	I20233	15.155	0.034	6.133	0.013	3.927	0.269	318
GRC_Mycenaean_Kastrouli_BA	I13580	18.145	0.011	2.285	0.131	8.695	0.034	333
BGR_Diamandievo_IA	I19481	3.781	0.805	5.183	0.023	7.924	0.048	334
BGR_RomByz	I18792	5.405	0.611	6.394	0.011	8.754	0.033	405

Table S 1 Mycenaean-like individuals. We show the Mahalanobis distance and associated p-value for individuals relative to the Mycenaean population. Samples are ordered by the sum of their distance ranks for the three tests (from most to least Mycenaean-like) (e.g., the first individual I19366 is closest/ranked 1st to the Mycenaean centroid on the ADMIXTURE and PCA measures and 3rd on the F4admix one and thus has sum of ranks=1+1+3=5).

S2: Runs of homozygosity in the Southern Arc

We applied hapROH a new Hidden Markov model method for detecting runs of homozygosity (ROH) in low coverage ancient individuals (*20*). Long ROH in an individual indicate that their parents were genetically closely related; this has been generally uncommon in much of the world since the onset of the Neolithic (*20*) and so we identify here interesting cases where this has occurred. We applied hapROH to all Southern Arc individuals with at least 400,000 autosomal SNPs covered and list in Table S 2 Long ROH segments in Southern Arc individuals.the individuals that had at least one ROH segment longer than 20cM and at least 50cM of such segments in total, together with the total length of such segments >20cM (=0.2M). We show some histograms of ROH distribution for a subset of the individuals of Table S 2 in Fig. S 5. We also summarize the distribution of long ROH segments in the highly inbred individuals in Fig. S 5 which indicates that two of these are probably the result of a 1st-degree relative pairing, two of an uncle/niece or aunt/nephew pairing, and the remaining of 1st cousin marriage.

ID	Sum of ROH >0.2M	Population
I14812	0.582	ARM_Karnut_KuraAraxes_EBA
I18274	1.408	ARM_Lhashen_LBA
I18467	1.449	ARM_Noratus_LBA
I17981	0.686	BGR_Dzhulyunitsa_N
I2521	2.670	BGR_Dzhulyunitsa_N
I13518	1.394	GRC_Mycenaean_Palace_of_Nestor_BA
I2937_all	0.653	GRC_Peloponnese_N
I5074	0.983	HRV_BA
I18727	1.537	HRV_Bezdanjača_BA
I18717	3.244	HRV_Bezdanjača_BA_brother.I18078.father.I18071
I18746	0.602	HRV_Cetina_BA
I1178	3.909	ISR_ChL
SFI-50.SG	1.205	LBN_IA
SFI-39.SG	2.257	LBN_IA
I15616	0.522	ROU_Bodrogkeresztur_ChL
I7129	0.694	ROU_Bodrogkeresztur_ChL
I4916	0.921	SRB_Iron_Gates_HG
I1131	1.450	SRB_N
I20224	1.892	TUR_Aegean_Muğla_Değirmendere_Anc
CBT016	0.525	TUR_C_ÇamlıbelTarlası_ChL
ART005	0.612	TUR_E_Arslantepe_ChL
ALA013	1.398	TUR_Hatay_Alalakh_MLBA
I14845	0.598	TUR_Marmara_Apollonia_Rom
I14844	4.528	TUR_Marmara_İznik_Y.kapı_PostMdv
I14635	2.306	TUR_SE_Batman_Anc
I4481	0.889	TUR_SE_Şırnak_BA
I4478	1.295	TUR_SE_Şırnak_BA_sibling.I4481

Table S 2 Long ROH segments in Southern Arc individuals.

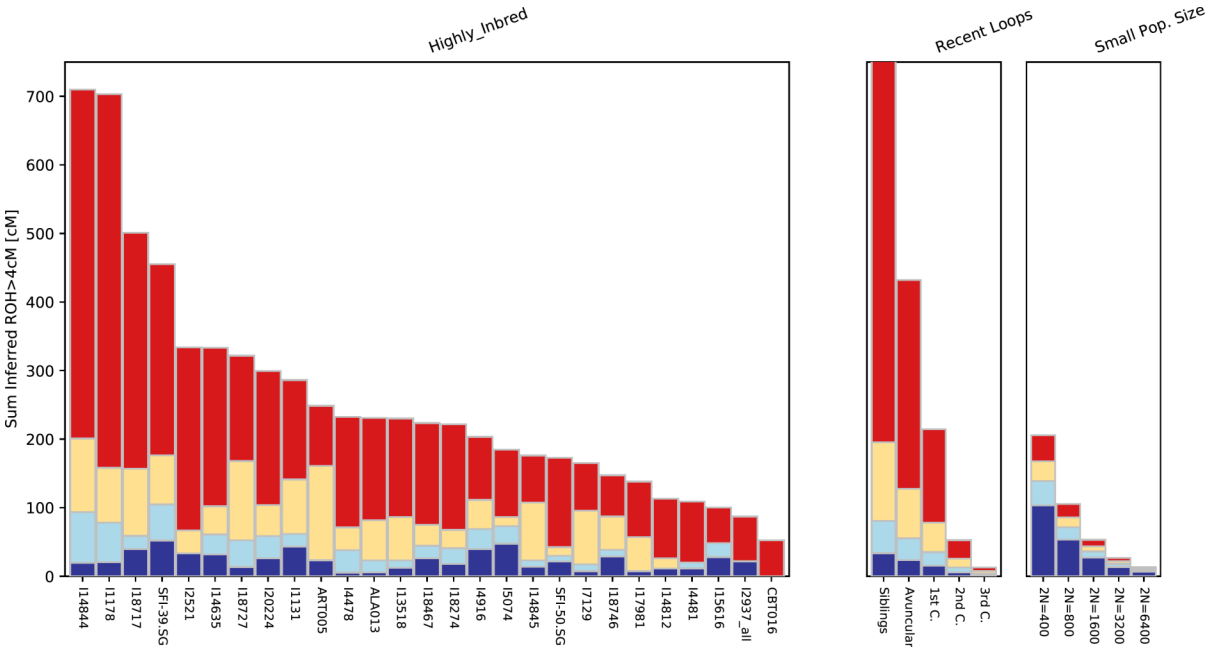

Fig. S 5 Summary of ROH distribution over all highly inbred individuals. The sum of ROH segments for different length classes (blue: 4-8cM; cyan: 8-12cM; yellow: 12-20cM; red: 20-300cM)

S3: Pigmentation variation of the Southern Arc in relation to West Eurasians

Ancient art and literature from the classical world abound with depictions and references to the phenotypes of people from different parts of the world. Most often these involved stereotypical descriptions of "exotic" populations such as a few indicative passages below:

"The Gauls are tall of body, with rippling muscles, and white of skin, and their hair is blond, and not only naturally so, but they also make it their practice by artificial means to increase the distinguishing colour which nature has given it." (Diodorus Siculus, The Library of History, Bk. IV, 28; 1st c. BCE)

"For my own part, I agree with those who think that the tribes of Germany are free from all taint of inter-marriages with foreign nations, and that they appear as a distinct, unmixed race, like none but themselves. Hence, too, the same physical peculiarities throughout so vast a population. All have fierce blue eyes, red hair, huge frames, fit only for a sudden exertion. They are less able to bear laborious work. Heat and thirst they cannot in the least endure; to cold and hunger their climate and their soil inure them." (Tacitus, Germania, 4; 1st c. CE)

"But mortals suppose that gods are born, wear their own clothes and have a voice and body. (frag. 14) Ethiopians say that their gods are snub-nosed and black; Thracians that theirs are blue-eyed and red-haired." (Xenophanes of Colophon, frag. 14, 16; 6th c. BCE)

"The Hellenes breakfasted and then started forward on their march, having first delivered the stronghold to their allies among the Mossynoecians. … The whole community, male and female alike, were fair-complexioned and white-skinned." (Xenophon, Anabasis, Bk. 5, IV; 4th c. BCE)

The Greek physician and philosopher Galen (De Temperamentis, 2.5; 2nd c. CE) contrasts the thin, straight, light "red" hair of inhabitants of cold and damp regions (Illyrians, Germans, Dalmatians, Sauromatians, and "all Scythians") with the thick, curly, black hair of warm and dry ones (Egyptians, Arabs, and Indians), and with the moderately dark hair of those of intermediate regions ("μελαίνας μετρίως καὶ παχείας συμμέτρως καὶ οὔτ' ἀκριβῶς οὔλας οὔτ' ἀκριβῶς εὐθείας." / melainas metriôs kai pakheias summetrôs kai out' akribôs oylas out' akribôs eutheias) This "climate theory" of light pigmentation was echoed by Vitruvius who suggested (On Architecture, 6.1.3; 1st c. BCE) that "the people of the north are so large in stature, so light in complexion, and have straight red hair, blue eyes, and are full of blood, for they are thus formed by the abundance of the moisture, and the coldness of their country."

While these descriptions correspond to a degree to what is known about the modern variation of pigmentation traits, they do not inform about the distribution of the different phenotypes in the ancient world, not do they inform about the statistical distribution of the different phenotypes in the different populations. For example, the description of the Gauls suggests that hair color may be a darker shade of blond that could be artificially lightened; the description of the Germans that they are all blue-eyed, a categorical statement not applicable to any known population: surely there were many in Germania compared to the Mediterranean world of Tacitus, but what fraction of the population did they make? The same question might apply to Xenophanes' Thracians in southeastern Europe who surely did not all have blue eyes, but probably more than the people of Colophon in northwestern Anatolia. Finally, in Xenophon's account of the escape from Mesopotamia to the Black Sea of his group of Greek mercenaries, he must have encountered a group of people with light skin pigmentation, an observation that might indirectly contrast to the numerous other tribes encountered during the long trek.

Ancient art which was often polychromatic(*56*) furnishes independent evidence of ancient phenotypes, but that too is limited by degradation over time, questions about the realistic vs. idealistic portrayal of human figures, and the choice of subjects depicted. In only rare cases, such as the desiccated corpses of the Tarim basin(*57*) is there direct evidence of the phenotypes exhibited, although in such cases too post-mortem chemical processes in the soil must be considered.

The question of what ancient populations looked came to be important during the 19th and 20th centuries with the rise of biological anthropology it became possible to study modern human phenotypic variation quantitatively and to infer how human populations came to be. An example of the association of phenotype with ideology was the emergence and promulgation of the "Aryan myth"(*42*). This idea, promoted by writers like Arthur de Gobineau(*58*) and Vacher de Lapouge(*59*) in France, but spreading to much of Europe, espoused the ideal of the "blond Aryan" master race, a theme that was later taken up by early 20th century writers such as Madison Grant(*60*) and Houston Chamberlain(*61*) and inspired racist ideologies and in some cases genocide in both the United States(*62, 63*) and Germany(*64-66*).

These ideas often conflated phenotypic features (such as pigmentation or skull shape) with ancestry, nationality, and with psychological and behavioral traits and leveraged the supposed history of the past to drive social policy in the present.

The association continues to be sometimes be made(*67, 68*) marshalling the evidence of biological anthropology, ancient art, and literature, that the Proto-Indo-Europeans had traits of depigmentation of the hair, eyes, and skin that largely correspond to the "Aryan myth" of past generations. For a useful summary of the persistence of this myth from its beginnings to the present see (*69*).

In this section we use the HIrisPlex-S system(*41, 70, 71*) to infer the pigmentation phenotypes of 4,118 ancient West Eurasians. Of these, 3,761 had data on at least one SNP of the system and could thus be submitted for phenotype prediction. Data was sufficient to make a prediction for 1,935 individuals in total. We limit our discussion to a subset of 1,899 individuals for which predictions were made for all phenotypes.

As in a previous publication(*4*) we simulated genotypes (10 random trials per individual) given genotype likelihoods at each SNP and a prior on the overall allele frequency of each SNP and submitted these to the online HIrisPlex-S website (https://hirisplex.erasmusmc.nl/). We averaged the results for the 10 trials and make phenotype prediction for the four categories based on these averages (HairSimple: "light" or "dark", HairDetailed: "red", "blond", "brown", "black", Eye: "blue", "intermediate", "brown", and Skin color: "very pale", "pale", "intermediate", "dark", "dark-to-black").

We enter three notes of caution. First, phenotypic prediction is not entirely accurate even for modern individuals with perfect genotype information and is less likely to be so in ancient ones. Second, we cannot exclude the possibility that pigmentation in ancient individuals may have been affected by loci not included in the HIrisPlex-S system. Third, the individual predictions of pigmentation are likely to be subject to noise, and so in our discussion we focus on general patterns observed among many individuals. These should be accurate to a degree for inferring the relative appearance of different groups using the best tool we currently possess and the available mostly low-coverage data. Thus, our results are provisional given these limitations, but show, nonetheless, some interesting patterns that we discuss below.

Our first observation (Table S 3) is that the modal phenotype of West Eurasians is one of dark brown hair, brown eyes, and intermediate skin, accounting for roughly ~1/3 of samples both

in the Southern Arc and outside it. The next two most frequent phenotypes have black instead of brown hair and either intermediate or dark skin. A wide variety of phenotypes are found both within the Southern Arc and outside it, although several rare depigmented phenotypes at the bottom of Table S 3 are found only outside the Southern Arc; this, however, should be considered with the knowledge of the larger sample size of non-Southern Arc individuals. The latter include individuals from Europe (outside the countries included in the Southern Arc), and the steppe-to-central South Asia.

Composite Phenotype	Southern Arc (n=705)	Non-Southern Arc (n=1194)	Southern Arc (%)	Non-Southern Arc (%)	
DarkHair_BrownHair_BrownEye_IntermediateSkin	255	379	36.2%	31.7%	
DarkHair_BlackHair_BrownEye_DarkSkin	108	138	15.3%	11.6%	
DarkHair_BlackHair_BrownEye_IntermediateSkin	93	150	13.2%	12.6%	
LightHair_BrownHair_BrownEye_IntermediateSkin	82	123	11.6%	10.3%	
DarkHair_BrownHair_BrownEye_DarkSkin	56	89	7.9%	7.5%	
DarkHair_BlackHair_BrownEye_DarkToBlackSkin	43	62	6.1%	5.2%	
LightHair_BrownHair_BlueEye_IntermediateSkin	18	59	2.6%	4.9%	
DarkHair_BrownHair_BlueEye_IntermediateSkin	10	25	1.4%	2.1%	
LightHair_BlondHair_BlueEye_IntermediateSkin	10	56	1.4%	4.7%	
DarkHair_BrownHair_BrownEye_DarkToBlackSkin	5	11	0.7%	0.9%	
DarkHair_BrownHair_BrownEye_PaleSkin	5	5	0.7%	0.4%	
LightHair_BlondHair_BrownEye_IntermediateSkin	5	17	0.7%	1.4%	
LightHair_BrownHair_BrownEye_DarkToBlackSkin	5	1	0.7%	0.1%	
LightHair_BrownHair_BrownEye_PaleSkin	4	14	0.6%	1.2%	
LightHair_BrownHair_BlueEye_PaleSkin	2	13	0.3%	1.1%	
DarkHair_BrownHair_BlueEye_DarkSkin	1	4	0.1%	0.3%	
LightHair_BlackHair_BrownEye_DarkToBlackSkin	1	0	0.1%	0.0%	
LightHair_BlondHair_BlueEye_PaleSkin	1	30	0.1%	2.5%	
LightHair_RedHair_BrownEye_IntermediateSkin	1	1	0.1%	0.1%	
DarkHair_BlackHair_BlueEye_IntermediateSkin	0	1	0.0%	0.1%	
DarkHair_BlackHair_BrownEye_PaleSkin	0	1	0.0%	0.1%	
DarkHair_BrownHair_BlueEye_DarkToBlackSkin	0	2	0.0%	0.2%	
DarkHair_BrownHair_BlueEye_PaleSkin	0	2	0.0%	0.2%	
LightHair_BlackHair_BrownEye_IntermediateSkin	0	1	0.0%	0.1%	
LightHair_BlondHair_BrownEye_PaleSkin	0	4	0.0%	0.3%	
LightHair_BrownHair_BlueEye_DarkSkin	0	1	0.0%	0.1%	
LightHair_BrownHair_BrownEye_DarkSkin	0	2	0.0%	0.2%	
LightHair_RedHair_BlueEye_IntermediateSkin	0	1	0.0%	0.1%	
LightHair_RedHair_BlueEye_PaleSkin	0	2	0.0%	0.2%	
Simple Phenotype					**Fisher's Exact Test**
DarkHair	576	869	81.7%	72.8%	1.1E-01
LightHair	129	325	18.3%	27.2%	**4.7E-04**
BlackHair	245	353	34.8%	29.6%	1.0E-01
BlondHair	16	107	2.3%	9.0%	**9.1E-09**
BrownHair	443	730	62.8%	61.1%	7.3E-01
RedHair	1	4	0.1%	0.3%	6.6E-01
BlueEye	42	196	6.0%	16.4%	**5.9E-10**
BrownEye	663	998	94.0%	83.6%	9.1E-02
DarkSkin	165	234	23.4%	19.6%	1.3E-01
DarkToBlackSkin	54	76	7.7%	6.4%	3.5E-01
IntermediateSkin	474	813	67.2%	68.1%	8.8E-01
PaleSkin	12	71	1.7%	5.9%	**1.1E-05**

Table S 3 Frequency of phenotypes in Southern Arc and non-Southern Arc populations

By examining simple phenotypes (Table S 3) we see that Southern Arc individuals have a lower frequency of light hair, blond hair, blue eyes, and pale skin compared to non-Southern Arc ones, a finding that is in agreement with the ancient sources that commented on the appearance of Celts, Germans, and Scytho-Sarmatians from Europe and Central Asia.

Note that these sources are from the 1st millennia BCE/CE, a narrower time range than that of our samples which extend millennia into the past when there were no written sources. In Table S 4 we tabulate phenotypic information for the populations of the Caucasian and Anatolian-Aegean bridge (of Fig. 2) which are from the Chalcolithic and Bronze Age. These show that the modal phenotype had dark brown hair, brown eyes, and intermediate skin pigmentation in most

populations. The Beaker group (with a large sample size) stands out with its higher frequency of blue eyes and blond hair; this group's territory coincided largely with that of the later historical Celtic and (partly) Germanic groups of Europe. But none of the individuals from the Early Bronze Age Yamnaya cluster exhibited these phenotypes, suggesting a turnover of phenotypes before the time of the written sources.

		Caucasian Bridge																								
		Counts												**Percentages**												
	Sample size	DarkHair	LightHair	BlackHair	**BlondHair**	BrownHair	**RedHair**	**BlueEye**	BrownEye	DarkSkin	DarkToBlackSkin	IntermediateSkin	**PaleSkin**	DarkHair	LightHair	BlackHair	**BlondHair**	BrownHair	**RedHair**	**BlueEye**	BrownEye	DarkSkin	DarkToBlackSkin	IntermediateSkin	**PaleSkin**
Anatolia	81	64	17	28	1	52	0	5	76	15	11	55	0	79.0%	21.0%	34.6%	1.2%	64.2%	0.0%	6.2%	93.8%	18.5%	13.6%	67.9%	0.0%
ARM_and_AZE	104	95	9	40	1	63	0	0	104	28	8	68	0	91.3%	8.7%	38.5%	1.0%	60.6%	0.0%	0.0%	100.0%	26.9%	7.7%	65.4%	0.0%
ARM_and_AZE_N	2	2	0	1	0	1	0	0	2	1	0	1	0	100.0%	0.0%	50.0%	0.0%	50.0%	0.0%	0.0%	100.0%	50.0%	0.0%	50.0%	0.0%
IRN	40	38	2	25	0	15	0	1	39	17	7	16	0	95.0%	5.0%	62.5%	0.0%	37.5%	0.0%	2.5%	97.5%	42.5%	17.5%	40.0%	0.0%
Levant	44	40	4	28	1	15	0	3	41	17	6	21	0	90.9%	9.1%	63.6%	2.3%	34.1%	0.0%	6.8%	93.2%	38.6%	13.6%	47.7%	0.0%
RUS_Eneol_Mountains	1	1	0	0	0	1	0	0	1	1	0	0	0	100.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	100.0%	0.0%	0.0%	0.0%
RUS_Eneol_Piedmont	2	2	0	1	0	1	0	0	2	1	0	1	0	100.0%	0.0%	50.0%	0.0%	50.0%	0.0%	0.0%	100.0%	50.0%	0.0%	50.0%	0.0%
RUS_MaykopCluster	5	4	1	1	0	4	0	1	4	1	0	4	0	80.0%	20.0%	20.0%	0.0%	80.0%	0.0%	20.0%	80.0%	20.0%	0.0%	80.0%	0.0%
RUS_Steppe_Maykop	4	4	0	3	0	1	0	0	4	2	1	1	0	100.0%	0.0%	75.0%	0.0%	25.0%	0.0%	0.0%	100.0%	50.0%	25.0%	25.0%	0.0%
RUS_YamnayaCluster	19	19	0	8	0	11	0	0	19	4	2	13	0	100.0%	0.0%	42.1%	0.0%	57.9%	0.0%	0.0%	100.0%	21.1%	10.5%	68.4%	0.0%
SE_Europe	114	89	25	23	5	85	1	10	104	19	3	90	2	78.1%	21.9%	20.2%	4.4%	74.6%	0.9%	8.8%	91.2%	16.7%	2.6%	78.9%	1.8%
		Anatolian-Aegean Bridge																								
		Counts												**Percentages**												
	Sample size	DarkHair	LightHair	BlackHair	**BlondHair**	BrownHair	**RedHair**	**BlueEye**	BrownEye	DarkSkin	DarkToBlackSkin	IntermediateSkin	**PaleSkin**	DarkHair	LightHair	BlackHair	**BlondHair**	BrownHair	**RedHair**	**BlueEye**	BrownEye	DarkSkin	DarkToBlackSkin	IntermediateSkin	**PaleSkin**
Beaker	96	61	35	12	12	71	1	21	75	13	2	75	6	63.5%	36.5%	12.5%	12.5%	74.0%	1.0%	21.9%	78.1%	13.5%	2.1%	78.1%	6.3%
Corded_Ware	15	12	3	7	0	8	0	2	13	3	0	12	0	80.0%	20.0%	46.7%	0.0%	53.3%	0.0%	13.3%	86.7%	20.0%	0.0%	80.0%	0.0%
GRC	12	9	3	3	1	8	0	0	12	2	0	10	0	75.0%	25.0%	25.0%	8.3%	66.7%	0.0%	0.0%	100.0%	16.7%	0.0%	83.3%	0.0%
MDA	5	5	0	0	0	5	0	0	5	1	0	4	0	100.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	20.0%	0.0%	80.0%	0.0%
RUS_YamnayaCluster	19	19	0	8	0	11	0	0	19	4	2	13	0	100.0%	0.0%	42.1%	0.0%	57.9%	0.0%	0.0%	100.0%	21.1%	10.5%	68.4%	0.0%
SE_Europe_BA_BIA	50	40	10	6	3	40	1	5	45	7	0	42	1	80.0%	20.0%	12.0%	6.0%	80.0%	2.0%	10.0%	90.0%	14.0%	0.0%	84.0%	2.0%
SE_Europe_ChL	46	34	12	14	1	31	0	5	41	8	3	34	1	73.9%	26.1%	30.4%	2.2%	67.4%	0.0%	10.9%	89.1%	17.4%	6.5%	73.9%	2.2%
SE_Europe_N	23	20	3	7	0	16	0	2	21	6	0	17	0	87.0%	13.0%	30.4%	0.0%	69.6%	0.0%	8.7%	91.3%	26.1%	0.0%	73.9%	0.0%
TUR_Aegean	3	2	1	1	0	2	0	0	3	0	0	3	0	66.7%	33.3%	33.3%	0.0%	66.7%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%
TUR_BlackSea	3	1	2	0	0	3	0	0	3	0	0	3	0	33.3%	66.7%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%
TUR_C	11	8	3	2	0	9	0	1	10	0	1	10	0	72.7%	27.3%	18.2%	0.0%	81.8%	0.0%	9.1%	90.9%	0.0%	9.1%	90.9%	0.0%
TUR_E	18	15	3	8	0	10	0	2	16	5	2	11	0	83.3%	16.7%	44.4%	0.0%	55.6%	0.0%	11.1%	88.9%	27.8%	11.1%	61.1%	0.0%
TUR_Hatay	24	21	3	10	1	13	0	2	22	5	5	14	0	87.5%	12.5%	41.7%	4.2%	54.2%	0.0%	8.3%	91.7%	20.8%	20.8%	58.3%	0.0%
TUR_Marmara	3	3	0	1	0	2	0	0	3	0	0	3	0	100.0%	0.0%	33.3%	0.0%	66.7%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%
TUR_Med	3	3	0	2	0	1	0	0	3	1	0	2	0	100.0%	0.0%	66.7%	0.0%	33.3%	0.0%	0.0%	100.0%	33.3%	0.0%	66.7%	0.0%
TUR_SE	16	11	5	4	0	12	0	0	16	4	3	9	0	68.8%	31.3%	25.0%	0.0%	75.0%	0.0%	0.0%	100.0%	25.0%	18.8%	56.3%	0.0%

Table S 4 Chalcolithic and Bronze Age phenotype distribution along Caucasian and Anatolian-Aegean bridges. Depigmented phenotypesare highlighted in bold.

		Counts												Percentages											
	Sample_Size	DarkHair	LightHair	BlackHair	BlondHair	BrownHair	RedHair	BlueEye	BrownEye	DarkSkin	DarkToBlackSk	IntermediateSki	PaleSkin	DarkHair	LightHair	BlackHair	BlondHair	BrownHair	RedHair	BlueEye	BrownEye	DarkSkin	DarkToBlackSk in	IntermediateSki n	PaleSkin
ALB_PostMdv	5	1	4	0	0	5	0	2	3	0	0	5	0	20.0%	80.0%	0.0%	0.0%	100.0%	0.0%	40.0%	60.0%	0.0%	0.0%	100.0%	0.0%
ARM_Bagheri_Tchala_EIA	8	8	0	3	0	5	0	0	8	3	1	4	0	100.0%	0.0%	37.5%	0.0%	62.5%	0.0%	0.0%	100.0%	37.5%	12.5%	50.0%	0.0%
ARM_Black_Fortress_LBA	6	5	1	4	0	2	0	0	6	2	1	3	0	83.3%	16.7%	66.7%	0.0%	33.3%	0.0%	0.0%	100.0%	33.3%	16.7%	50.0%	0.0%
ARM_Brardzryal_Urartian	5	5	0	2	0	3	0	0	5	1	0	4	0	100.0%	0.0%	40.0%	0.0%	60.0%	0.0%	0.0%	100.0%	20.0%	0.0%	80.0%	0.0%
ARM_Harjis_LateUrartian	6	6	0	1	0	5	0	0	6	3	0	3	0	100.0%	0.0%	16.7%	0.0%	83.3%	0.0%	0.0%	100.0%	50.0%	0.0%	50.0%	0.0%
ARM_Karashamb_LBA	21	21	0	9	0	12	0	0	21	6	1	14	0	100.0%	0.0%	42.9%	0.0%	57.1%	0.0%	0.0%	100.0%	28.6%	4.8%	66.7%	0.0%
ARM_Lhashen_LBA	16	15	1	7	0	9	0	0	16	7	1	8	0	93.8%	6.3%	43.8%	0.0%	56.3%	0.0%	0.0%	100.0%	43.8%	6.3%	50.0%	0.0%
ARM_Nerkin_Getashen_LBA	7	6	1	2	0	5	0	0	7	2	1	4	0	85.7%	14.3%	28.6%	0.0%	71.4%	0.0%	0.0%	100.0%	28.6%	14.3%	57.1%	0.0%
ARM_Noratus_EIA	7	7	0	3	0	4	0	0	7	1	0	6	0	100.0%	0.0%	42.9%	0.0%	57.1%	0.0%	0.0%	100.0%	14.3%	0.0%	85.7%	0.0%
ARM_Noratus_LBA	8	6	2	3	1	4	0	0	8	2	0	6	0	75.0%	25.0%	37.5%	12.5%	50.0%	0.0%	0.0%	100.0%	25.0%	0.0%	75.0%	0.0%
AUT_LBK_EN	7	6	1	4	0	3	0	1	6	2	2	3	0	85.7%	14.3%	57.1%	0.0%	42.9%	0.0%	14.3%	85.7%	28.6%	28.6%	42.9%	0.0%
BGR_ChL	7	6	1	3	0	4	0	0	7	1	1	5	0	85.7%	14.3%	42.9%	0.0%	57.1%	0.0%	0.0%	100.0%	14.3%	14.3%	71.4%	0.0%
BGR_KapitanAndreevo_IA	6	2	4	0	0	6	0	1	5	0	0	5	1	33.3%	66.7%	0.0%	0.0%	100.0%	0.0%	16.7%	83.3%	0.0%	0.0%	83.3%	16.7%
CHE_EBA_2	7	6	1	2	0	5	0	0	7	2	0	5	0	85.7%	14.3%	28.6%	0.0%	71.4%	0.0%	0.0%	100.0%	28.6%	0.0%	71.4%	0.0%
CHE_LN	18	14	4	3	1	14	0	2	16	3	1	13	1	77.8%	22.2%	16.7%	5.6%	77.8%	0.0%	11.1%	88.9%	16.7%	5.6%	72.2%	5.6%
CZE_Bell_Beaker	25	18	7	2	2	21	0	5	20	5	0	18	2	72.0%	28.0%	8.0%	8.0%	84.0%	0.0%	20.0%	80.0%	20.0%	0.0%	72.0%	8.0%
CZE_Corded_Ware	7	5	2	2	0	5	0	2	5	1	0	6	0	71.4%	28.6%	28.6%	0.0%	71.4%	0.0%	28.6%	71.4%	14.3%	0.0%	85.7%	0.0%
CZE_EBA	8	4	4	0	1	7	0	1	7	0	0	6	2	50.0%	50.0%	0.0%	12.5%	87.5%	0.0%	12.5%	87.5%	0.0%	0.0%	75.0%	25.0%
DEU_Bell_Beaker	17	11	6	3	3	11	0	2	15	0	2	14	1	64.7%	35.3%	17.6%	17.6%	64.7%	0.0%	11.8%	88.2%	0.0%	11.8%	82.4%	5.9%
DEU_BellBeaker_Lech	5	4	1	1	1	3	0	1	4	1	0	4	0	80.0%	20.0%	20.0%	20.0%	60.0%	0.0%	20.0%	80.0%	20.0%	0.0%	80.0%	0.0%
DEU_Early_Medieval	27	5	22	1	13	13	0	15	12	0	0	17	10	18.5%	81.5%	3.7%	48.1%	48.1%	0.0%	55.6%	44.4%	0.0%	0.0%	63.0%	37.0%
DEU_Early_Medieval_lc	8	1	7	1	5	2	0	5	3	0	0	3	5	12.5%	87.5%	12.5%	62.5%	25.0%	0.0%	62.5%	37.5%	0.0%	0.0%	37.5%	62.5%
DEU_EBA_Lech	11	4	7	0	1	10	0	5	6	0	0	11	0	36.4%	63.6%	0.0%	9.1%	90.9%	0.0%	45.5%	54.5%	0.0%	0.0%	100.0%	0.0%
DEU_EN_LBK	8	6	2	5	1	2	0	1	7	0	2	6	0	75.0%	25.0%	62.5%	12.5%	25.0%	0.0%	12.5%	87.5%	0.0%	25.0%	75.0%	0.0%
DEU_LBK_EN	5	3	2	3	1	1	0	3	2	1	0	4	0	60.0%	40.0%	60.0%	20.0%	20.0%	0.0%	60.0%	40.0%	20.0%	0.0%	80.0%	0.0%
DEU_SouthernDEU_Singen_EBA	5	5	0	2	0	3	0	0	5	0	0	5	0	100.0%	0.0%	40.0%	0.0%	60.0%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%
ESP_BA	12	10	2	3	0	9	0	0	12	1	0	11	0	83.3%	16.7%	25.0%	0.0%	75.0%	0.0%	0.0%	100.0%	8.3%	0.0%	91.7%	0.0%
ESP_C	17	14	3	7	0	10	0	1	16	6	3	8	0	82.4%	17.6%	41.2%	0.0%	58.8%	0.0%	5.9%	94.1%	35.3%	17.6%	47.1%	0.0%
ESP_EN	5	4	1	2	1	2	0	0	5	1	0	4	0	80.0%	20.0%	40.0%	20.0%	40.0%	0.0%	0.0%	100.0%	20.0%	0.0%	80.0%	0.0%
ESP_MN	8	8	0	5	0	3	0	0	8	3	2	3	0	100.0%	0.0%	62.5%	0.0%	37.5%	0.0%	0.0%	100.0%	37.5%	25.0%	37.5%	0.0%
FRA_MN	37	30	7	12	1	24	0	1	36	7	3	27	0	81.1%	18.9%	32.4%	2.7%	64.9%	0.0%	2.7%	97.3%	18.9%	8.1%	73.0%	0.0%
GBR_England_Bell_Beaker	19	11	8	3	3	13	0	9	10	4	0	14	1	57.9%	42.1%	15.8%	15.8%	68.4%	0.0%	47.4%	52.6%	21.1%	0.0%	73.7%	5.3%
GBR_England_CA_EBA	12	6	6	1	1	10	0	2	10	0	0	11	1	50.0%	50.0%	8.3%	8.3%	83.3%	0.0%	16.7%	83.3%	0.0%	0.0%	91.7%	8.3%
GBR_England_MBA	15	10	5	4	2	9	0	4	11	2	0	10	3	66.7%	33.3%	26.7%	13.3%	60.0%	0.0%	26.7%	73.3%	13.3%	0.0%	66.7%	20.0%
GBR_England_N	5	5	0	0	0	5	0	0	5	1	0	4	0	100.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	20.0%	0.0%	80.0%	0.0%
GBR_England_N_all.SG	5	4	1	1	1	3	0	2	3	0	0	5	0	80.0%	20.0%	20.0%	20.0%	60.0%	0.0%	40.0%	60.0%	0.0%	0.0%	100.0%	0.0%
GBR_England_Saxon	5	0	5	0	2	3	0	2	3	0	0	3	2	0.0%	100.0%	0.0%	40.0%	60.0%	0.0%	40.0%	60.0%	0.0%	0.0%	60.0%	40.0%
GBR_Scotland_CA_EBA	5	3	2	0	1	4	0	2	3	0	0	4	1	60.0%	40.0%	0.0%	20.0%	80.0%	0.0%	40.0%	60.0%	0.0%	0.0%	80.0%	20.0%
GBR_Scotland_N	25	25	0	14	0	11	0	0	25	8	2	15	0	100.0%	0.0%	56.0%	0.0%	44.0%	0.0%	0.0%	100.0%	32.0%	8.0%	60.0%	0.0%
GRC_Minoan_Lasithi_BA	5	3	2	1	1	3	0	0	5	1	0	4	0	60.0%	40.0%	20.0%	20.0%	60.0%	0.0%	0.0%	100.0%	20.0%	0.0%	80.0%	0.0%
GRC_Mycenaean	7	6	1	2	0	5	0	0	7	1	0	6	0	85.7%	14.3%	28.6%	0.0%	71.4%	0.0%	0.0%	100.0%	14.3%	0.0%	85.7%	0.0%
GRC_Peloponnese_N	5	4	1	2	0	3	0	0	5	1	0	4	0	80.0%	20.0%	40.0%	0.0%	60.0%	0.0%	0.0%	100.0%	20.0%	0.0%	80.0%	0.0%
HRV_BA	6	5	1	2	1	3	0	0	6	1	0	5	0	83.3%	16.7%	33.3%	16.7%	50.0%	0.0%	0.0%	100.0%	16.7%	0.0%	83.3%	0.0%
HRV_Bezdanjača_BA	11	9	2	0	0	10	1	0	11	3	0	8	0	81.8%	18.2%	0.0%	0.0%	90.9%	9.1%	0.0%	100.0%	27.3%	0.0%	72.7%	0.0%
HRV_IA	8	4	4	1	1	6	0	2	6	0	0	7	1	50.0%	50.0%	12.5%	12.5%	75.0%	0.0%	25.0%	75.0%	0.0%	0.0%	87.5%	12.5%
HRV_Trogir_Byz	5	2	3	0	0	5	0	0	5	0	0	5	0	40.0%	60.0%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%
HUN_Baden_LCA	6	5	1	1	0	5	0	0	6	0	0	6	0	83.3%	16.7%	16.7%	0.0%	83.3%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%
HUN_Bell_Beaker	8	3	5	0	1	6	1	1	7	0	0	8	0	37.5%	62.5%	0.0%	12.5%	75.0%	12.5%	12.5%	87.5%	0.0%	0.0%	100.0%	0.0%
HUN_Langobard_Mdv	27	6	21	0	9	17	1	11	16	0	0	21	6	22.2%	77.8%	0.0%	33.3%	63.0%	3.7%	40.7%	59.3%	0.0%	0.0%	77.8%	22.2%
HUN_LBK_MN	5	5	0	2	0	3	0	0	5	0	1	4	0	100.0%	0.0%	40.0%	0.0%	60.0%	0.0%	0.0%	100.0%	0.0%	20.0%	80.0%	0.0%
IND_RoopkundB	7	6	1	1	0	6	0	1	6	3	0	4	0	85.7%	14.3%	14.3%	0.0%	85.7%	0.0%	14.3%	85.7%	42.9%	0.0%	57.1%	0.0%
IRL_MN.SG	13	13	0	4	0	9	0	1	12	4	1	8	0	100.0%	0.0%	30.8%	0.0%	69.2%	0.0%	7.7%	92.3%	30.8%	7.7%	61.5%	0.0%
IRN_DinkhaTepe_BIA_A	6	5	1	5	0	1	0	0	6	4	0	2	0	83.3%	16.7%	83.3%	0.0%	16.7%	0.0%	0.0%	100.0%	66.7%	0.0%	33.3%	0.0%
IRN_DinkhaTepe_BIA_B	8	8	0	6	0	2	0	0	8	5	0	3	0	100.0%	0.0%	75.0%	0.0%	25.0%	0.0%	0.0%	100.0%	62.5%	0.0%	37.5%	0.0%
IRN_Hasanlu_IA	17	15	2	9	0	8	0	0	17	8	1	7	1	88.2%	11.8%	52.9%	0.0%	47.1%	0.0%	0.0%	100.0%	47.1%	5.9%	41.2%	5.9%
IRN_Shahr_I_Sokhta_BA	7	7	0	6	0	1	0	0	7	2	4	1	0	100.0%	0.0%	85.7%	0.0%	14.3%	0.0%	0.0%	100.0%	28.6%	57.1%	14.3%	0.0%
IRN_TepeHissar_ChL	8	8	0	6	0	2	0	0	8	5	1	2	0	100.0%	0.0%	75.0%	0.0%	25.0%	0.0%	0.0%	100.0%	62.5%	12.5%	25.0%	0.0%
ISL_Pre_Christian	7	2	5	0	3	4	0	3	4	0	0	5	2	28.6%	71.4%	0.0%	42.9%	57.1%	0.0%	42.9%	57.1%	0.0%	0.0%	71.4%	28.6%
ISR_Canaanite_MLBA	7	7	0	5	0	2	0	0	7	4	1	2	0	100.0%	0.0%	71.4%	0.0%	28.6%	0.0%	0.0%	100.0%	57.1%	14.3%	28.6%	0.0%
ISR_ChL	7	5	2	2	1	4	0	2	5	1	0	6	0	71.4%	28.6%	28.6%	14.3%	57.1%	0.0%	28.6%	71.4%	14.3%	0.0%	85.7%	0.0%
ITA_Imperial.SG	18	14	4	0	2	16	0	3	15	3	0	14	1	77.8%	22.2%	0.0%	11.1%	88.9%	0.0%	16.7%	83.3%	16.7%	0.0%	77.8%	5.6%
ITA_LateAntiquity.SG	7	6	1	3	0	4	0	1	6	2	0	5	0	85.7%	14.3%	42.9%	0.0%	57.1%	0.0%	14.3%	85.7%	28.6%	0.0%	71.4%	0.0%
ITA_LateAntiquity_oCentralEuropean.SG	5	2	3	0	0	5	0	1	4	0	0	3	2	40.0%	60.0%	0.0%	0.0%	100.0%	0.0%	20.0%	80.0%	0.0%	0.0%	60.0%	40.0%
ITA_Medieval_EarlyModern.SG	14	10	4	0	0	14	0	0	14	0	0	13	1	71.4%	28.6%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%	0.0%	92.9%	7.1%
ITA_Medieval_EarlyModern_oCentralEuropean.SG	8	5	3	0	1	7	0	2	6	0	0	7	1	62.5%	37.5%	0.0%	12.5%	87.5%	0.0%	25.0%	75.0%	0.0%	0.0%	87.5%	12.5%
ITA_North_EarlyMedieval_Langobards	13	7	6	1	2	10	0	3	10	0	0	12	1	53.8%	46.2%	7.7%	15.4%	76.9%	0.0%	23.1%	76.9%	0.0%	0.0%	92.3%	7.7%
ITA_Sardinia_BA_Nuragic	9	9	0	2	0	7	0	0	9	3	0	6	0	100.0%	0.0%	22.2%	0.0%	77.8%	0.0%	0.0%	100.0%	33.3%	0.0%	66.7%	0.0%
ITA_Sardinia_Chalcolithic	5	5	0	4	0	1	0	0	5	2	0	3	0	100.0%	0.0%	80.0%	0.0%	20.0%	0.0%	0.0%	100.0%	40.0%	0.0%	60.0%	0.0%
ITA_Sardinia_EBA	9	9	0	6	0	3	0	0	9	4	1	4	0	100.0%	0.0%	66.7%	0.0%	33.3%	0.0%	0.0%	100.0%	44.4%	11.1%	44.4%	0.0%
ITA_Sardinia_LateC	5	5	0	2	0	3	0	0	5	0	0	5	0	100.0%	0.0%	40.0%	0.0%	60.0%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%
ITA_Sardinia_Nuragic_BA	8	8	0	2	0	6	0	0	8	1	1	6	0	100.0%	0.0%	25.0%	0.0%	75.0%	0.0%	0.0%	100.0%	12.5%	12.5%	75.0%	0.0%
JOR_LBA	15	15	0	11	0	4	0	0	15	5	3	7	0	100.0%	0.0%	73.3%	0.0%	26.7%	0.0%	0.0%	100.0%	33.3%	20.0%	46.7%	0.0%
KAZ_Aktogai_MLBA	5	0	5	0	3	2	0	2	3	0	0	4	1	0.0%	100.0%	0.0%	60.0%	40.0%	0.0%	40.0%	60.0%	0.0%	0.0%	80.0%	20.0%
KAZ_Maitan_MLBA_Alakul	7	3	4	0	1	6	0	2	5	0	0	6	1	42.9%	57.1%	0.0%	14.3%	85.7%	0.0%	28.6%	71.4%	0.0%	0.0%	85.7%	14.3%
KAZ_Zevakinskiy_LBA	5	4	1	1	0	4	0	0	5	0	1	4	0	80.0%	20.0%	20.0%	0.0%	80.0%	0.0%	0.0%	100.0%	0.0%	20.0%	80.0%	0.0%
LBN_IA	6	4	2	3	0	3	0	0	6	2	0	4	0	66.7%	33.3%	50.0%	0.0%	50.0%	0.0%	0.0%	100.0%	33.3%	0.0%	66.7%	0.0%
LBN_Rom	5	5	0	3	0	2	0	0	5	2	0	3	0	100.0%	0.0%	60.0%	0.0%	40.0%	0.0%	0.0%	100.0%	40.0%	0.0%	60.0%	0.0%
LTU_EMN_Narva	5	3	2	0	0	5	0	4	1	1	0	3	1	60.0%	40.0%	0.0%	0.0%	100.0%	0.0%	80.0%	20.0%	20.0%	0.0%	60.0%	20.0%
LVA_BA	8	1	7	1	1	5	1	5	3	0	0	5	3	12.5%	87.5%	12.5%	12.5%	62.5%	12.5%	62.5%	37.5%	0.0%	0.0%	62.5%	37.5%
LVA_HG	14	13	1	1	0	13	0	4	10	6	1	7	0	92.9%	7.1%	7.1%	0.0%	92.9%	0.0%	28.6%	71.4%	42.9%	7.1%	50.0%	0.0%
LVA_MN	5	3	2	3	1	1	0	1	4	2	0	3	0	60.0%	40.0%	60.0%	20.0%	20.0%	0.0%	20.0%	80.0%	40.0%	0.0%	60.0%	0.0%
MKD_Anc	9	5	4	0	1	8	0	0	9	0	0	9	0	55.6%	44.4%	0.0%	11.1%	88.9%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%
MNE_LBA	7	7	0	1	0	6	0	0	7	0	0	6	1	100.0%	0.0%	14.3%	0.0%	85.7%	0.0%	0.0%	100.0%	0.0%	0.0%	85.7%	14.3%
NLD_Bell_Beaker	6	1	5	0	1	5	0	1	5	0	0	5	1	16.7%	83.3%	0.0%	16.7%	83.3%	0.0%	16.7%	83.3%	0.0%	0.0%	83.3%	16.7%
PAK_Katelai_IA	16	16	0	9	0	7	0	0	16	9	3	4	0	100.0%	0.0%	56.3%	0.0%	43.8%	0.0%	0.0%	100.0%	56.3%	18.8%	25.0%	0.0%
PAK_Loebanr_IA	19	17	2	14	0	5	0	0	19	12	0	7	0	89.5%	10.5%	73.7%	0.0%	26.3%	0.0%	0.0%	100.0%	63.2%	0.0%	36.8%	0.0%
PAK_Saidu_Sharif_H	9	8	1	5	0	4	0	0	9	3	2	4	0	88.9%	11.1%	55.6%	0.0%	44.4%	0.0%	0.0%	100.0%	33.3%	22.2%	44.4%	0.0%
PAK_Udegram_IA	9	9	0	7	0	2	0	0	9	2	0	7	0	100.0%	0.0%	77.8%	0.0%	22.2%	0.0%	0.0%	100.0%	22.2%	0.0%	77.8%	0.0%
POL_BKG	6	6	0	3	0	3	0	1	5	0	0	6	0	100.0%	0.0%	50.0%	0.0%	50.0%	0.0%	16.7%	83.3%	0.0%	0.0%	100.0%	0.0%
POL_Koszyce_GAC	6	3	3	0	0	6	0	2	4	0	0	6	0	50.0%	50.0%	0.0%	0.0%	100.0%	0.0%	33.3%	66.7%	0.0%	0.0%	100.0%	0.0%
PRT_LN_C	7	7	0	3	0	4	0	0	7	3	0	3	1	100.0%	0.0%	42.9%	0.0%	57.1%	0.0%	0.0%	100.0%	42.9%	0.0%	42.9%	14.3%
ROU_Bodrogkeresztur_ChL	33	23	10	10	1	22	0	5	28	6	2	24	1	69.7%	30.3%	30.3%	3.0%	66.7%	0.0%	15.2%	84.8%	18.2%	6.1%	72.7%	3.0%
RUS_Afanasievo	15	14	1	7	1	7	0	1	14	4	1	10	0	93.3%	6.7%	46.7%	6.7%	46.7%	0.0%	6.7%	93.3%	26.7%	6.7%	66.7%	0.0%
RUS_Krasnoyarsk_MLBA	11	2	9	0	4	7	0	6	5	1	0	8	2	18.2%	81.8%	0.0%	36.4%	63.6%	0.0%	54.5%	45.5%	9.1%	0.0%	72.7%	18.2%
RUS_Poltavka	5	5	0	3	0	2	0	0	5	4	0	1	0	100.0%	0.0%	60.0%	0.0%	40.0%	0.0%	0.0%	100.0%	80.0%	0.0%	20.0%	0.0%
RUS_Sintashta_MLBA	13	6	7	0	5	8	0	4	9	0	0	12	1	46.2%	53.8%	0.0%	38.5%	61.5%	0.0%	30.8%	69.2%	0.0%	0.0%	92.3%	7.7%

RUS_Srubnaya	7	5	2	0	1	6	0	2	5	0	0	7	0	71.4%	28.6%	0.0%	14.3%	85.7%	0.0%	28.6%	71.4%	0.0%	0.0%	100.0%	0.0%
RUS_Yamnaya_Samara_EBA	5	5	0	2	0	3	0	0	5	0	1	4	0	100.0%	0.0%	40.0%	0.0%	60.0%	0.0%	0.0%	100.0%	0.0%	20.0%	80.0%	0.0%
SRB_Iron_Gates_HG	25	24	1	19	0	6	0	0	25	6	10	9	0	96.0%	4.0%	76.0%	0.0%	24.0%	0.0%	0.0%	100.0%	24.0%	40.0%	36.0%	0.0%
SWE_Motala_HG	5	2	3	1	2	2	0	4	1	0	0	5	0	40.0%	60.0%	20.0%	40.0%	40.0%	0.0%	80.0%	20.0%	0.0%	0.0%	100.0%	0.0%
SYR_Ebla_EMBA	6	5	1	4	0	2	0	0	6	3	0	3	0	83.3%	16.7%	66.7%	0.0%	33.3%	0.0%	0.0%	100.0%	50.0%	0.0%	50.0%	0.0%
TUR_Aegean_Muğla_Çapalıbağ_Mdv	12	9	3	2	0	10	0	2	10	2	1	8	1	75.0%	25.0%	16.7%	0.0%	83.3%	0.0%	16.7%	83.3%	16.7%	8.3%	66.7%	8.3%
TUR_Aegean_Muğla_Değirmendere_Anc	10	7	3	2	0	8	0	0	10	2	0	7	1	70.0%	30.0%	20.0%	0.0%	80.0%	0.0%	0.0%	100.0%	20.0%	0.0%	70.0%	10.0%
TUR_Aegean_Muğla_Stratonikeia_Byz	11	10	1	2	0	9	0	0	11	1	0	9	1	90.9%	9.1%	18.2%	0.0%	81.8%	0.0%	0.0%	100.0%	9.1%	0.0%	81.8%	9.1%
TUR_C_ÇamlıbelTarlası_ChL	7	4	3	2	0	5	0	1	6	0	1	6	0	57.1%	42.9%	28.6%	0.0%	71.4%	0.0%	14.3%	85.7%	0.0%	14.3%	85.7%	0.0%
TUR_C_Gordion_Anc	5	3	2	1	0	4	0	0	5	0	1	3	1	60.0%	40.0%	20.0%	0.0%	80.0%	0.0%	0.0%	100.0%	0.0%	20.0%	60.0%	20.0%
TUR_E_Arslantepe_ChL	13	11	2	6	0	7	0	1	12	2	1	10	0	84.6%	15.4%	46.2%	0.0%	53.8%	0.0%	7.7%	92.3%	15.4%	7.7%	76.9%	0.0%
TUR_Hatay_Alalakh_MLBA	21	19	2	8	0	13	0	1	20	4	4	13	0	90.5%	9.5%	38.1%	0.0%	61.9%	0.0%	4.8%	95.2%	19.0%	19.0%	61.9%	0.0%
TUR_Marmara_Balıkesir_Byz	6	4	2	0	2	4	0	2	4	2	0	4	0	66.7%	33.3%	0.0%	33.3%	66.7%	0.0%	33.3%	66.7%	33.3%	0.0%	66.7%	0.0%
TUR_Marmara_Barcın_N	18	15	3	11	2	5	0	3	15	7	1	10	0	83.3%	16.7%	61.1%	11.1%	27.8%	0.0%	16.7%	83.3%	38.9%	5.6%	55.6%	0.0%
TUR_Marmara_İznik_Y.kapı_Byz	5	3	2	1	0	4	0	0	5	0	0	5	0	60.0%	40.0%	20.0%	0.0%	80.0%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%
TUR_SE_Kilis_MBA	6	5	1	4	0	2	0	0	6	1	2	3	0	83.3%	16.7%	66.7%	0.0%	33.3%	0.0%	0.0%	100.0%	16.7%	33.3%	50.0%	0.0%
TUR_SE_Mardin_RomByz	6	4	2	3	0	3	0	1	5	2	0	4	0	66.7%	33.3%	50.0%	0.0%	50.0%	0.0%	16.7%	83.3%	33.3%	0.0%	66.7%	0.0%
UKR_N	9	9	0	6	0	3	0	0	9	4	2	3	0	100.0%	0.0%	66.7%	0.0%	33.3%	0.0%	0.0%	100.0%	44.4%	22.2%	33.3%	0.0%
UZB_Dzharkutan1_BA	6	6	0	6	0	0	0	0	6	3	1	2	0	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	100.0%	50.0%	16.7%	33.3%	0.0%
UZB_Sappali_Tepe_BA	9	9	0	8	0	1	0	0	9	5	2	2	0	100.0%	0.0%	88.9%	0.0%	11.1%	0.0%	0.0%	100.0%	55.6%	22.2%	22.2%	0.0%

Table S 5 Phenotype distribution in West Eurasian populations with at least 5 individuals.

To get a better picture of phenotype variation in West Eurasia beyond the Southern Arc we tabulated phenotype distribution (

Table S 5) in all populations of our dataset with at least 5 individuals. We infer the presence of depigmented phenotypes in the Southern Arc, listing examples of early regional presence below:

- Blue eyes were present in the Chalcolithic of the Levant (Israel)(*72*), Neolithic of Anatolia (Turkey) at Barcın(*5*) and Chalcolithic at Arslantepe and Çamlıbel Tarlası(*31*), and Chalcolithic Southeastern Europe (Romania at Bodrogkeresztur).
- Blond hair was present in the Neolithic of Anatolia (Turkey) at Barcın(*5*), Chalcolithic Southeastern Europe (Romania at Bodrogkeresztur), Chalcolithic of the Levant (Israel)(*72*), and a Minoan from Lasithi.(*4*)
- Pale skin was inferred for Chalcolithic Southeastern Europe (Romania at Bodrogkeresztur), Iron Age Iran (Hasanlu), Croatia and Bulgaria, and Late Bronze Age Montenegro.

Did steppe groups possess these traits to a higher frequency than the inhabitants of the Southern Arc?

Blue eyes were not inferred for all 19 individuals of the Yamnaya cluster examined (Table S 4) and for 1/15 individuals of the Afanasievo culture. They were found at a higher frequency (~29-55%) at the later Middle-to-Late Bronze Age samples of the Srubnaya, Sintashta cultures and at Krasnoyarsk in Russia(*5, 34, 52, 73, 74*) and Kazakhstan (Aktogai and Maitan Alakul),(*52*) i.e., populations with elevated Anatolian/European farmer ancestry.(*5*) They were also present in Early/Middle Neolithic farmers from Central Europe including the LBK (first farmers of central Europe) and Globular Amphora culture,(*75*) and at the highest observed frequencies in farmers from Scandinavia and the Baltics (EBN Narva in Lithuania(*76*) and Motala in Sweden(*5, 10, 35*)). Similarly, blond hair was inferred for 1/34 individuals of the combined Yamnaya and Afanasievo cluster, but reached ~14-60% in the aforementioned later steppe groups. Interestingly, light pigmentation phenotype prevalence was nominally higher in the Beaker group than in Corded Ware than in the Yamnaya cluster (where as we have seen it was rare), in reverse relationship to steppe ancestry, and thus inconsistent with the theory that steppe groups were spreading this set of phenotypes.

As for the category of pale skin that is very limited in samples from the Southern Arc as a whole (1.7%), it appears to have been rare in all the studied samples in general, exceeding 1/4 in frequency only in Medieval Germany, Saxons from England, Central European outliers from

Late Antique Italy, Pre-Christian Icelanders, with the earliest high frequency found in Bronze Age Latvians at 37.5% (3/8).

Our survey of pigmentation phenotypes is not meant to be a comprehensive treatment of how these varied in space and time, but we highlight three key observations:

- The modal phenotype of the Southern Arc and West Eurasia was as expected one with dark hair, eyes, and intermediate skin pigmentation, similar to other Eurasians.
- The distinctive depigmentation found in modern groups was not associated with a particular type of ancestry in the past, as light eyes and hair were found in both West Asia and Europe, and among early farming, steppe pastoralist, as well as hunter-gatherer groups.
- The frequency of these traits could have been shaped by migration or by selection, but is more complex than simplistic stories, e.g., of these traits arising due to sexual selection in boreal hunter-gatherers(*77*) or spread by steppe Indo-Europeans.(*68*)

Surveying the history of thought on human pigmentation differences, we can remark that the ancient writers of the classical world more or less accurately described the average lighter pigmentation of populations of central/northern Europe and the Eurasian steppe, although they lacked the statistical vocabulary to express these in relative terms and exaggerated what various ancient groups (such as the "Celts" or "Scythians") looked like. Their theory that these differences were linked to climate was fundamentally flawed, as we know that people with quite different pigmentation lived in more less similar conditions of e.g., central Europe at the time of the farmers or the medieval period or the steppe in the Early Bronze Age or the time of the Scytho-Sarmatians with which they were familiar.

The promulgators of the Aryan myth also started with the present-day distribution of pigmentation phenotypes and came to a different conclusion: that these were not due to climate dictating a different phenotype for the cold north and temperate south, but rather of the existence of a primordial "race" of pale, blond, blue-eyed Proto-Indo-Europeans spreading their languages together with their phenotypes. Thus, they extrapolated the phenotype of some of their contemporaries and medieval ancestors backwards in time, postulating that it was a *survival* from the remote past that had decreased in frequency as this supposed "race" encountered and admixed with other populations. On the contrary, our survey of ancient phenotypes suggests that aspects of this phenotype were distributed in the past among diverse ancestral populations and did not coincide in any single population except as isolated individuals, and certainly not in any of the proposed homelands of the Indo-European language family.

1. I. Lazaridis, S. Alpaslan-Roodenberg et al., The genetic history of the Southern Arc: a bridge between West Asia and Europe *(in submission)*, (2022).
2. I. Lazaridis, S. Alpaslan-Roodenberg et al., Ancient DNA from Mesopotamia suggests distinct Pre-Pottery and Pottery Neolithic migrations into Anatolia. *(in submission)*, (2022).
3. Perseus Digital Library. Ed. Gregory R. Crane. Tufts University. http://www.perseus.tufts.edu (accessed 2021-2022).
4. I. Lazaridis *et al.*, Genetic origins of the Minoans and Mycenaeans. *Nature* **548**, 214-218 (2017).
5. I. Mathieson *et al.*, Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499-503 (2015).

6. Z. Hofmanová *et al.*, Early farmers from across Europe directly descended from Neolithic Aegeans. *Proceedings of the National Academy of Sciences* **113**, 6886 (2016).
7. I. Mathieson *et al.*, The genomic history of southeastern Europe. *Nature* **555**, 197-203 (2018).
8. J. L. Davis, S. R. Stocker, The Lord of the Gold Rings: The Griffin Warrior of Pylos. *Hesperia: The Journal of the American School of Classical Studies at Athens* **85**, 627-655 (2016).
9. E. R. Jones *et al.*, Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nature Communications* **6**, 8912 (2015).
10. W. Haak *et al.*, Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207-211 (2015).
11. I. Lazaridis *et al.*, Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419-424 (2016).
12. F. Clemente *et al.*, The genomic history of the Aegean palatial civilizations. *Cell*, (2021).
13. D. Adamov, V. M. Gurianov, S. Karzhavin, V. Tagankin, V. Urasin, Defining a New Rate Constant for Y-Chromosome SNPs based on Full Sequencing Data. *Russian Journal of Genetic Genealogy* **7**, 1920-2997 (2015).
14. L. D. Jack, R. S. Sharon, The Lord of the Gold Rings: The Griffin Warrior of Pylos. *Hesperia: The Journal of the American School of Classical Studies at Athens* **85**, 627-655 (2016).
15. C.-C. Wang *et al.*, Ancient human genome-wide data from a 3000-year interval in the Caucasus corresponds with eco-geographic regions. *Nature Communications* **10**, 590 (2019).
16. N. Patterson *et al.*, Large-scale migration into Britain during the Middle to Late Bronze Age. *Nature* **601**, 588-594 (2022).
17. M. B. Sakellariou, *Les proto-grecs*. (Ekdotik?? Athenon, Athens, 1980).
18. Detailed information is provided in the supplementary materials
19. R. Yaka *et al.*, Variable kinship patterns in Neolithic Anatolia revealed by ancient genomes. *Current Biology*, (2021).
20. H. Ringbauer, J. Novembre, M. Steinrücken, Parental relatedness through time revealed by runs of homozygosity in ancient DNA. *Nature Communications* **12**, 5425 (2021).
21. I. Olalde *et al.*, The genomic history of the Iberian Peninsula over the past 8000 years. *Science* **363**, 1230 (2019).
22. M. Feldman *et al.*, Ancient DNA sheds light on the genetic origins of early Iron Age Philistines. *Science Advances* **5**, eaax0061 (2019).
23. J. Diakonoff, S. A. e. Starostin, Hurro-Urartian as an Eastern Caucasian Language. *Muenchener Studien zur sprachwissenschaft*, (1986).
24. M. Narasimhan Vagheesh *et al.*, The formation of human populations in South and Central Asia. *Science* **365**, eaat7487 (2019).
25. A. S. Kassian *et al.*, Rapid radiation of the inner Indo-European languages: an advanced approach to Indo-European lexicostatistics. *Linguistics* **59**, 949-979 (2021).
26. P. A. Underhill *et al.*, The phylogenetic and geographic structure of Y-chromosome haplogroup R1a. *European Journal of Human Genetics* **23**, 124-131 (2015).
27. V. Grugni *et al.*, Ancient Migratory Events in the Middle East: New Clues from the Y-Chromosome Variation of Modern Iranians. *PLOS ONE* **7**, e41252 (2012).

28. S. Sengupta *et al.*, Polarity and temporality of high-resolution y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralists. *Am J Hum Genet* **78**, 202-221 (2006).
29. M. L. Antonio *et al.*, Ancient Rome: A genetic crossroads of Europe and the Mediterranean. *Science* **366**, 708 (2019).
30. T. Saupe *et al.*, Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula. *Curr Biol* **31**, 2576-2591.e2512 (2021).
31. E. Skourtanioti *et al.*, Genomic History of Neolithic to Bronze Age Anatolia, Northern Levant, and Southern Caucasus. *Cell* **181**, 1158-1175.e1128 (2020).
32. M. Unterländer *et al.*, Ancestry and demography and descendants of Iron Age nomads of the Eurasian Steppe. *Nature Communications* **8**, 14615 (2017).
33. P. d. B. Damgaard *et al.*, 137 ancient human genomes from across the Eurasian steppes. *Nature* **557**, 369-374 (2018).
34. M. Krzewińska *et al.*, Ancient genomes suggest the eastern Pontic-Caspian steppe as the source of western Iron Age nomads. *Science Advances* **4**, eaat4457 (2018).
35. I. Lazaridis *et al.*, Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409-413 (2014).
36. P. Ralph, G. Coop, The Geography of Recent Genetic Ancestry across Europe. *PLOS Biology* **11**, e1001555 (2013).
37. George B. J. Busby *et al.*, The Role of Recent Admixture in Forming the Contemporary West Eurasian Genomic Landscape. *Current Biology* **25**, 2518-2526 (2015).
38. E. James, *Europe's Barbarians, AD 200-600*. (Pearson Longman, 2009).
39. I. Olalde *et al.*, Cosmopolitanism at the Roman Danubian Frontier, Slavic Migrations, and the Genomic Formation of Modern Balkan Peoples. *bioRxiv*, 2021.2008.2030.458211 (2021).
40. N. Patterson *et al.*, Ancient Admixture in Human History. *Genetics* **192**, 1065-1093 (2012).
41. L. Chaitanya *et al.*, The HIrisPlex-S system for eye, hair and skin colour prediction from DNA: Introduction and forensic developmental validation. *Forensic Science International: Genetics* **35**, 123-135 (2018).
42. L. Poliakov, *Le mythe aryen. essai sur les sources du racisme et du nationalisme*. (Calmann-Levy, Paris, 1971).
43. J. V. Day, *Indo-European Origins: The Anthropological Evidence*. (Institute for the Study of Man, 2001).
44. D. Ju, I. Mathieson, The evolution of skin pigmentation-associated variation in West Eurasia. *Proceedings of the National Academy of Sciences* **118**, e2009227118 (2021).
45. S. Wilde *et al.*, Direct evidence for positive selection of skin, hair, and eye pigmentation in Europeans during the last 5,000 y. *Proceedings of the National Academy of Sciences* **111**, 4832-4837 (2014).
46. S. Alpaslan-Roodenberg *et al.*, Ethics of DNA research on human remains: five globally applicable guidelines. *Nature* **599**, 41-46 (2021).
47. L. Baringhaus, C. Franz, On a new multivariate two-sample test. *Journal of Multivariate Analysis* **88**, 190-206 (2004).
48. Q. Fu *et al.*, The genetic history of Ice Age Europe. *Nature* **534**, 200-205 (2016).

49. M. van de Loosdrecht *et al.*, Pleistocene North African genomes link Near Eastern and sub-Saharan African human populations. *Science* **360**, 548 (2018).
50. I. Lazaridis *et al.*, Paleolithic DNA from the Caucasus reveals core of West Eurasian ancestry. *bioRxiv*, 423079 (2018).
51. P. de Barros Damgaard *et al.*, The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* **360**, eaar7711 (2018).
52. V. M. Narasimhan *et al.*, The formation of human populations in South and Central Asia. *Science* **365**, eaat7487 (2019).
53. P. Moorjani *et al.*, A genetic method for dating ancient genomes provides a direct estimate of human generation interval in the last 45,000 years. *Proceedings of the National Academy of Sciences* **113**, 5652 (2016).
54. C. E. G. Amorim *et al.*, Understanding 6th-century barbarian social organization and migration through paleogenomics. *Nature Communications* **9**, 3547 (2018).
55. J. H. Marcus *et al.*, Genetic history from the Middle Neolithic to present on the Mediterranean island of Sardinia. *Nature Communications* **11**, 939 (2020).
56. R. Panzanelli, E. D. Schmidt, K. D. S. Lapatin, J. P. G. Museum, *The color of life : polychromy in sculpture from antiquity to the present*. (J. Paul Getty Museum : The Getty Research Institute, Los Angeles, 2008).
57. J. P. Mallory, V. H. Mair, Thames, Hudson, *The Tarim mummies : ancient China and the mystery of the earliest peoples from the West*. (Thames & Hudson, New York, 2008).
58. A. de Gobineau, *Essai sur l'Inégalité des Races humaines*. (F. Didot, 1855).
59. G. V. de Lapouge, *L'Aryen; son rôle social*. (A. Fontemoing, 1899).
60. M. Grant, The passing of the great race or, The racial basis of European history. (1916).
61. H. S. Chamberlain, *Die Grundlagen des neunzehnten Jahrhunderts*. (Bruckmann, München, 1898).
62. J. P. Spiro, Defending the master race : conservation, eugenics, and the legacy of Madison Grant. (2009).
63. L. Stoddard, *The rising tide of color against white world-supremacy*. (Charles Scribner's Sons, New York, 1920).
64. H. F. K. Günther, G. C. W. C. Wheeler, *The racial elements of European history*. (E.P. Dutton and Co., New York, 1927).
65. H. F. K. Günther, *Rassenkunde des deutschen Volkes*. (1937).
66. J. Chapoutot, R. R. Nybakken, Greeks, Romans, Germans : how the Nazis usurped Europe's classical past. (2017).
67. R. Peterson, The Greek Face. *Journal of Indo-European Studies* **2**, 385-406 (1974).
68. J. V. Day, *Indo-European origins : the anthropological evidence*. (Institute for the Study of Man, Washington, D.C., 2001).
69. B. O. Skaarup, Physical anthropology and the Aryan question: A historical overview and a few words of warning. *in LANGUAGE AND PREHISTORYOF THE INDO-EUROPEAN PEOPLES: A cross-disciplinary perspective*, 173-186 (2017).
70. S. Walsh *et al.*, Global skin colour prediction from DNA. *Hum Genet* **136**, 847-863 (2017).
71. S. Walsh *et al.*, Developmental validation of the HIrisPlex system: DNA-based eye and hair colour prediction for forensic and anthropological usage. *Forensic Sci Int Genet* **9**, 150-161 (2014).

72. É. Harney *et al.*, Ancient DNA from Chalcolithic Israel reveals the role of population mixture in cultural transformation. *Nature Communications* **9**, 3336 (2018).
73. M. E. Allentoft *et al.*, Population genomics of Bronze Age Eurasia. *Nature* **522**, 167-172 (2015).
74. M. Järve *et al.*, Shifts in the Genetic Landscape of the Western Eurasian Steppe Associated with the Beginning and End of the Scythian Dominance. *Current Biology* **29**, 2430-2441.e2410 (2019).
75. H. Schroeder *et al.*, Unraveling ancestry, kinship, and violence in a Late Neolithic mass grave. *Proceedings of the National Academy of Sciences* **116**, 10705-10710 (2019).
76. A. Mittnik *et al.*, The genetic prehistory of the Baltic Sea region. *Nature Communications* **9**, 442 (2018).
77. P. Frost, European hair and eye color: A case of frequency-dependent sexual selection? *Evolution and Human Behavior* **27**, 85-103 (2006).