

## RESEARCH ARTICLE SUMMARY

## HUMAN GENETICS

The genetic history of the Southern Arc:  
A bridge between West Asia and EuropeIosif Lazaridis, Songül Alpaslan-Roodenberg *et al.*

**INTRODUCTION:** For thousands of years, humans moved across the “Southern Arc,” the area bridging Europe through Anatolia with West Asia. We report ancient DNA data from 727 individuals of this region over the past 11,000 years, which we co-analyzed with the published archaeogenetic record to understand the origins of its people. We focused on the Chalcolithic and Bronze Ages about 7000 to 3000 years ago, when Indo-European language speakers first appeared.

**RATIONALE:** Genetic data are relevant for understanding linguistic evolution because they can identify movement-driven opportunities for language spread. We investigated how the changing ancestral landscape of the Southern Arc, as reflected in DNA, corresponds to the structure inferred by linguistics, which links Anatolian (e.g., Hittite and Luwian) and Indo-European (e.g., Greek, Armenian, Latin, and Sanskrit) languages as twin daughters of a Proto-Indo-Anatolian language.

**RESULTS:** Steppe pastoralists of the Yamnaya culture initiated a chain of migrations linking Europe in the west to China and India in the East. Some people across the Balkans (about 5000 to 4500 years ago) traced almost all their genes to this expansion. Steppe migrants soon admixed with locals, creating a tapestry of diverse ancestry from which speakers of the Greek, Paleo-Balkan, and Albanian languages arose.

The Yamnaya expansion also crossed the Caucasus, and by about 4000 years ago, Armenia had become an enclave of low but pervasive steppe ancestry in West Asia, where the patrilineal descendants of Yamnaya men, virtually extinct on the steppe, persisted. The Armenian language was born there, related to Indo-European languages of Europe such as Greek by their shared Yamnaya heritage.

Neolithic Anatolians (in modern Turkey) were descended from both local hunter-gatherers and Eastern populations of the Caucasus, Mesopotamia, and the Levant. By about 6500 years ago and thereafter, Anatolians became more

genetically homogeneous, a process driven by the flow of Eastern ancestry across the peninsula. Earlier forms of Anatolian and non-Indo-European languages such as Hattic and Hurrian were likely spoken by migrants and locals participating in this great mixture.

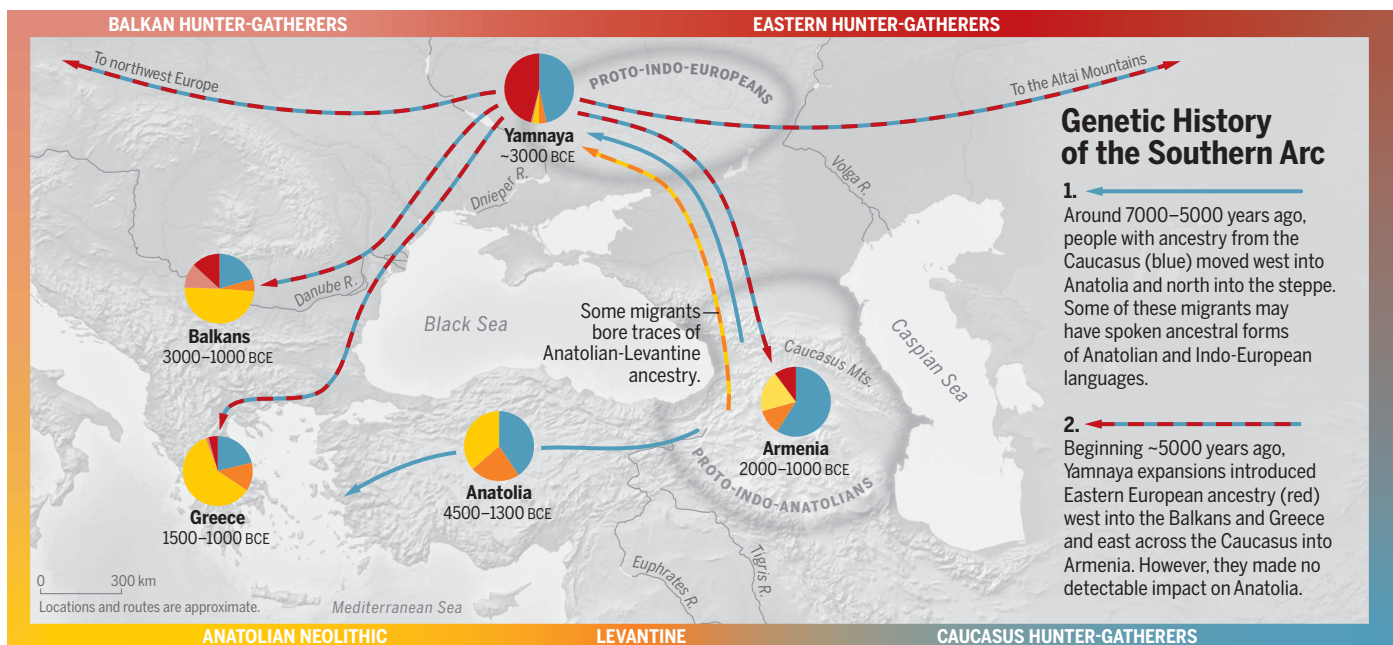
Anatolia is remarkable for its lack of steppe ancestry down to the Bronze Age. The ancestry of the Yamnaya was, by contrast, only partly local; half of it was West Asian, from both the Caucasus and the more southern Anatolian-Levantine continuum. Migration into the steppe started by about 7000 years ago, making the later expansion of the Yamnaya into the Caucasus a return to the homeland of about half their ancestors.

**CONCLUSION:** All ancient Indo-European speakers can be traced back to the Yamnaya culture, whose southward expansions into the Southern Arc left a trace in the DNA of the Bronze Age people of the region. However, the link connecting the Proto-Indo-European-speaking Yamnaya with the speakers of Anatolian languages was in the highlands of West Asia, the ancestral region shared by both. ■

The full author list and the list of author affiliations are available in the full article online.

Corresponding authors: Iosif Lazaridis, [lazaridis@genetics.med.harvard.edu](mailto:lazaridis@genetics.med.harvard.edu); Songül Alpaslan-Roodenberg, [msglalpaslan@gmail.com](mailto:msglalpaslan@gmail.com); Ron Pinhasi, [ron.pinhasi@univie.ac.at](mailto:ron.pinhasi@univie.ac.at); David Reich, [reich@genetics.med.harvard.edu](mailto:reich@genetics.med.harvard.edu)  
Cite this article as I. Lazaridis *et al.*, *Science* **377**, eabm4247 (2022). DOI: 10.1126/science.abm4247

**S READ THE FULL ARTICLE AT**  
<https://doi.org/10.1126/science.abm4247>



**Many partings, many meetings: How migration and admixture drove early language spread.** Westward and northward migrations out of the West Asian highlands split the Proto-Indo-Anatolian language into Anatolian and Indo-European branches. Yamnaya pastoralists, formed on the steppe by a fusion of

newcomers and locals, admixed again as they expanded far and wide, splitting the Proto-Indo-European language into its daughter languages across Eurasia. Border colors represent the ancestry and locations of five source populations before the migrations (arrows) and mixture (pie charts) documented here.

## RESEARCH ARTICLE

## HUMAN GENETICS

## The genetic history of the Southern Arc: A bridge between West Asia and Europe

Iosif Lazaridis<sup>1,2,\*†</sup>, Songül Alpaslan-Roodenberg<sup>2,3,\*†</sup>, Ayşe Acar<sup>4</sup>, Ayşen Açıkkol<sup>5</sup>, Anagnostis Agelarakis<sup>6</sup>, Levon Aghikyan<sup>7</sup>, Uğur Akyüz<sup>8</sup>, Desislava Andreeva<sup>9</sup>, Gojko Andrijašević<sup>10</sup>, Dragana Antonović<sup>11</sup>, Ian Armit<sup>12</sup>, Alper Atmaca<sup>13</sup>, Pavel Avetisyan<sup>7</sup>, Ahmet İhsan Aytekin<sup>14</sup>, Krum Bacvarov<sup>15</sup>, Ruben Badalyan<sup>7</sup>, Stefan Bakardzhiev<sup>16</sup>, Jacqueline Balen<sup>17</sup>, Lorenc Bejko<sup>18</sup>, Rebecca Bernardos<sup>2</sup>, Andreas Bertsatos<sup>19</sup>, Hanifi Biber<sup>20</sup>, Ahmet Bilir<sup>21</sup>, Mario Bodružić<sup>22</sup>, Michelle Bonogofsky<sup>23</sup>, Clive Bonsall<sup>24</sup>, Dušan Borić<sup>25</sup>, Nikola Borovinić<sup>26</sup>, Guillermo Bravo Morante<sup>3</sup>, Katharina Buttinger<sup>3</sup>, Kim Callan<sup>2,27</sup>, Francesca Candilio<sup>28</sup>, Mario Carić<sup>29</sup>, Olivia Cheronet<sup>3</sup>, Stefan Chohadzhiev<sup>30</sup>, Maria-Eleni Chovalopoulou<sup>19</sup>, Stella Chryssoulaki<sup>31</sup>, Ion Ciobanu<sup>32,33</sup>, Natalija Čondić<sup>34</sup>, Mihai Constantinescu<sup>35</sup>, Emanuela Cristiani<sup>36</sup>, Brendan J. Culleton<sup>37</sup>, Elizabeth Curtis<sup>2,27</sup>, Jack Davis<sup>38</sup>, Ruben Davtyan<sup>7</sup>, Tatiana I. Demcenco<sup>39</sup>, Valentin Dergachev<sup>40</sup>, Zafer Derin<sup>41</sup>, Sylvia Deskač<sup>42</sup>, Seda Devejyan<sup>7</sup>, Vojislav Djordjević<sup>43</sup>, Kellie Sara Duffett Carlson<sup>3</sup>, Laurie R. Eccles<sup>44</sup>, Nedko Elenski<sup>45</sup>, Atila Engin<sup>46</sup>, Nihat Erdoğan<sup>47</sup>, Sabiha Eriş-Pazarci<sup>48</sup>, Daniel M. Fernandes<sup>3,49</sup>, Matthew Ferry<sup>2,27</sup>, Suzanne Freilich<sup>3</sup>, Alin Frînculeasa<sup>50</sup>, Michael L. Galaty<sup>42</sup>, Beatriz Gamarra<sup>51,52,53</sup>, Boris Gasparyan<sup>7</sup>, Biserka Gaydarska<sup>54</sup>, Elif Genç<sup>55</sup>, Timur Gültekin<sup>56</sup>, Serkan Gündüz<sup>57</sup>, Tamás Hajdu<sup>58</sup>, Volker Heyd<sup>59</sup>, Suren Hobosyan<sup>7</sup>, Nelli Hovhannisyán<sup>60</sup>, Iliya Iliev<sup>16</sup>, Lora Iliev<sup>2,27</sup>, Stanislav Iliev<sup>61</sup>, İlkay İvgin<sup>62</sup>, Ivor Janković<sup>29</sup>, Lence Jovanova<sup>63</sup>, Panagiotis Karkanas<sup>64</sup>, Berna Kavaz-Kundüglü<sup>65</sup>, Esra Hilal Kaya<sup>66</sup>, Denise Keating<sup>3</sup>, Douglas J. Kennett<sup>37,67</sup>, Seda Deniz Kesici<sup>68</sup>, Anahit Khudaverdyan<sup>7</sup>, Krisztián Kiss<sup>58,69</sup>, Sinan Kılıç<sup>20</sup>, Paul Klostermann<sup>70</sup>, Sinem Kostak Boca Negra Valdes<sup>68</sup>, Saša Kovačević<sup>71</sup>, Marta Krenz-Niedbala<sup>72</sup>, Maja Krznarić Škrivanko<sup>73</sup>, Rovenka Kurti<sup>74</sup>, Pasko Kuzman<sup>75</sup>, Ann Marie Lawson<sup>2,27</sup>, Catalin Lazar<sup>76</sup>, Krassimir Leshtakov<sup>77</sup>, Thomas E. Levy<sup>78</sup>, Ioannis Liritzis<sup>79,80</sup>, Kirsi O. Lorentz<sup>81</sup>, Sylwia Łukasik<sup>72</sup>, Matthew Mah<sup>2,27,82</sup>, Swapan Mallick<sup>2,27</sup>, Kirsten Mandl<sup>3</sup>, Kristine Martirosyan-Olshansky<sup>83</sup>, Roger Matthews<sup>84</sup>, Wendy Matthews<sup>84</sup>, Kathleen McSweeney<sup>24</sup>, Varduhi Melikyan<sup>7</sup>, Adam Micco<sup>2</sup>, Megan Michel<sup>1,2,27</sup>, Lidija Milašinović<sup>85</sup>, Alissa Mittnik<sup>1,2,86</sup>, Janet M. Monge<sup>87</sup>, Georgi Nekhrizov<sup>15</sup>, Rebecca Nicholls<sup>88</sup>, Alexey G. Nikitin<sup>89</sup>, Vassil Nikolov<sup>15</sup>, Mario Novak<sup>29</sup>, Ifigo Olalde<sup>2,90</sup>, Jonas Oppenheimer<sup>2,27</sup>, Anna Osterholtz<sup>91</sup>, Celal Özdemir<sup>13</sup>, Kadir Toykan Özdoğan<sup>3</sup>, Nurettin Öztürk<sup>65</sup>, Nikos Papadimitriou<sup>92</sup>, Niki Papakonstantinou<sup>93</sup>, Anastasia Papanthanasidou<sup>94</sup>, Lujana Paraman<sup>95</sup>, Evgeny G. Paskary<sup>96</sup>, Nick Patterson<sup>1,82</sup>, Ilian Petrakiev<sup>45</sup>, Levon Petrosyan<sup>7</sup>, Vanya Petrova<sup>77</sup>, Anna Philippa-Touchais<sup>97</sup>, Ashot Piliposyan<sup>98</sup>, Nada Pocuca Kuzman<sup>75</sup>, Hrvoje Potrebica<sup>99</sup>, Bianca Preda-Bălăneț<sup>99</sup>, Zrinka Premužić<sup>100</sup>, T. Douglas Price<sup>101</sup>, Lijun Qiu<sup>2,27</sup>, Siniša Radović<sup>102</sup>, Kamal Raeuf Aziz<sup>103</sup>, Petra Rajčić Šikanjić<sup>29</sup>, Kamal Rasheed Raheem<sup>103</sup>, Sergei Razumov<sup>104</sup>, Amy Richardson<sup>84</sup>, Jacob Roodenberg<sup>105</sup>, Rudenc Ruka<sup>74</sup>, Victoria Russeva<sup>106</sup>, Mustafa Şahin<sup>57</sup>, Aysegül Şarbak<sup>107</sup>, Emre Savaş<sup>68</sup>, Constanze Schattke<sup>3</sup>, Lynne Schepartz<sup>108</sup>, Tayfun Selçuk<sup>68</sup>, Ayla Sevim-Erol<sup>109</sup>, Michel Shamoon-Pour<sup>110</sup>, Henry M. Shephard<sup>111</sup>, Athanasios Sideris<sup>112</sup>, Angela Simalcsik<sup>32,113</sup>, Hakob Simonyan<sup>114</sup>, Vitalij Sinika<sup>104</sup>, Kendra Sirak<sup>2</sup>, Ghenadie Sirbu<sup>115</sup>, Mario Šlaus<sup>116</sup>, Andrei Soficaru<sup>35</sup>, Bilal Söğüt<sup>117</sup>, Arkadiusz Sołtysiak<sup>118</sup>, Çilem Sönmez-Sözer<sup>109</sup>, Maria Stathi<sup>119</sup>, Martin Steskal<sup>120</sup>, Kristin Stewardson<sup>2,27</sup>, Sharon Stocker<sup>38</sup>, Fadime Suata-Alpaslan<sup>121</sup>, Alexander Suvorov<sup>59</sup>, Anna Szécsényi-Nagy<sup>122</sup>, Tamás Szeniczey<sup>58</sup>, Nikolai Telnov<sup>104</sup>, Strahil Temov<sup>123</sup>, Nadezhda Todorova<sup>77</sup>, Ulsi Tota<sup>74,124</sup>, Gilles Touchais<sup>125</sup>, Sevi Triantaphyllou<sup>93</sup>, Atila Türker<sup>126</sup>, Marina Ugarković<sup>71</sup>, Todor Valchev<sup>16</sup>, Fanica Veljanovska<sup>123</sup>, Zlatko Videvski<sup>123</sup>, Cristian Virag<sup>127</sup>, Anna Wagner<sup>3</sup>, Sam Walsh<sup>128</sup>, Piotr Włodarczak<sup>129</sup>, J. Noah Workman<sup>2</sup>, Aram Yardumian<sup>130,131</sup>, Evgenii Yarovoy<sup>132</sup>, Alper Yener Yavuz<sup>133</sup>, Hakan Yılmaz<sup>20</sup>, Fatma Zalzal<sup>2,27</sup>, Anna Zetti<sup>3</sup>, Zhao Zhang<sup>2</sup>, Rafet Çavuşoğlu<sup>20</sup>, Nadin Rohland<sup>2</sup>, Ron Pinhasi<sup>3,134,\*</sup>, David Reich<sup>1,2,27,82,\*</sup>

By sequencing 727 ancient individuals from the Southern Arc (Anatolia and its neighbors in Southeastern Europe and West Asia) over 10,000 years, we contextualize its Chalcolithic period and Bronze Age (about 5000 to 1000 BCE), when extensive gene flow entangled it with the Eurasian steppe. Two streams of migration transmitted Caucasus and Anatolian/Levantian ancestry northward, and the Yamnaya pastoralists, formed on the steppe, then spread southward into the Balkans and across the Caucasus into Armenia, where they left numerous patrilineal descendants. Anatolia was transformed by intra-West Asian gene flow, with negligible impact of the later Yamnaya migrations. This contrasts with all other regions where Indo-European languages were spoken, suggesting that the homeland of the Indo-Anatolian language family was in West Asia, with only secondary dispersals of non-Anatolian Indo-Europeans from the steppe.

The Balkans and Anatolia are often portrayed as being geographically peripheral to Europe and Asia rather than as central to an interconnected region spanning both continents. Here, we take a different view by providing a systematic genetic history of what we refer to as the “Southern Arc,” a region (Fig. 1A) centered on the large Anatolian peninsula (Turkey), including in the west (in Europe) the Balkans and the Aegean, and in the south and east, Cyprus, Mesopotamia, the Levant, Armenia, Azerbaijan, and Iran. We present new genome-wide DNA data from 777 individuals from the Southern Arc: 727 previously unsampled and 50 previously published for which we report new data from 1094 newly generated ancient DNA libraries (1). As a resource to guide future sampling

efforts, we also report negative results for 476 samples that we screened using 537 libraries and that failed to yield ancient DNA data meeting the criteria for authenticity (1). Finally, we provide 239 new radiocarbon dates on the same skeletal elements analyzed for DNA (1). We studied these along with the previously published individuals for a total sample size of 1317 ancient individuals in the region (Fig. 1B) (1).

Our newly reported data fill many sampling gaps in space and time in the Southern Arc. In Turkey, our new sampling has a particular focus on the western (Aegean, Marmara), northern (Black Sea), and eastern (Eastern Anatolia, Southeastern Anatolia) regions connecting it with the rest of the Southern Arc. Another area of high-density sampling is Armenia, with sub-

stantial coverage of the Bronze and Iron Ages representing an order of magnitude more individuals than previously available. Many individuals of the Bronze-to-Iron Age time frame are also sampled from the Iranian highlands at Hasanlu, where only a single individual has previously been studied (2), and from Dinkha Tepe, neighboring Anatolia, Mesopotamia, Armenia, and the Caucasus. In the southern part of Southeastern Europe, we sample Mycenaean-era individuals from multiple regions of the Aegean. From the Southern Balkans, we present a full time transect of Albania; numerous individuals from North Macedonia, where previously data from only a single Neolithic individual had been published (3); and more than double the previously available body of ancient DNA data from Bulgaria. Farther north, at the western wing of the Southern Arc, we sample individuals from Croatia, Montenegro, and Serbia in the west and Romania and Moldova in the east, which interface with the extensively studied worlds of Central Europe and the Eurasian steppe. This dataset includes >100 Bronze Age individuals, including many from Cetina Valley and Bezdanjača Cave in Croatia, which add to only five previously published from the entire area (3, 4). Some of the Balkan individuals include culturally Yamnaya individuals from Serbia and Bulgaria,

allowing us to compare them with those of the Eurasian steppe. With this greatly enhanced dataset across the entire region, we are able to fill in major gaps in sampling in time, space, and cultural context. Our large sample sizes also allow us to identify main clusters as well as genetic outliers, providing insights about within-population patterns of variation and contact networks with neighboring groups. Details of all studied individuals can be found in (I) (figs. S5 to S21).

To discuss the geographic distribution of these individuals, we take a flexible approach, in some cases using the names of ecological or

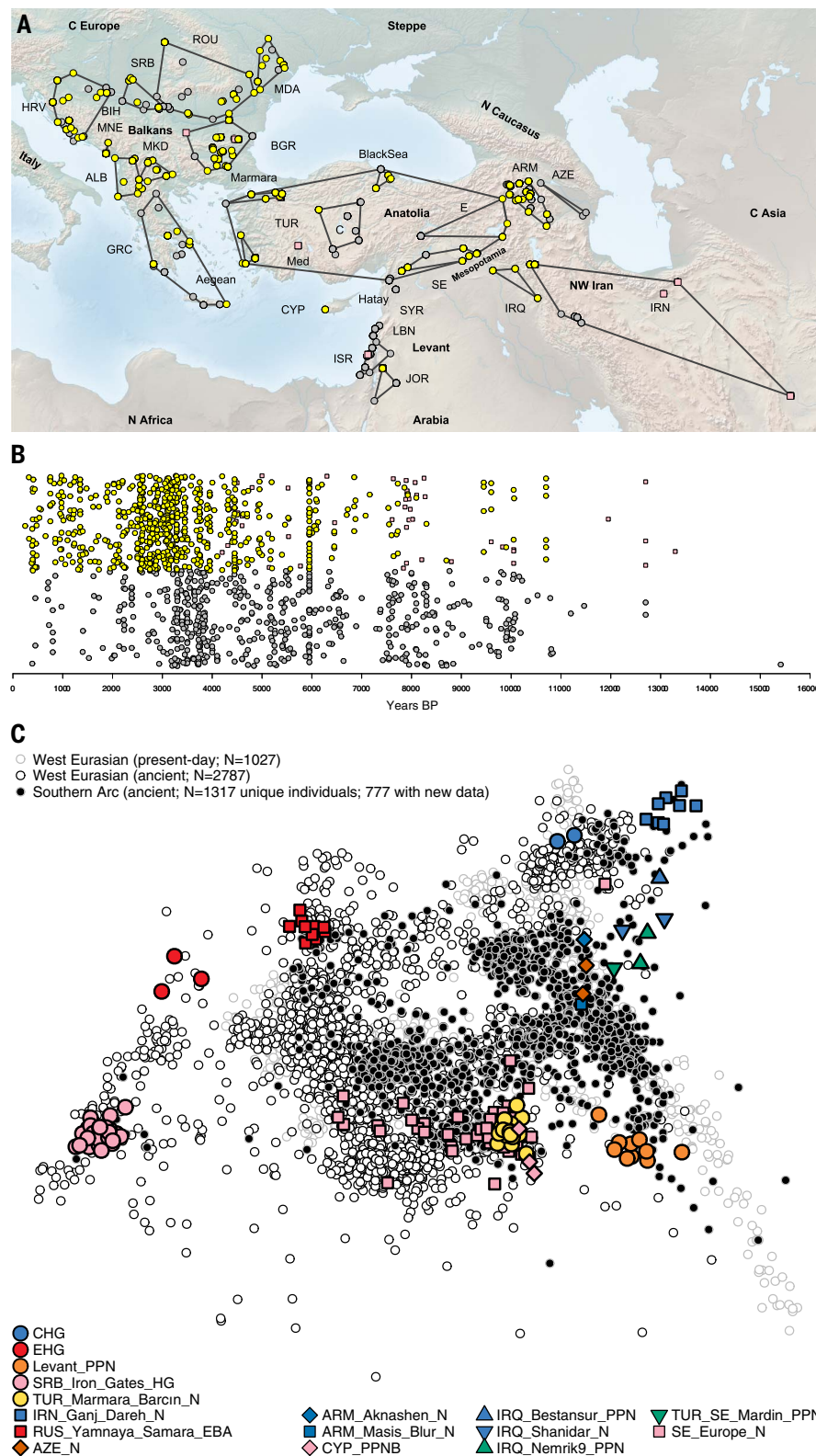
topographical regions and in others the names of present-day countries depending on how well these align with genetic patterns. In some cases, we also use more specific regional location information to add precision (5). In the interest of having a uniform nomenclature that is easily accessible to readers familiar with the current political map of the Southern Arc, we also refer to groups of individuals with labels prefixed with three-letter International Standards Organization (ISO) codes for countries, as in Fig. 1. Multiple toponyms have been used for the same sites during the Southern Arc's long history, and we typically choose labels

appropriate for the period and/or present-day usage. To designate the period in which individuals lived, we use conventional archaeological designations for each region; e.g., Eneolithic and Chalcolithic both denote copper-using cultures in different parts of the archaeological literature. We caution that the transition between the Eneolithic or Chalcolithic and the Bronze Age did not occur simultaneously in different parts of the Southern Arc. Detailed archaeological information for each individual is presented in (I), specifying the analysis labels we use integrating information from chronology, geography, archaeology, and genetics.

<sup>1</sup>Department of Human Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA. <sup>2</sup>Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. <sup>3</sup>Department of Evolutionary Anthropology, University of Vienna, 1030 Vienna, Austria. <sup>4</sup>Department of Anthropology, Faculty of Letters, Mardin Artuklu University, 47510 Artuklu, Mardin, Turkey. <sup>5</sup>Department of Anthropology, Faculty of Letters, Sivas Cumhuriyet University, 58140 Sivas, Turkey. <sup>6</sup>Department of History, Adelphi University, Garden City, NY 11530, USA. <sup>7</sup>Institute of Archaeology and Ethnography, NAS RA, 0025 Yerevan, Armenia. <sup>8</sup>Samsun Museum of Archeology and Ethnography, Kale Mahallesi, Merkez, İlkadim, 55030 Samsun, Turkey. <sup>9</sup>Iskra Museum of History, 6100 Kazanlak, Bulgaria. <sup>10</sup>Historical Museum in Kotor, 85330 Kotor, Montenegro. <sup>11</sup>Institute of Archaeology, 11000 Belgrade, Serbia. <sup>12</sup>Department of Archaeology, University of York, York YO1 7EP, UK. <sup>13</sup>Amasya Archaeology Museum, Mustafa Kemal Paşa Caddesi, 05000 Amasya, Turkey. <sup>14</sup>Department of Anthropology, Faculty of Arts and Science, Burdur Mehmet Akif University, 15100 Burdur, Turkey. <sup>15</sup>National Institute of Archaeology and Museum, Bulgarian Academy of Sciences, 1000 Sofia, Bulgaria. <sup>16</sup>Yambol Regional Historical Museum, 8600 Yambol, Bulgaria. <sup>17</sup>Archaeological Museum in Zagreb, 10000 Zagreb, Croatia. <sup>18</sup>Department of Archaeology and Heritage Studies, University of Tirana, 1010 Tirana, Albania. <sup>19</sup>Department of Animal and Human Physiology, Faculty of Biology, School of Sciences, National and Kapodistrian University of Athens, 10679 Athens, Greece. <sup>20</sup>Department of Archaeology, Faculty of Humanities, Van Yüzüncü Yil University, 65090 Tuşba, Van, Turkey. <sup>21</sup>Department of Archaeology, Faculty of Science and Letters, Düzce University, 81620 Düzce, Turkey. <sup>22</sup>Stratum Ltd., 21218 Seget Donji, Croatia. <sup>23</sup>Independent Researcher, Berkeley, CA 94720, USA. <sup>24</sup>School of History, Classics and Archaeology, University of Edinburgh, Edinburgh EH8 9AG, UK. <sup>25</sup>The Italian Academy for Advanced Studies in America, Columbia University, New York, NY 10027, USA. <sup>26</sup>Center for Conservation and Archaeology of Montenegro, 81250 Cetinje, Montenegro. <sup>27</sup>Howard Hughes Medical Institute, Harvard Medical School, Boston, MA 02115, USA. <sup>28</sup>Servizio di Bioarcheologia, Museo delle Civiltà, 00144 Rome, Italy. <sup>29</sup>Centre for Applied Bioanthropology, Institute for Anthropological Research, 10000 Zagreb, Croatia. <sup>30</sup>Department of Archaeology, University of Veliko Tarnovo "St. Cyril and St. Methodius," 5003 Veliko Tarnovo, Bulgaria. <sup>31</sup>Hellenic Ministry of Culture and Sports, Ephorate of Antiquities of Piraeus and the Islands, 10682 Piraeus, Greece. <sup>32</sup>"Orheiul Vechi" Cultural-Natural Reserve, Institute of Bioarchaeological and Ethnocultural Research, 3552 Butuceni, Moldova. <sup>33</sup>National Archaeological Agency, 2012 Chişinău, Moldova. <sup>34</sup>Archaeological Museum in Zadar, 23000 Zadar, Croatia. <sup>35</sup>Francisc I. Rainer" Institute of Anthropology, 050711 Bucharest, Romania. <sup>36</sup>Department of Oral and Maxillo-Facial Sciences, Sapienza University of Rome, 00161 Rome, Italy. <sup>37</sup>Institutes of Energy and the Environment, The Pennsylvania State University, University Park, PA 16802, USA. <sup>38</sup>Department of Classics, University of Cincinnati, Cincinnati, OH 45221, USA. <sup>39</sup>Independent Researcher, Aberystwyth SY23 4UH, UK. <sup>40</sup>Center of Archaeology, Institute of Cultural Heritage, Academy of Science of Moldova, 2001 Chişinău, Moldova. <sup>41</sup>Department of Archaeology, Faculty of Letters, Ege University, 35100 Bornova-Izmir, Turkey. <sup>42</sup>Museum of Anthropological Archaeology, University of Michigan, Ann Arbor, MI 48109, USA. <sup>43</sup>Narodni muzej Pančevo, 26101 Pančevo, Serbia. <sup>44</sup>Human Paleocology and Isotope Geochemistry Lab, Department of Anthropology, The Pennsylvania State University, University Park, PA 16802, USA. <sup>45</sup>Regional Museum of History - Veliko Tarnovo, 5000 Veliko Tarnovo, Bulgaria. <sup>46</sup>Department of Archaeology, Faculty of Science and Letters, Gaziantep University, 27310 Gaziantep, Turkey. <sup>47</sup>Mardin Archaeological Museum, Şar, Cumhuriyet Meydanı üstü, 47100 Artuklu, Mardin, Turkey. <sup>48</sup>Muğla İl Kültür ve Turizm Müdürlüğü, 48000 Muğla, Turkey. <sup>49</sup>Research Centre for Anthropology and Health (CIAS), Department of Life Sciences, University of Coimbra, 3000-456 Coimbra, Portugal. <sup>50</sup>Prahova County Museum of History and Archaeology, 100042 Ploieşti, Romania. <sup>51</sup>Institutul Català de Paleoecologia Humana i Evolució Social, 43007 Tarragona, Spain. <sup>52</sup>Departament d'Història i Història de l'Art, Universitat Rovira i Virgili, 43002 Tarragona, Spain. <sup>53</sup>School of Archaeology and Earth Institute, University College Dublin, Belfield, Dublin 4, Ireland. <sup>54</sup>Department of Archaeology, Durham University, Durham DH1 3LE, UK. <sup>55</sup>Department of Archaeology, Faculty of Science and Letters, Çukurova University, 01330 Balçalı-Sarıçam-Adana, Turkey. <sup>56</sup>Department of Anthropology, Faculty of Humanities, Ankara University, 06100 Sıhhiye, Ankara, Turkey. <sup>57</sup>Department of Archaeology, Faculty of Science and Letters, Bursa Uludağ University, 16059 Görkük, Bursa, Turkey. <sup>58</sup>Department of Biological Anthropology, Institute of Biology, Eötvös Loránd University, 1053 Budapest, Hungary. <sup>59</sup>Department of Cultures, University of Helsinki, 00100 Helsinki, Finland. <sup>60</sup>Department of Ecology and Nature Protection, Yerevan State University, 0025 Yerevan, Armenia. <sup>61</sup>Regional Museum of History, 6300 Haskovo, Bulgaria. <sup>62</sup>Ministry of Culture and Tourism, İsmet İnönü Bulvarı, 06100 Emek, Ankara, Turkey. <sup>63</sup>Museum of the City of Skopje, 1000 Skopje, North Macedonia. <sup>64</sup>Malcolm H. Wiener Laboratory, American School of Classical Studies at Athens, 10676 Athens, Greece. <sup>65</sup>Department of Archaeology, Faculty of Letters, Atatürk University, 25100 Erzurum, Turkey. <sup>66</sup>Muğla Archaeological Museum and Yatağan Thermal Power Generation Company, Rescue Excavations, 48000 Muğla, Turkey. <sup>67</sup>Department of Anthropology, University of California, Santa Barbara, Santa Barbara, CA 93106, USA. <sup>68</sup>Bodrum Museum of Underwater Archeology, Çarşı Neighbourhood, 48400 Bodrum, Muğla, Turkey. <sup>69</sup>Department of Anthropology, Hungarian Natural History Museum, 1117 Budapest, Hungary. <sup>70</sup>Department of Anthropology, Natural History Museum Vienna, 1010 Vienna, Austria. <sup>71</sup>Institute of Archaeology, 10000 Zagreb, Croatia. <sup>72</sup>Faculty of Biology, Adam Mickiewicz University in Poznań, 61-614 Poznań, Poland. <sup>73</sup>Municipal Museum Vinkovci, 32100 Vinkovci, Croatia. <sup>74</sup>Prehistory Department, Albanian Institute of Archaeology, Academy of Albanian Studies, 1000 Tirana, Albania. <sup>75</sup>National Museum in Ohrid, 6000 Ohrid, North Macedonia. <sup>76</sup>ArchaeoSciences Division, Research Institute of the University of Bucharest, University of Bucharest, 050663 Bucharest, Romania. <sup>77</sup>Department of Archaeology, St. Kliment Ohridski University of Sofia, 1504 Sofia, Bulgaria. <sup>78</sup>Department of Anthropology, University of California, San Diego, La Jolla, CA 92093, USA. <sup>79</sup>Key Research Institute of Yellow River Civilization and Sustainable Development and the Collaborative Innovation Center on Yellow River Civilization of Henan Province, Laboratory of Yellow River Cultural Heritage, Henan University, 475001 Kaifeng, China. <sup>80</sup>European Academy of Sciences and Arts, 5020 Salzburg, Austria. <sup>81</sup>Science and Technology in Archaeology and Culture Research Center, The Cyprus Institute, 2121 Aglantzia, Nicosia, Cyprus. <sup>82</sup>Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA. <sup>83</sup>Cotsen Institute of Archaeology, University of California, Los Angeles, Los Angeles, CA 90095, USA. <sup>84</sup>Department of Archaeology, University of Reading, Reading RG6 6AB, UK. <sup>85</sup>National Museum of Kikinda, 23300 Kikinda, Serbia. <sup>86</sup>Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany. <sup>87</sup>University of Pennsylvania Museum of Archaeology and Anthropology, Philadelphia, PA 19104, USA. <sup>88</sup>School of Archaeological and Forensic Sciences, Faculty of Life Sciences, University of Bradford, Bradford BD7 1DP, UK. <sup>89</sup>Department of Biology, Grand Valley State University, Allendale, MI 49401, USA. <sup>90</sup>BIOMICS Research Group, University of the Basque Country UPV/EHU, 01006 Vitoria-Gasteiz, Spain. <sup>91</sup>Department of Anthropology and Middle Eastern Cultures, Mississippi State University, Mississippi State, MS 39762, USA. <sup>92</sup>Paul and Alexandra Canellopoulos Museum, 105-55 Athens, Greece. <sup>93</sup>Faculty of Philosophy, School of History and Archaeology, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece. <sup>94</sup>Ephorate of Paleoanthropology and Speleology, Greek Ministry of Culture, 11636 Athens, Greece. <sup>95</sup>Trogir Town Museum, 21220 Trogir, Croatia. <sup>96</sup>Moldavian Historic-Geographical Society, 2044 Chişinău, Moldova. <sup>97</sup>French School of Archaeology at Athens, 10680 Athens, Greece. <sup>98</sup>Department of Armenian History, Armenian State Pedagogical University After Khachatur Abovyan, 0010 Yerevan, Armenia. <sup>99</sup>Department of Archaeology, Faculty of Humanities and Social Sciences, University of Zagreb, 10000 Zagreb, Croatia. <sup>100</sup>Independent Researcher, 10000 Zagreb, Croatia. <sup>101</sup>Laboratory for Archaeological Chemistry, University of Wisconsin-Madison, Madison, WI 53706, USA. <sup>102</sup>Institute for Quaternary Paleontology and Geology, Croatian Academy of Sciences and Arts, 10000 Zagreb, Croatia. <sup>103</sup>Sulaymaniyah Directorate of Antiquities and Heritage, 46010 Sulaymaniyah, Iraq. <sup>104</sup>Pridnestrovian University named after Taras Shevchenko, 3300 Tiraspol, Moldova. <sup>105</sup>The Netherlands Institute for the Near East, 2311 Leiden, Netherlands. <sup>106</sup>Institute of Experimental Morphology, Pathology and Archeology with Museum, Bulgarian Academy of Science, 1113 Sofia, Bulgaria. <sup>107</sup>Department of Anthropology, Faculty of Science and Letters, Hitit University, 19040 Çorum, Turkey. <sup>108</sup>School of Anatomical Sciences, The University of the Witwatersrand, 2193 Johannesburg, South Africa. <sup>109</sup>Department of Anthropology, Faculty of Language and History - Geography, Ankara University, 06100 Sıhhiye, Ankara, Turkey. <sup>110</sup>Department of Anthropology, Binghamton University, Binghamton, NY 13902, USA. <sup>111</sup>Archaeological Institute of America, Boston, MA 02108, USA. <sup>112</sup>Institute of Classical Archaeology, Charles University, 11636 Prague, Czechia. <sup>113</sup>"Olga Necrasov" Centre of Anthropological Research, Romanian Academy Iaşi Branch, 2012 Iaşi, Romania. <sup>114</sup>Scientific Research Center of the Historical and Cultural Heritage, 0010 Yerevan, Armenia. <sup>115</sup>Thracology Scientific Research Laboratory of the State University of Moldova, Department of Academic Management, Academy of Science of Moldova, 2009 Chişinău, Moldova. <sup>116</sup>Anthropological Center of the Croatian Academy of Sciences and Arts, 10000 Zagreb, Croatia. <sup>117</sup>Department of Archaeology, Faculty of Science and Arts, Pamukkale University, 20070 Denizli, Turkey. <sup>118</sup>Faculty of Archaeology, University of Warsaw, 00-927 Warszawa, Poland. <sup>119</sup>Ephorate of Antiquities of East Attica, Ministry of Culture and Sports, 10682 Athens, Greece. <sup>120</sup>Austrian Archaeological Institute at the Austrian Academy of Sciences, 1190 Vienna, Austria. <sup>121</sup>Department of Anthropology, Faculty of Letters, Istanbul University, 34134 Istanbul, Turkey. <sup>122</sup>Institute of Archaeogenetics, Research Centre for the Humanities, Eötvös Loránd Research Network, 1097 Budapest, Hungary. <sup>123</sup>Archaeology Museum of North Macedonia, 1000 Skopje, North Macedonia. <sup>124</sup>Culture and Patrimony Department, University of Avignon, F-84029 Avignon, France. <sup>125</sup>Department of the History of Art and Archaeology, Université Paris 1 Panthéon-Sorbonne, 75006 Paris, France. <sup>126</sup>Department of Archaeology, Faculty of Science and Letters, Ondokuz Mayıs University, 55139 Atakum-Samsun, Turkey. <sup>127</sup>Satu Mare County Museum, 440031 Satu Mare, Romania. <sup>128</sup>School of Natural Sciences, University of Central Lancashire, Preston PR1 2HE, UK. <sup>129</sup>Institute of Archaeology and Ethnology, Polish Academy of Sciences, 31-016 Kraków, Poland. <sup>130</sup>Department of History and Social Sciences, Bryn Athyn College, Bryn Athyn, PA 19009, USA. <sup>131</sup>Penn Museum, University of Pennsylvania, Philadelphia, PA 19104, USA. <sup>132</sup>History of the Ancient World and Middle Ages Department, Moscow Region State University, Moscow Region, 141014 Mytishki, Russia. <sup>133</sup>Department of Anthropology, Burdur Mehmet Akif Ersoy University, Istiklal Campus, 15100 Burdur, Turkey. <sup>134</sup>Human Evolution and Archaeological Sciences, University of Vienna, 1030 Vienna, Austria.

\*Corresponding author. Email: lazarisid@genetics.med.harvard.edu (I.L.A.); msglalpaslan@gmail.com (S.A.-R.); ron.pinhasi@univie.ac.at (R.P.); reich@genetics.med.harvard.edu (D.R.)

†These authors contributed equally to this work.



**Fig. 1. Studied individuals and PCA analysis.** (A) The geography of the Southern Arc. Sampling locations of previously published individuals are shown as gray circles, new data on published individuals are shown by pink squares, and new individuals are shown as yellow circles. Convex hulls of individuals from each present-day country are also shown. (B) Timeline of studied individuals (random uniform jitter applied to the vertical dimension). (C) Principal components analysis of ancient individuals projected on modern West Eurasian variation. Country names are represented by three-letter International Standards Organization (ISO) codes.

### Overview of genetic variation in the Southern Arc

To understand genetic variation in the Southern Arc, we began with ADMIXTURE (fig. S1) analysis, which allowed us to detect individuals with non-West Eurasian-associated ancestry (6) and to appreciate the broad pattern of variation in terms of the four West Eurasian components that appear in the ADMIXTURE analysis: Iran/Caucasus-related, “Eastern hunter-gatherer,” Anatolian/Levantine-related, and “Balkan hunter-gatherer.” Principal components analysis (Fig. 1C) of Southern Arc individuals together with other West Eurasian individuals demonstrates the central position of the Southern Arc within the continuum of West Eurasian variation, with a long “bridge” of individuals joining Europe (left) to West Asia (right), but with individuals spread across the entire range of variation.

To quantify the ancestry of Southern Arc individuals, we developed a five-source modeling framework (using qpAdm and F4admix) (1) that allows a high-resolution description of the ancestry of the Southern Arc population as a whole and as individuals. To generate this model, we used an automated procedure that did not preselect a specific set of surrogates for the source populations, but instead explored many possible sets and identified those that, for as many individuals as possible, maximized the quality of the statistical fit of the model while minimizing the standard errors in inferences of ancestry proportions (tables S1 to S21 and figs. S22 to S27). After applying this procedure, the five sources of ancestry that we used are: Caucasus hunter-gatherers (CHG) (7), Eastern hunter-gatherers (EHG) from Europe (8, 9), Levantine Pre-Pottery Neolithic (10), Balkan hunter-gatherers from the Iron Gates in Serbia (3), and Northwestern Anatolian Neolithic from Barcin (9). These correspond to the four-source ADMIXTURE model, with further distinction between the Anatolian and Levantine ends of the “Mediterranean” interaction zone (11). These five sources should not be unduly emphasized beyond their utility as a descriptive convenience because (i) they could be swapped for related ones [e.g., Neolithic Iran captures much of the same deep ancestry as Caucasus hunter-gatherers do (10, 11)], (ii) they were themselves derived from earlier (more “distal”) populations [e.g., Levantine Pre-Pottery Neolithic from earlier Natufian hunter-gatherers (10)], and (iii) they transmitted their ancestry through later (more “proximal”) sources [e.g., Eastern hunter-gatherers through Yamnaya steppe pastoralists (8)]. The inferred proportions of ancestry for individuals are summarized in figs. S2 to S4 and figs. S28 to S76 and are discussed in detail in (1).

### The Anatolian core of the Southern Arc

When we apply our five-way model to individuals from Anatolia (Fig. 2, A to E), it is



the Neolithic ones that preceded them (Fig. 2F). We observed that Northwest Anatolian-related ancestry varied between ~100% (at Barcin, Menteşe, and Ilımar in the Marmara region; we use the high-quality data we have from Barcin to define this component of ancestry) to ~16% (the Pre-Pottery Neolithic individual from Mardin in Southeast Anatolia/North Mesopotamia). Conversely, Caucasus/Levantine ancestry varied between ~50 and ~32% in North Mesopotamia to ~0% in Northwest Anatolia.

The Chalcolithic period in Anatolia has a wide temporal range (Fig. 2) that spans from the end of the Neolithic (~6000 BCE) to the beginning of the Bronze Age (~3000 BCE). Individuals in our analysis are mostly from the Late Chalcolithic (after ~4500 BCE) and from the entirety of the Bronze Age (down to 1300 BCE). Both Chalcolithic and Bronze Age populations from all regions generally had intermediate admixture proportions within the Neolithic ranges of ancestry. This suggests that they could be modeled as drawn from mixtures of the preceding Neolithic populations. In the Marmara region, Caucasus hunter-gatherer ancestry increased from ~0 to ~33% between the Neolithic and Chalcolithic periods [to define the Chalcolithic, we added four individuals from Ilımar to a single one from Barcin previously published (10)]. In the Central region, we document an increase from ~10 to 15% at Neolithic Çatalhöyük (12) and Tepecik-Çiftlik (13) to a similar ~33% at Chalcolithic Çamlıbel Tarlası (14) and ~42% at Bronze Age Kalehöyük and Ovaören (15). In the Mediterranean region (Southwest Anatolia), the same approximate one-third proportion was present at Harmanören Göndürle (16) in the Bronze Age. In the Aegean region (Western Anatolia), we observe a similar ~29% in the Bronze Age. Thus, individuals from more western regions of Anatolia (Marmara, Aegean, Central, and Mediterranean) all had more Caucasus-related ancestry (and correspondingly less Anatolian-related ancestry) during the Chalcolithic and Bronze Age than the preceding Neolithic populations of the area, suggesting that a spread of this ancestry westward across the peninsula occurred after the Neolithic, a pattern also observed in the Levant (11). In the more eastern regions of Anatolia [East, in Arslantepe (14); Southeast, from Batman, Gaziantep, Kilis, and Şırnak (new data) and Tiritiş Höyük (14); Black Sea, from Devret Höyük in Amasya and Samsun (new data) and İkiztepe (14)], populations of the Chalcolithic and Bronze Age periods had, conversely, more Western Anatolian Neolithic-related, and less Caucasus-related ancestry, than the Pre-Pottery Neolithic individual from Mardin. This pattern is also observed when we compare the Chalcolithic with the Bronze Age. Differences are small but all in the direction

of more Western Anatolian Neolithic-related ancestry (an increase of ~3 to 7% in the East, Southeast, and Black Sea regions) except in the Hatay Province (14), where Western Anatolian Neolithic-related ancestry decreased and Caucasus-related ancestry increased (from ~14 to 43%) between the Early Chalcolithic (~5500 BCE) and the Middle to Late Bronze Age (after ~2000 BCE).

Taken as a whole, the genetic history of Anatolia during the Chalcolithic and Bronze Age can be characterized as one of homogenization. Neolithic populations differed by as much as ~80% in terms of Western Anatolian Neolithic-related and by ~50% in terms of Caucasus-related ancestry. In the Chalcolithic and Bronze Age, the range of these differences narrowed substantially. That of Western Anatolian Neolithic-related ancestry halved to ~40% (becoming ~20 to 60%) and that of Caucasus-related ancestry to ~15% (becoming ~30 to 45% except in the Hatay Province). Despite this homogenization, some ancestry differences persisted. The eastern regions retained more Caucasus-related ancestry than the western ones, but the overall pattern was one of attenuated differentiation after intra-Anatolian gene flow stemming from the highly differentiated Neolithic populations of Western/Central Anatolia on the one hand and Northern Mesopotamia on the other (as well as hitherto unsampled others).

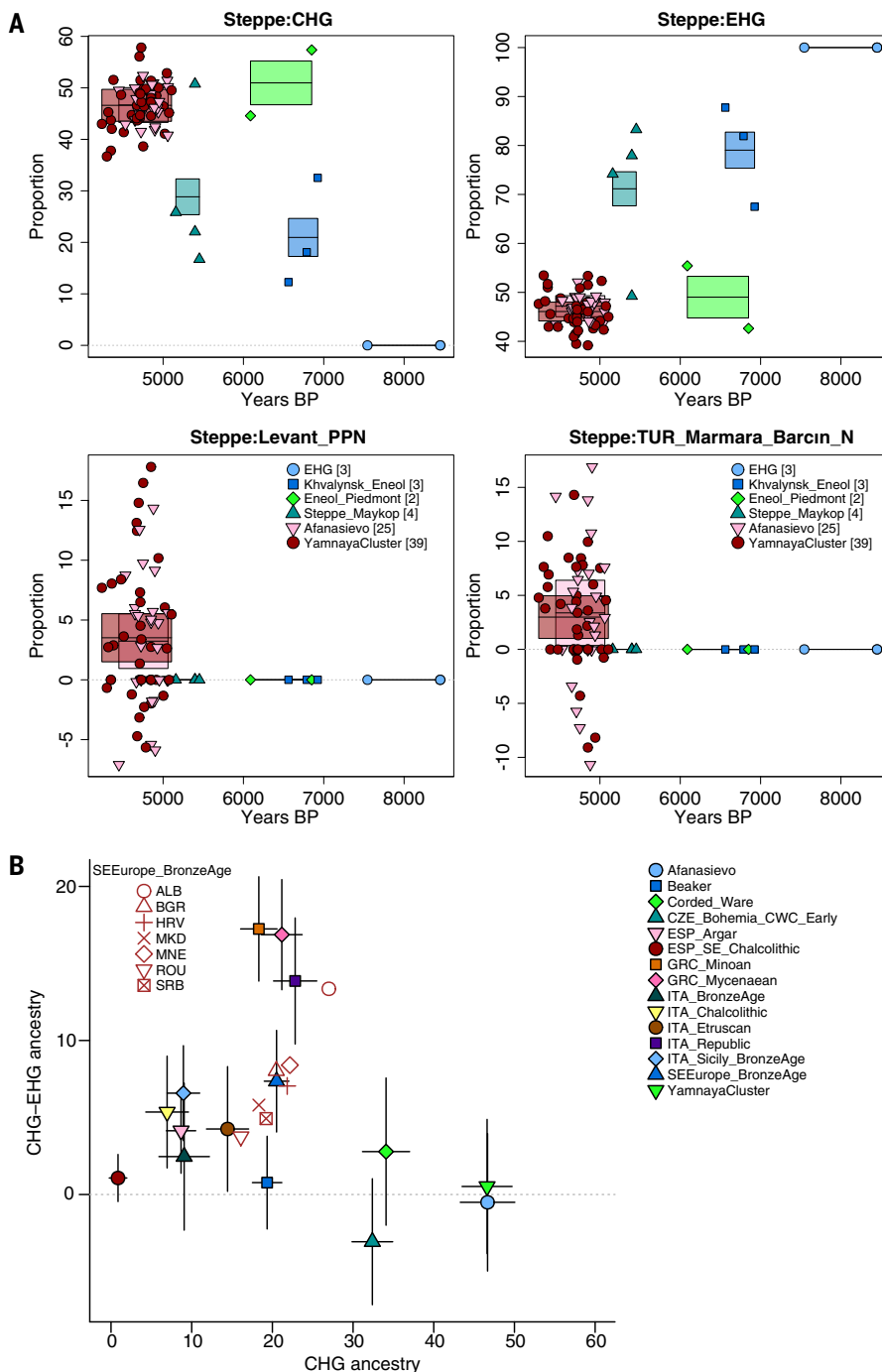
Homogenization in Anatolia was coupled by impermeability to exogenous gene flow from Europe, which could be explained by either a large and stable population base that attenuated the demographic impact of external immigration or cultural factors impeding it. The asymmetry of gene flow between Anatolia and its neighbors is evident, for example, in the fact that Caucasus hunter-gatherer-related ancestry flowed westward across Anatolia into the Balkans and northward into the Eurasian steppe, but Balkan hunter-gatherer ancestry did not flow into Anatolia or further eastward, and Eastern hunter-gatherer ancestry entered West Asia only as far south as Armenia and, to a lesser extent, Iran (as we will see below). This was true even down to the Urartian period of the Iron Age, when a population lacking Eastern hunter-gatherer ancestry still existed in the center of the Kingdom of Van (6).

### The origin and expansion of steppe pastoralists

The absence of European hunter-gatherer admixture in Anatolia during the Chalcolithic and Bronze Age periods contrasts with developments to the north of the Southern Arc and north of the Black and Caspian Seas, which saw the formation of Eneolithic (a term used instead of Chalcolithic for this area) and Bronze Age pastoralist populations that harbored a mixture of populations from Eastern Europe and the Southern Arc (8, 9, 17). Examining

individuals from the steppe (Fig. 3), we observe that in the post-5000 BCE period, Caucasus-related ancestry is added to the previous Eastern hunter-gatherer population, forming the Eneolithic populations at Khvalynsk (9) and Progress-2 (17); this ancestry persisted in the Steppe Maykop population of the 4th millennium BCE (17). However, all of these populations before ~3000 BCE lack any detectable Anatolian/Levantine-related ancestry, contrasting with all contemporaneous ones from the Southern Arc, which have at least some such ancestry at least since the Neolithic (11). In all later periods in the Southern Arc, Caucasus hunter-gatherer-related ancestry is never found by itself but rather is always admixed, to various degrees, with Anatolian/Levantine ancestry. This suggests that whatever the source of the Caucasus-related ancestry in the Eneolithic steppe, it cannot have been from the range of variation sampled in the Southern Arc because this would have introduced Anatolian/Levantine-related ancestry. This implies that the proximal source of the Caucasus-related ancestry in the Eneolithic steppe should be sought in an unsampled group that did not experience Anatolian/Levantine-related gene flow until the Eneolithic. Plausibly, this population existed in the North Caucasus, from which Caucasus hunter-gatherer-related, but not Anatolian/Levantine-related, ancestry could have entered the Eneolithic steppe.

The Eneolithic steppe population contrasts with that of the Yamnaya cluster of individuals ~3000 BCE, which does have significant Anatolian (3 ± 1%) and Levantine (3.5 ± 1%) related ancestry [Fig. 3A; steppe individuals in this analysis are listed in (1)]. This inference is further supported by detailed analysis of Yamnaya ancestry at different time depths (tables S22 to S28) (1), which indicates that they derived from at least two southern sources. The first source dates to the Neolithic and includes Caucasus hunter-gatherer ancestry only. The second source dates to before the formation of the Yamnaya cluster and includes Anatolian/Levantine-related ancestry in addition to Caucasus hunter-gatherer (as deep sources), ancestry related to Neolithic people of Armenia (more proximally), or ancestry related to Chalcolithic people of the Caucasus to Southeast Anatolia (even more proximally). A more direct and geographically proximate source in the Maykop population of the North Caucasus of the 4th millennium BCE has also been proposed (18). Although the exact source cannot at present be determined (all of the candidates have different combinations of the same Anatolian/Levantine/Caucasus ancestry; fig. S1), it was people drawn from this metapopulation in the Chalcolithic Caucasus, Armenia, and East/Southeast Anatolia that must have been responsible for the second pulse of Southern Arc ancestry into the precursors of Yamnaya



**Fig. 3. Yamnaya origins and expansions.** (A) The earliest inhabitants of the steppe (EHG) were followed by CHG-admixed populations by ~5000 BCE and by Anatolian/Levantine-admixed populations by ~3000 BCE with the emergence of the Yamnaya-Afanasio genetic cluster. The proportion of Balkan hunter-gatherer-related ancestry (not shown) is  $0.8 \pm 0.6\%$  in the Yamnaya cluster and  $-0.5 \pm 0.5\%$  in the Afanasievo. (B) The Yamnaya had nearly half their ancestry from CHG, higher than any Bronze Age Europeans from the Balkans, Italy, or Central/Northern Europe, but their CHG-EHG balance was equal, similar to the Corded Ware/Beaker clusters of Central/Northern Europe and contrasting with Southeastern and Mediterranean Europe, where CHG was significantly higher than EHG. 95% confidence intervals of  $\pm 1.96$  SE are shown.

steppe pastoralists. The genetic contribution of the second pulse may have been as low as 6.5%, the sum of Anatolian and Levantine ancestry in the Yamnaya, or as high as 53.1%, the

totality of the combined Caucasus hunter-gatherer and Anatolian/Levantine ancestry. The low end is unlikely because Caucasus hunter-gatherer ancestry was ubiquitous in

West Asia during the Chalcolithic period and some of it should be added to the 6.5% figure. The high end is also unlikely because it suggests that all Caucasus hunter-gatherer ancestry flowed northward with the second pulse, thus ignoring the evidence for its independent flow into the Eneolithic steppe. Our modeling suggests intermediate values of ~21 to 26% (table S28), in the middle of the 6.5 to 53.1% range, an estimate that may be updated in the future as better proximate sources in both West Asia and the steppe come to light.

Archaeological evidence documents how western steppe populations interacted with European farmer groups such as the Cucuteni-Trypillia and Globular Amphora cultures, and it was previously suggested that ancestry from such groups contributed to the ancestry of the Yamnaya (17). Our genetic results contradict this scenario because European farmers were themselves a mixture of Anatolian Neolithic and European hunter-gatherer ancestry, but the Yamnaya lacked the European hunter-gatherer ancestry differentiating European from West Asian farmers, and had an ~1:1 ratio of Levantine-to-Anatolian ancestry in our five-way model, contrasting with the overwhelming predominance of Anatolian ancestry in European farmers. The Caucasus hunter-gatherer/Eastern hunter-gatherer/Western hunter-gatherer/Anatolian Neolithic model of (17) fails ( $P < 1 \times 10^{-10}$ ) because it underestimates shared genetic drift with Levantine farmers ( $Z = 5.6$ ), whose contribution into the Yamnaya cannot be explained under that model. These results shift the quest for the ancestral origins of a component of Yamnaya ancestry firmly to the south of the steppe and the eastern wing of the Southern Arc. Determining the proximate source of the two movements into the steppe from the south will depend on further sampling across the Anatolia-Caucasus-Mesopotamia-Zagros area where populations with variations of the three components existed. Similarly, on the steppe side, study of Eneolithic (pre-Yamnaya) individuals could disclose the source dynamics of Caucasus hunter-gatherer infiltration northward and identify the likely geographical region for the emergence of the distinctive Yamnaya cluster, which we show has an autosomal signal of admixture dating to the mid-5th millennium BCE [fig. S5 and (19)], coinciding with the direct evidence of the first southern influence provided by the Eneolithic individuals of the steppe.

The role of Yamnaya-like populations in spreading both Eastern hunter-gatherer and West Asian ancestry into mainland Europe has been previously recognized (8), but it has also become apparent that some of the latter entered Europe independently of steppe expansions into the Aegean (9, 16), Sicily (20), and even as far west as Iberia (21) by the Bronze Age. We observe that the Caucasus

minus Eastern hunter-gatherer ancestry difference in the Yamnaya is  $\sim 0\%$  (Fig. 4B), and this allows us to both test whether steppe migrants into mainland Europe may have originated from a different steppe population (with a nonequal balance of Caucasus and Eastern hunter-gatherer components) and whether additional migrations (with either more Eastern or Caucasus hunter-gatherer ancestry, thus shifting the difference away from zero) occurred. We find that the Corded Ware and Bell Beaker complex individuals from Europe are all consistent with a balanced presence of the two components (consistent with having been transmitted through a Yamnaya-like population). Even in the early Corded Ware from Bohemia, where a third “northern” source has been suggested to have been substantially involved (22), the difference is one of a small  $3.1 \pm 2.1\%$  excess of Eastern hunter-gatherer ancestry, which is entirely consistent with being transmitted entirely by the Yamnaya to

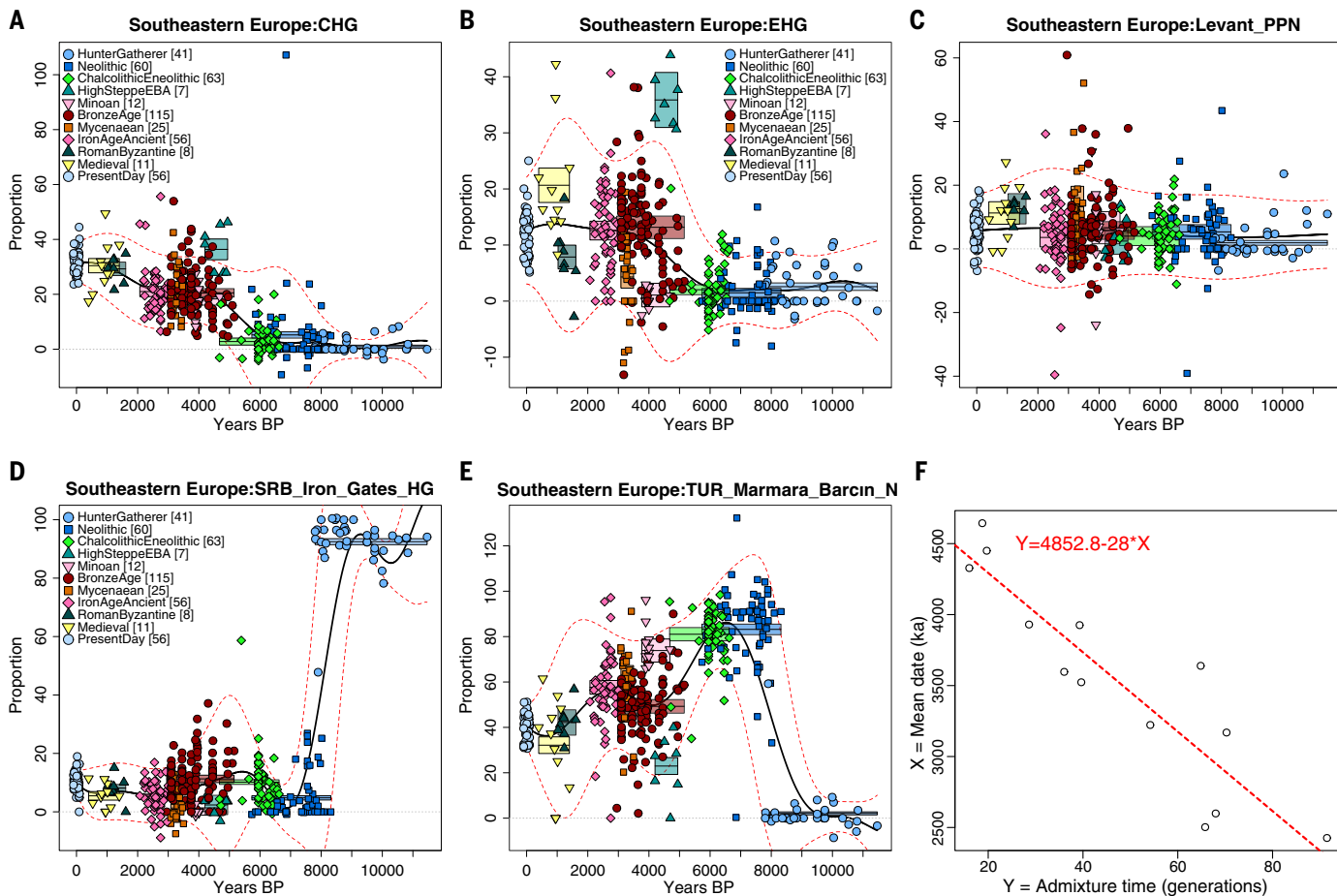
the limits of the resolution of our statistical analysis. This is not the case for Southeastern Europe, where Bronze Age individuals had an excess of Caucasus over Eastern hunter-gatherer ancestry not only in the Aegean ( $\sim 17\%$  in both Minoans and Mycenaeans) (16), but throughout the Balkan peninsula (Fig. 3B), where the overall Bronze Age excess is  $7.4 \pm 1.7\%$  (with by-country estimates of  $\sim 4$  to  $13\%$ ). A possible explanation for this excess is the existence of a small  $5.2 \pm 0.6\%$  Caucasus hunter-gatherer component in the Neolithic substratum of Southeastern Europe (Fig. 4A); we estimated that this proportion is  $\sim 0$  to  $1\%$  in four separate Early Neolithic populations from Hungary (Starčevo-Körös cultural complex), France, Spain, and the Linearbandkeramik of Austria, Germany, and Hungary (3, 23–30). Thus, the Bronze Age Caucasus hunter-gatherer ancestry in Southeastern Europe compared with Central/Northern/Western Europe may replicate this contrast from the Neolithic. However,

the even higher levels observed in the Aegean [Fig. 3B and (6)] suggest additional gene flow after the Neolithic by the time of the Early Bronze Age (31).

### Interplay of local, steppe, and West Asian ancestries in Southeastern Europe

Southeastern Europe interfaces geographically with both the Eurasian steppe and Anatolia, and its genetic history (Fig. 4) bears traces of both connections, starting from the partial replacement of its local Balkan hunter-gatherers by Anatolian Neolithic farmers beginning  $\sim 8500$  years ago, followed by the expansion of Eastern hunter-gatherer-ancestry-bearing steppe populations  $\sim 5000$  years ago (3). Although the Bronze Age was a period of partial homogenization in Anatolia, as we have seen, in Southeastern Europe, it was a time of substantial contrasts.

One aspect of this heterogeneity was the retention of the local Balkan hunter-gatherer



**Fig. 4. Genetic heterogeneity in Southeastern Europe after the Yamnaya expansion.** (A to E) Five components of ancestry in Southeastern Europe. The replacement of hunter-gatherer by early farmer ancestry [(D) and (E)] was followed by the rise of CHG and EHG ancestry over the past 5000 years [(A) and (B)], with Levantine ancestry being relatively unimportant and showing no discernible temporal pattern (C). In (F), we show a linear regression of population

dates (using directly radiocarbon-dated individuals for each population) on admixture times in generations; more recent populations have older admixture times, and the regression places admixture between populations related to the Southeast European Neolithic and Yamnaya at  $4853 \pm 205$  years ago and the generation length at  $28 \pm 4$  years, virtually identical to its independent empirical estimation of 28 years.

ancestry itself, which was detected only in the Balkans (within the Southern Arc), thus precluding any substantial migration from the area to the rest of the Southern Arc. Balkan hunter-gatherer ancestry was variable during the Bronze Age and related to geography. A marked contrast is found within Romania, where our new data show that it makes up ~12% of the ancestry of 42 individuals from the Bodrogkeresztúr Chalcolithic and ~24 to 30% in 10 Bronze Age individuals from Cărlomănești (Arman) and from Ploiești and Târgșoru Vechi south of the Carpathian Mountains. Together with another Bronze Age individual from Padina in Serbia [2460 to 2296 calibrated (cal) BCE] near the Iron Gates, whose Balkan hunter-gatherer ancestry was ~37%, these results prove substantial hunter-gatherer ancestry preservation in the North Balkans postdating the arrival of both Anatolian Neolithic and steppe ancestry in the region. This contrasts with the southern end of the Balkan peninsula in the Aegean (6), where neither the Neolithic nor the Bronze Age populations had any significant Balkan hunter-gatherer ancestry, raising the question of whether the region's pre-Neolithic population was more similar to that of the North Balkans (Balkan hunter-gatherer-like) or Western Anatolia (and thus similar to the Neolithic population).

The key driver of the Bronze Age heterogeneity was the appearance of Eastern hunter-gatherer ancestry that became ubiquitous in Southeastern Europe after its sporadic Chalcolithic appearance (3). This is most evident (~31 to 44%) in Moldova at several Bronze Age sites, including those of the Catacomb and Multi-cordoned Ware cultures, and individuals from Romania (Trestiana and Smeeni) on the eastern/southeastern slopes of the Carpathians, which contrast with the high-Balkan hunter-gatherer group from Arman. We also detect a contrast between Catacomb culture individuals from Moldova and those from the Caucasus (17), driven by an individual from Purcari with substantial (17 ± 4%) Anatolian Neolithic ancestry, suggesting some heterogeneity within this culture on opposite sides of the Black Sea. For the rest of the Balkans, the amount of Eastern hunter-gatherer ancestry is ~15% and drops to ~4% in Mycenaean Greece and to negligible levels in Minoan Crete (6, 16).

Our study identifies a “high-steppe ancestry” set of individuals, a term we use to refer to individuals from the Balkans during the Early Bronze Age who had unusually high proportions of Eastern hunter-gatherer ancestry compared with their contemporaries (Fig. 4B). This includes two previously published individuals from Nova Zagora in Bulgaria and Vucedol in Croatia (3), as well as five newly reported individuals, including an Early Bronze Age individual from Çinamak in Albania (2663 to 2472 calBCE) and four that are culturally

Yamnaya: one from Vojlovica-Humka in Serbia, two from Boyanovo, and one from Mogila in Bulgaria. In aggregate, this group of Balkan individuals has 35.9 ± 2.5% Eastern hunter-gatherer, 36.4 ± 1.9% Caucasus hunter-gatherer, and 23.0 ± 1.9% Anatolian Neolithic ancestry compared with the Yamnaya cluster individuals (46.1 ± 1.0%, 46.6 ± 1.6%, and 3.0 ± 1.0%, respectively), i.e., the same Caucasus/Eastern hunter-gatherer balance as the Yamnaya but diluted by about one-fifth by local Neolithic ancestry of ultimately Anatolian origin.

When we use DATES (19) to date the admixture of steppe ancestry in populations of Southeastern Europe (Fig. 5F and fig. S6), we arrive at an estimate that this took place ~4850 years ago, i.e., precisely after the Yamnaya expansion, and within the time frame of our “high-steppe” cluster individuals. This suggests that (as a first approximation) steppe ancestry in Southeastern Europe from the Bronze Age onward was largely mediated by descendants of Yamnaya and local Balkan populations and not by earlier waves out of the steppe that affected the region sporadically. This admixture need not have taken place in one locality, as indicated by the presence of Yamnaya-like individuals in several regions of the Balkans, spatially beyond both the cultural transition zone between steppe pastoralist and settled populations (32), and the geographical one from the Eastern European flatlands into mountainous areas.

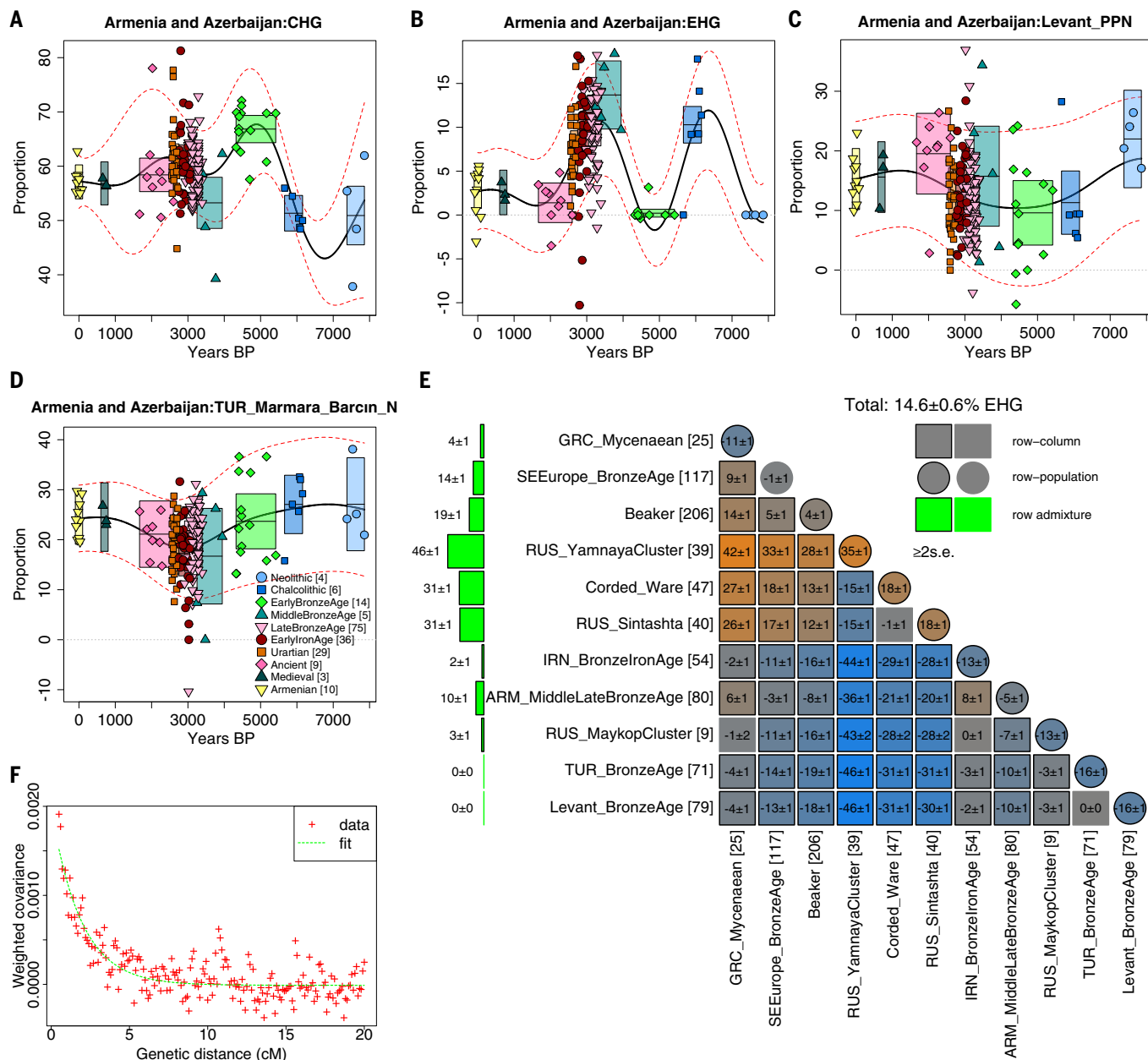
#### Armenia: Fluctuating steppe ancestry against a persistent West Asian genetic background

Armenia is situated in the highlands of West Asia to the east of Anatolia and to the south of the Caucasus mountains separating West Asia from the Eurasian steppe to the north. When we examine the trajectory of ancestry there (Fig. 5), we observe that the local Caucasus hunter-gatherer-related ancestry (Fig. 5A) has always been the most important component of the population from the Neolithic to the present, making up ~50 to 70% of ancestry over the past 8000 years. As in Anatolia, the two other components of West Asian ancestry had a strong presence as well, making up most of the remainder.

The most noticeable feature of the history of Armenia compared with all other Asian regions of the Southern Arc is the tentative appearance of Eastern hunter-gatherer ancestry in the Chalcolithic at Areni-1 Cave (10) ~6000 years ago (Fig. 5B), followed by its disappearance ~5000 years ago with the Early Bronze Age Kura-Araxes culture and its reappearance at the Middle Bronze Age, when a level of ~14% was followed by ~10% in the Late Bronze Age and Iron Age and then diluted to ~7% by the Urartian period of the first half of the 1st millennium BCE and to the ~1 to 3% levels observed since the second half of that millen-

nium at sites such as Aghitu and through the medieval period (at Agarak) down to present-day Armenians. When we compare the Middle/Late Bronze Age individuals from Armenia (when Eastern hunter-gatherer ancestry was highest and from which we have individuals from >20 sites) with other West Asian European and steppe populations (Fig. 5E), it is evident that Armenia is an outlier. Populations from Armenia have significantly more such ancestry than all surrounding populations: Anatolia and the Levant, where this ancestry is undetected during the Bronze Age; Iran, where it makes up ~2% overall; and even the Maykop cluster populations of the North Caucasus (17), where it reaches ~3%. These analyses in Armenia show that Eastern hunter-gatherer ancestry flowed from the steppe not only west of the Black Sea into Southeastern Europe, attaining its minimum in the Aegean and east of it, but also across the Caucasus into Armenia. However, substantial proportions of steppe ancestry spread no further into Anatolia from either west or east.

The appearance of Eastern hunter-gatherer ancestry at Areni-1 Cave is the first known genetic influence of peoples of the Eurasian steppe on West Asia, although with our current sparse sampling of the Eneolithic steppe, we do not know the precise geographical source of this ancestry within the steppe. The Areni individuals date to the same 5th millennium BCE, in which we saw that the Eneolithic steppe came to be influenced by Caucasus hunter-gatherer-related ancestry from the south and to which our admixture dating of Yamnaya origins also points. However, it was only during the Middle/Late Bronze Age that Eastern hunter-gatherer ancestry became entrenched in Armenia, at least for a while, forming an “enclave” of steppe influence in West Asia that eventually dissipated during the 1st millennium BCE. This period of relatively high-steppe ancestry corresponds to the “Lchashen-Metsamor” culture of the Bronze-to-Iron Age (1). Linkage disequilibrium dating of steppe admixture (Fig. 5F) in our extensive set of individuals of average late 2nd millennium BCE date suggests it occurred a millennium and a half earlier, at the middle of the 3rd millennium BCE, and thus in parallel to the transformation of mainland Europe and the Balkans. In Armenia itself, the mid-3rd millennium BCE corresponds to the demise of the Kura-Araxes culture and its succession by the “Early Kurgan” culture, followed during the end of that millennium by the “Trialeti-Vanadzor” complex from which an individual from Tavshut (2127 to 1900 calBCE) already has the ~10% Eastern hunter-gatherer ancestry of the Lchashen-Metsamor population, the first documented steppe descendant in Armenia two millennia after the Chalcolithic. The analysis of Y chromosomes to which we now turn provides an independent line of evidence



**Fig. 5. A genetic history of Armenia.** Shown are changes in the four components of ancestry. **(A)** CHG is the most important component in all ages, rising to its maximum in the Kura-Araxes culture of the Early Bronze Age. **(B)** EHG ancestry first appears in the Chalcolithic at Areni Cave, disappears during the Kura-Araxes period, reappears strongly in the Middle-to-Late Bronze Age period, and decreases to about one-third of its peak value by ~2000 years ago. **(C and D)** Levantine and Anatolian ancestry were present in all periods as minority components. Balkan hunter-gatherer ancestry (not shown) is <1% in all

periods. All individuals shown are from Armenia save for two Neolithic and a Chalcolithic individual previously published from Azerbaijan. **(E)** During the Middle-to-Late Bronze Age peak, Armenia had more EHG ancestry than its neighbors in West Asia (Anatolia, the Levant, and Iran). **(F)** C<sup>14</sup>-dated Bronze-to-Iron Age individuals from Armenia admixed 52.2 ± 8.0 generations (1460 ± 224 years) before their average date of 1119 BCE, or ~2579 BCE (mid-3rd millennium BCE), assuming a generation length of 28 years (54). We use Early Bronze Age Armenia and Yamnaya cluster individuals from Russia as proxy sources.

for a link between the Yamnaya and populations of Armenia after this 3rd millennium BCE reappearance of Eastern hunter-gatherer ancestry.

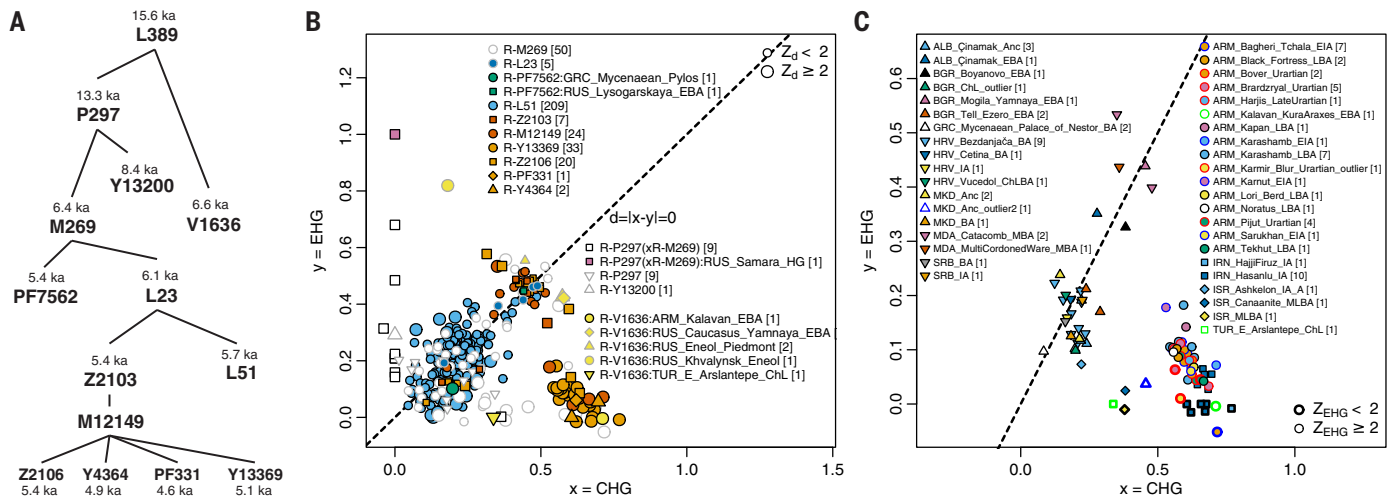
**Y-chromosome links between the steppe and West Asia in their genome-wide context**

Y-chromosome variation (tables S29 to S34 and figs. S77 and S79) (7) can be used to pro-

vide confident upper bounds on the date when two populations shared ancestors because the large number of mutations that can be analyzed over almost 10 million nucleotides of alignable sequence means that the split times in the genealogy are accurately known. The ancient individuals' Y-chromosome analysis also has the potential to provide insight into social processes.

Subclades of Y-chromosome haplogroup R-L389 are particularly informative for tracing connections between the Southern Arc and the Eurasian steppe (Fig. 6). First, haplogroup R-V1636, with an inferred common ancestor in the 5th millennium BCE, documents gene flow between the steppe and the Southern Arc in the Eneolithic/Chalcolithic period (Fig. 6B). R-V1636 is present in two individuals

Downloaded from https://www.science.org on September 13, 2023



**Fig. 6. Y-chromosome links between the Southern Arc and the Eurasian steppe.** (A) Phylogeny of haplogroup R-L389 (R1b1a1) with TMRCA estimates of yfull. com. (B) CHG/EHG ancestral composition of R-L389 Y-chromosome individuals. (C) R-L389 individuals from the Southern Arc, representing a subset of the individuals plotted in (B). Individuals >2000 years old are shown. ka, thousand years ago.

from the Late Chalcolithic at Arslantepe (Turkey) (14) and the Early Bronze Age in Armenia at Kalavan (10). It is also found in the piedmont of the North Caucasus at Progress-2 (17), the open steppe at Khvalynsk II (9), and the Single Grave Culture of Northern Europe (Gjerrild) (33). The individuals from Armenia and Arslantepe lack any detectible Eastern hunter-gatherer autosomal ancestry (Fig. 6C), which is maximized in the Khvalynsk individuals, an observation that provides some evidence for a southern origin for the R-V1636 haplogroup (we caution, however, that the haplogroup occurs earlier in several sites in the north, which could be consistent with an alternative scenario in which male migrants from the steppe introduced it into Southern Arc populations during the Chalcolithic, but their autosomal genetic legacy was diluted by the much more numerous locals). The earliest individuals from the R-L389 clade belong to the R-P297 sister clade of R-V1636, including the hunter-gatherer from Lebyazhinka IV (8, 9) and hunter-gatherers from the Baltic region (3), both without Caucasus hunter-gatherer ancestry, suggesting an Eastern European origin of this clade that would eventually give rise to the R-M269 clade that spread extremely widely in the Bronze Age.

Haplogroup R-M269, which is inferred to have a shared common ancestor in the 5th millennium BCE, is crucial for understanding steppe expansions because it was the dominant lineage of the Yamnaya-Afnasiev group (4, 8, 34) in its 4th millennium BCE R-Z2103→R-M12149 sublineage. In the Balkans, a group of six Bronze Age individuals from the 3rd millennium BCE carrying R-M269 (Fig. 6C) are associated with >30% Eastern hunter-gatherer ancestry, and this includes not only Catacomb and Multi-cordoned Ware individuals from Moldova, adjacent to the steppe, but also

from farther south, including two Yamnaya males from Bulgaria (Boyanovo and Mogila, the latter associated with Yamnaya burial custom and with the R-Z2103 haplogroup typical of the steppe Yamnaya) and one from Albania (Çinamak) belonging to the high-steppe ancestry group. By the Late Bronze Age (late 2nd millennium BCE) and later, no high-steppe ancestry individuals are observed, but steppe-associated Y chromosomes persist, including R-Z2106, a lineage that links North Macedonia (Ulanci-Veles), Albania (Çinamak), the steppe, and Armenia. The population of Southeastern Europe contrasts strongly with those of the Central/Northern Europe and Eurasian steppe archaeological cultures of ~3000 to 2000 BCE that were strongly associated with particular Y-chromosome lineages: Afanasievo (4, 34) with the same R-Z2103 as the Yamnaya, Corded Ware/Fatyanovo/Sintashta (4, 8, 34, 35) with R-M417, and Beaker (36) with R-L51. In Southeastern Europe during the Bronze Age, we detect 32/30/21/11 Y chromosomes belonging to haplogroups R/J/I/G linking it with Central/Northern Europe and the steppe/West Asia/local hunter-gatherers/Anatolian-European Neolithic farmers, respectively. Together with the extraordinary heterogeneity in autosomal ancestry in the Balkans, a picture emerges of a fragmented genetic landscape that may well parallel the poorly understood linguistic diversity in the ancient Balkans, which among Indo-European languages includes Paleo-Balkan speakers before the spread of Latin and Slavic, with Albanian as the only surviving representative. Did the early Indo-European language become successful in Southeastern Europe because it functioned as a “lingua franca,” facilitating communication among speakers of the diverse languages of previous farmer and hunter-gatherer populations?

Our newly reported data reveal that a large proportion of individuals in Armenia and Northwest Iran belonged to the R-Z2103→R-M12149 haplogroup during the 2nd and early 1st millennium BCE, providing a genetic link with the Yamnaya in these regions where no archaeological presence of the Yamnaya culture itself is attested. It definitely represents a more direct link than either R-V1636 or the early appearance of Eastern hunter-gatherer ancestry at Areni-1 cave in Armenia (10) during the Chalcolithic at the end of the 5th millennium BCE, which provides evidence of converse movement of Caucasus hunter-gatherer ancestry into the steppe Eneolithic.

Despite the Y-chromosome movement southward attested by our data, any association between R-haplogroup bearers and Eastern hunter-gatherer ancestry was lost south of the steppe because these had similar proportions of Eastern hunter-gatherer ancestry as I-Y16419 bearers (the second most prevalent lineage in Armenia). Two Bronze-to-Iron Age sites with substantial sample sizes [unrelated males from Bagheri Tchala ( $n = 7$ ) and Noratus ( $n = 12$ )] have contrasting haplogroup distributions dominated by R-M12149 and I-Y16419, respectively (Fisher's exact test  $P < 0.001$ ), suggesting founder events, high genetic drift, or a patrilineal mating system ~1000 BCE in Armenia. During the same period at Hasanlu in Northwest Iran, many individuals have no trace of Eastern hunter-gatherer ancestry at all despite the presence of R-M12149 there (6), suggesting that the initial association of this lineage with Eastern hunter-gatherer ancestry on the steppe had vanished as R-M12149 bearers reproduced with Southern Arc individuals without Eastern hunter-gatherer ancestry (Fig. 6C).

We observe that, on the steppe, R-M12149 Y chromosomes (within haplogroup R1b) at

the beginning of the 3rd millennium BCE, associated with the Yamnaya, were replaced by the beginning of the next millennium by R-Z93 Y chromosomes (within haplogroup R1a), associated with Corded Ware/Fatiano (35) steppe descendants such as those of the Sintashta culture (34). Genetic data cannot distinguish whether this Y-chromosome replacement was the result of competition between patrilineal groups from the steppe, one of which may have had cultural adaptations such as usage of an improved variety of domesticated horse (37), or whether one group simply filled an ecological niche vacated by earlier groups. A fuller understanding of the reason for this profound genetic change requires combined analysis of genetic and archaeological data.

Whatever the reason for their demise on the steppe itself, the Yamnaya-descended R-Z2103 patrilineages survived in Armenia down to the present day, where this clade is present in appreciable frequencies in all studied Armenian groups (38) despite the substantial dilution of autosomal steppe ancestry documented in our study. The persistent and lasting presence of Yamnaya patrilineal descendants in Armenia contrasts with mainland Europe and South Asia, where steppe ancestry was introduced by people who were not patrilineal descendants of the dominant R-M12149 lineage of the Yamnaya population. Instead, they belonged to different descent groups who had received autosomal steppe admixture while carrying different predominant Y-chromosome lineages. Armenia also contrasts with Anatolia, for which no R-M269 Y-chromosomes are observed at all during the Chalcolithic, Bronze Age, or Ancient (pre-Roman) periods [ $n = 80$  unrelated individuals; 95% confidence interval (CI): 0 to 4.5%] and in which haplogroups J (36 individuals) and G (17 individuals) are most common. Haplogroup J is still common at a frequency of about one-third in present-day people from Turkey (39), having achieved such prominence despite occurring in only one in 18 Neolithic male individuals from Barcin and Ilpınar in the Marmara region during the pre-Chalcolithic period. A likely explanation for the haplogroup J increase is that it accompanied the spread of Caucasus hunter-gatherer ancestry inferred by our admixture analysis (Fig. 2). This inference is made plausible by the fact that both Caucasus hunter-gatherer individuals from Kotias and Satsurblia (7) and a Mesolithic individual from Hotu Cave (10, 34) in Iran belonged to this lineage, suggesting its very old presence in the Caucasus/Iran region, and in contrast with haplogroup G, which occurred in the majority (10/18) of individuals from the Neolithic Marmara region. By the Chalcolithic, haplogroups G and J were ubiquitous in Anatolia, each making up 10/28 males from that period, paralleling the homogenization that had occurred by that time.

### The Indo-Hittite hypothesis in the light of genetic data

We discuss the implications of our genetic findings for hypotheses about the origins and spread of Indo-European and Anatolian languages. We also highlight a caveat: In contrast to findings about movements of people, the relevance of genetics to debates about language origins is more indirect because languages can be replaced with little or no genetic change and populations can migrate and mix with little or no linguistic change. Nevertheless, the detection of migration is important because it identifies a plausible vector for language change (40).

The discoveries of massive migrations from the steppe both westward into Central and Western Europe (4, 8), and eastward into South Siberia (4) and Central/South Asia (34), have provided powerful evidence for the theory of steppe Indo-European origins by linking populations all the way from Northwest Europe (36) to India and China through common steppe ancestry. The present study adds further support to the theory by the discovery of ubiquitous ancestry from the steppe in the Bronze Age Balkans [where, indubitably, Indo-European Paleo-Balkan languages such as Thracian and Illyrian (41) were spoken], including individuals of predominantly steppe ancestry; by documenting the ubiquity of steppe ancestry in Bronze and Iron Age Armenia where Armenian is first attested and links between Armenia, the steppe, and the Balkans; and by the further documentation of steppe ancestry in the Aegean (6) during the Mycenaean period when the Greek language is first attested, albeit at lower levels. All ancient and present-day branches of the Indo-European language family can be derived or at least linked to the early Bronze Age Yamnaya pastoralists of the steppe or genetically similar populations.

A link to the steppe cannot be established for the speakers of Anatolian languages because of the absence of Eastern hunter-gatherer ancestry in Anatolia (4, 10, 14, 16), which our study reinforces in three ways: (i) by documenting its paucity in ~100 new Anatolian individuals from the Chalcolithic to pre-Roman antiquity, (ii) by contrasting western parts of Anatolia with its immediate Aegean-Balkan neighbors to the west, and (iii) by contrasting eastern/northern parts of Anatolia with its neighbors in Armenia in the east. Certainly, the absence of Eastern hunter-gatherer ancestry in Anatolia can never be categorically proven (because more sampling can always disclose some such ancestry); however, at present, and despite extensive sampling, such ancestry is not detected either at possible entry points (west and east by land or even north by sea) or in the population as a whole.

The Indo-Hittite hypothesis, first proposed by E. H. Sturtevant in 1926 (42), has been partially

supported by more modern phylolinguistic analyses, indicating that Anatolian languages such as Hittite are basal to the rest of the Indo-European family tree (43) and suggesting an early split between the two. We have shown that Anatolia was indeed transformed by the Late Chalcolithic through the spread of Caucasus hunter-gatherer-related ancestry to its westernmost edges, as were apparently Eneolithic populations of the steppe, which included also Anatolian/Levantine-related ancestry by the time of the formation of the Yamnaya pastoralists. It is premature to identify the proximate sources of these movements before all the candidate source populations of Anatolia, North Mesopotamia, Western Iran, Armenia, Azerbaijan, and the Caucasus have been adequately sampled.

Our analyses show that there were at least two gene flows from two groups related to West Asians into the steppe, which transformed the steppe's population and may have induced linguistic change there. The reverse movement is more tentative, with early influences from the north such as at Areni Cave (10) or possibly associated with R-V1636 Y-chromosomes, not making a sizable genetic impact on the population of Anatolia. The evidence is consistent with two hypotheses.

Hypothesis A postulates that Proto-Indo-Anatolian (including both Anatolian languages and Proto-Indo-European) was spoken by a population with high Eastern hunter-gatherer ancestry that had a disproportionate linguistic impact on Anatolia while contributing little if any ancestry. In the post-Bronze Age landscape of Anatolia, we do find outliers marked by European or steppe influence (6), but this is a period when Anatolia is influenced by numerous linguistically non-Anatolian Indo-European populations, including Phrygians, Greeks, Persians, Galatians, and Romans, to name only a few. However, in individuals from Gordion, a Central Anatolian city that was under the control of Hittites before becoming the Phrygian capital and then coming under the control of Persian and Hellenistic rulers, the proportion of Eastern hunter-gatherer ancestry is only ~2%, a tiny fraction for a region controlled by at least four different Indo-European-speaking groups. In medieval times, Central Asian ancestry associated with Turkic speakers was added (6), and it persists to the present. Clearly, Anatolia has not been impervious to linguistic change during its recorded history, and the harbingers of that change are also detected genetically, even if as outliers. By contrast, the complete absence of Eastern hunter-gatherer ancestry in the Chalcolithic and Bronze Age either as isolated outliers or as a general low-level presence challenges the steppe theory to suggest a plausible mechanism of how a population that made little, if any, genetic impact could nonetheless effect

large-scale linguistic change. A common vocabulary for wheeled vehicles is not attested for both Anatolian languages and the rest of the Indo-European languages (44), thus potentially removing a technological advantage regarded as potentially crucial in the dissemination of Indo-European languages (45).

Hypothesis B postulates that Proto-Indo-Anatolian was spoken by a population of West Asia and the Caucasus with low or no Eastern hunter-gatherer ancestry, which affected both Anatolia and the steppe. Hypothesis B may help to explain the linguistic diversity observed in Bronze Age Anatolia in which both Anatolian (Hittite, Luwian, and Palaic) speakers, as well as speakers of other languages including Hattic (a non-Indo-European linguistic isolate of Central/Northern Anatolia) and Hurrian [a non-Indo-European language from Eastern Anatolia and North Mesopotamia related to the later Iron Age Urartian language (6)], coexisted. The non-Indo-European Hattic language, attested only in Anatolia, would most economically represent the linguistic substratum, spoken by a population of high Anatolian-related ancestry, whereas the Indo-European Anatolian languages would be spoken by a population of high Caucasus hunter-gatherer-related ancestry. The spread of people of high Caucasus hunter-gatherer ancestry across the peninsula from the east, at least some of whom may have spoken early forms of Anatolian languages, would simultaneously explain both the genetic homogenization before the Late Chalcolithic (Fig. 2) and the coexistence of the two linguistic groups. How many of the peoples associated with the spread of Caucasus hunter-gatherer ancestry spoke Anatolian languages? People speaking other languages related to the diverse non-Indo-European language families of the Caucasus, such as Kartvelian and Northwest/Northeast Caucasian, may have also participated in the westward movements.

As for the steppe, at least two streams of migration from the south (Eneolithic and Yamnaya-specific) present the opportunity for an early (Chalcolithic) split of Yamnaya linguistic ancestors from the Anatolian linguistic ancestors, followed 1000 to 2000 years later by the dispersal of Indo-European languages from the steppe with the expansion of the Yamnaya culture. Linguistic borrowings (46) between Proto-Indo-European and other language families such as Kartvelian (spoken primarily in Georgia) could be useful for localizing the Proto-Indo-Anatolian homeland, but these may have alternatively come about by long-range mobility since the Chalcolithic, proven by such evidence as the presence of R-VI636 descendants ~3000 km apart from Khvalynsk to Anatolia during this period. Contributions of Indo-European to Uralic languages (spoken in the forest zone of Eastern Europe and Siberia)

appear to have involved only Indo-Iranian speakers ~4200 years ago (47). This is important because it constrains the migratory history of Proto-Indo-Iranian, consistent with genetic evidence (34) that it spread through the steppe to South Asia and ruling out the possibility that it spread from West Asia to South Asia over the Iranian plateau. However, the contribution of Indo-Iranian to Uralic languages does not shed light on the deeper question of early Indo-Anatolian origins. A challenge for the theory that Proto-Indo-Anatolian was formed in the south in a Caucasus hunter-gatherer-rich population will be to trace the origins of the autosomal ancestry of the Yamnaya in the Caucasus or West Asia [where some existing proposals place the Proto-Indo-Anatolian homeland (32, 48, 49)] and to identify the place from which the R-M269 ancestral lineage expanded, because this will be a most plausible secondary homeland of Indo-European expansion outside of Anatolia.

The scenario of a West Asian source of Proto-Indo-Anatolian is consistent with a linguistic analysis (50) that places the split of Tocharian from the remaining (Inner Indo-European) languages ~3000 BCE associated with the Yamnaya expansion and the disintegration of the remaining languages during the 3rd millennium BCE, consistent with our inferences of major steppe admixture into the Balkans and Armenia for the subset of Indo-European languages of these regions. The Anatolian split is placed by that study at ~3700 BCE (4314 to 3450 BCE, 95% highest posterior density interval), a period during which the Caucasus hunter-gatherer ancestry first appears as far west as the Chalcolithic individuals from Northwest Anatolia (at Ilpınar) sampled in our study and during which the flow of Caucasus hunter-gatherer ancestry into the steppe had already commenced.

Overall, we suggest that a scenario in which Anatolian and Indo-European languages are descended from a common West Asian progenitor matches the evidence of population change provided by ancient DNA for four reasons. First, the genetic transformation of Anatolia after the Neolithic and before the Late Chalcolithic (Fig. 2) was a clear opportunity for linguistic spread resulting in the coexistence of Hattic and Anatolian languages. Second, the two transformations of steppe populations during the Eneolithic and before the Bronze Age, with their strong south-north directionality (Fig. 3), were opportunities for linguistic spread and match exactly the Anatolia/Indo-European split inferred by linguists. Third, steppe migrations into regions where Indo-European daughter languages were spoken, such as the Balkans (Fig. 4), Armenia (Fig. 5), Central/Northern Europe (4, 8, 36), and Central/South Asia (4, 34), were clear opportunities for the disintegration of

Proto-Indo-European and the dispersal of its daughter languages across Eurasia. Fourth, the absence of such migrations into Anatolia (Fig. 2F), in contrast to both neighboring Armenia and Southeastern Europe [Figs. 4 and 5 and (6)], makes Anatolia the only exception in the association of steppe ancestry with Indo-Anatolian languages.

This outline of events points toward a concrete research program of investigating the archaeological cultures of West Asia, the Caucasus, and the Eurasian steppe to identify a population driving transformations of both the steppe and Anatolia, linking the two regions. The discovery of such a “missing link” (corresponding to Proto-Indo-Anatolians if our reconstruction is correct) would bring to an end the centuries-old quest for a common source binding through language and some ancestry many of the peoples of Asia and Europe (41, 51).

## REFERENCES AND NOTES

- Detailed information is provided in the supplementary materials.
- F. Broushaki *et al.*, Early Neolithic genomes from the eastern Fertile Crescent. *Science* **353**, 499–503 (2016). doi: [10.1126/science.aaf7943](https://doi.org/10.1126/science.aaf7943); pmid: 27417496
- I. Mathieson *et al.*, The genomic history of Southeastern Europe. *Nature* **555**, 197–203 (2018). doi: [10.1038/nature25778](https://doi.org/10.1038/nature25778); pmid: 29466330
- M. E. Allentoft *et al.*, Population genomics of Bronze Age Eurasia. *Nature* **522**, 167–172 (2015). doi: [10.1038/nature14507](https://doi.org/10.1038/nature14507); pmid: 26062507
- S. Eisenmann *et al.*, Reconciling material cultures in archaeology with genetic data: The nomenclature of clusters emerging from archaeogenomic analysis. *Sci. Rep.* **8**, 13003 (2018). doi: [10.1038/s41598-018-31123-z](https://doi.org/10.1038/s41598-018-31123-z); pmid: 30158639
- I. Lazaridis *et al.*, A genetic probe into the ancient and medieval history of Southern Europe and West Asia. *Science* **377**, 940–951 (2022).
- E. R. Jones *et al.*, Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat. Commun.* **6**, 8912 (2015). doi: [10.1038/ncomms9912](https://doi.org/10.1038/ncomms9912); pmid: 26567969
- W. Haak *et al.*, Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015). doi: [10.1038/nature14317](https://doi.org/10.1038/nature14317); pmid: 25731166
- I. Mathieson *et al.*, Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503 (2015). doi: [10.1038/nature16152](https://doi.org/10.1038/nature16152); pmid: 26595274
- I. Lazaridis *et al.*, Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419–424 (2016). doi: [10.1038/nature19310](https://doi.org/10.1038/nature19310); pmid: 27459054
- I. Lazaridis *et al.*, Ancient DNA from Mesopotamia suggests distinct Pre-Pottery and Pottery Neolithic migrations into Anatolia. *Science* **377**, 982–987 (2022).
- R. Yaka *et al.*, Variable kinship patterns in Neolithic Anatolia revealed by ancient genomes. *Curr. Biol.* **31**, 2455–2468.e18 (2021). doi: [10.1016/j.cub.2021.03.050](https://doi.org/10.1016/j.cub.2021.03.050); pmid: 33857427
- G. M. Kiliç *et al.*, The demographic development of the first farmers in Anatolia. *Curr. Biol.* **26**, 2659–2666 (2016). doi: [10.1016/j.cub.2016.07.057](https://doi.org/10.1016/j.cub.2016.07.057); pmid: 27498567
- E. Skourtanioti *et al.*, Genomic history of Neolithic to Bronze Age Anatolia, Northern Levant, and Southern Caucasus. *Cell* **181**, 1158–1175.e28 (2020). doi: [10.1016/j.cell.2020.04.044](https://doi.org/10.1016/j.cell.2020.04.044); pmid: 32470401
- P. de Barros Damgaard *et al.*, The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* **360**, earr7711 (2018). doi: [10.1126/science.aar7711](https://doi.org/10.1126/science.aar7711); pmid: 29743352
- I. Lazaridis *et al.*, Genetic origins of the Minoans and Mycenaeans. *Nature* **548**, 214–218 (2017). doi: [10.1038/nature23310](https://doi.org/10.1038/nature23310); pmid: 28783727
- C.-C. Wang *et al.*, Ancient human genome-wide data from a 3000-year interval in the Caucasus corresponds with ecogeographic regions. *Nat. Commun.* **10**, 590 (2019). doi: [10.1038/s41467-018-08220-8](https://doi.org/10.1038/s41467-018-08220-8); pmid: 30713341

18. K. Kristiansen, "The archaeology of Proto-Indo-European and Proto-Anatolian: Locating the split," in *Dispersals and Synchronisation: Linguistic and Archaeological Perspectives on the Early Stages of Indo-European*, M. Serangeli, T. Olander, Eds. (Brill, 2019), pp. 157–165.
19. M. Chintalapati, N. Patterson, P. Moorjani, Reconstructing the spatiotemporal patterns of admixture during the European Holocene using a novel genomic dating method. *bioRxiv* 2022.2001.2018.476710 (2022). doi: [10.1101/2022.01.18.476710](https://doi.org/10.1101/2022.01.18.476710)
20. D. M. Fernandes *et al.*, The spread of steppe and Iranian-related ancestry in the islands of the Western Mediterranean. *Nat. Ecol. Evol.* **4**, 334–345 (2020). doi: [10.1038/s41559-020-1102-0](https://doi.org/10.1038/s41559-020-1102-0); pmid: 32094539
21. V. Villalba-Mouco *et al.*, Genomic transformation and social organization during the Copper Age-Bronze Age transition in southern Iberia. *Sci. Adv.* **7**, eabi7038 (2021). doi: [10.1126/sciadv.abi7038](https://doi.org/10.1126/sciadv.abi7038); pmid: 34788096
22. L. Papac *et al.*, Dynamic changes in genomic and social structures in third millennium BCE central Europe. *Sci. Adv.* **7**, eabi6941 (2021). doi: [10.1126/sciadv.abi6941](https://doi.org/10.1126/sciadv.abi6941); pmid: 34433570
23. S. Brunel *et al.*, Ancient genomes from present-day France unveil 7,000 years of its demographic history. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 12791–12798 (2020). doi: [10.1073/pnas.1918034117](https://doi.org/10.1073/pnas.1918034117); pmid: 32457149
24. R. Fregel *et al.*, Ancient genomes from North Africa evidence prehistoric migrations to the Maghreb from both the Levant and Europe. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 6774–6779 (2018). doi: [10.1073/pnas.1800851115](https://doi.org/10.1073/pnas.1800851115); pmid: 29895688
25. I. Lazaridis *et al.*, Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413 (2014). doi: [10.1038/nature13673](https://doi.org/10.1038/nature13673); pmid: 25230663
26. M. Lipson *et al.*, Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* **551**, 368–372 (2017). doi: [10.1038/nature24476](https://doi.org/10.1038/nature24476); pmid: 29144465
27. I. Olalde *et al.*, A common genetic origin for early farmers from Mediterranean Cardial and Central European LBK Cultures. *Mol. Biol. Evol.* **32**, 3132–3142 (2015). doi: [10.1093/molbev/msv181](https://doi.org/10.1093/molbev/msv181); pmid: 26337550
28. M. Rivollat *et al.*, Ancient genome-wide DNA from France highlights the complexity of interactions between Mesolithic hunter-gatherers and Neolithic farmers. *Sci. Adv.* **6**, eaaz5344 (2020). doi: [10.1126/sciadv.aaz5344](https://doi.org/10.1126/sciadv.aaz5344); pmid: 32523989
29. C. Valdiosera *et al.*, Four millennia of Iberian biomolecular prehistory illustrate the impact of prehistoric migrations at the far end of Eurasia. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 3428–3433 (2018). doi: [10.1073/pnas.1717762115](https://doi.org/10.1073/pnas.1717762115); pmid: 29531053
30. V. Villalba-Mouco *et al.*, Survival of Late Pleistocene hunter-gatherer ancestry in the Iberian Peninsula. *Curr. Biol.* **29**, 1169–1177.e7 (2019). doi: [10.1016/j.cub.2019.02.006](https://doi.org/10.1016/j.cub.2019.02.006); pmid: 30880015
31. F. Clemente *et al.*, The genomic history of the Aegean palatial civilizations. *Cell* **184**, 2565–2586.e21 (2021). doi: [10.1016/j.cell.2021.03.039](https://doi.org/10.1016/j.cell.2021.03.039); pmid: 33930288
32. C. Renfrew, *Archaeology and Language: The Puzzle of Indo-European Origins* (CUP Archive, 1990).
33. A. F.-H. Egjford *et al.*, Genomic Steppe ancestry in skeletons from the Neolithic Single Grave Culture in Denmark. *PLoS ONE* **16**, e0244872 (2021). doi: [10.1371/journal.pone.0244872](https://doi.org/10.1371/journal.pone.0244872); pmid: 33444387
34. V. M. Narasimhan *et al.*, The formation of human populations in South and Central Asia. *Science* **365**, eaat7487 (2019). doi: [10.1126/science.aat7487](https://doi.org/10.1126/science.aat7487); pmid: 31488661
35. L. Saag *et al.*, Genetic ancestry changes in Stone to Bronze Age transition in the East European plain. *Sci. Adv.* **7**, eabd6535 (2021). doi: [10.1126/sciadv.abd6535](https://doi.org/10.1126/sciadv.abd6535); pmid: 33523926
36. I. Olalde *et al.*, The Beaker phenomenon and the genomic transformation of Northwest Europe. *Nature* **555**, 190–196 (2018). doi: [10.1038/nature25738](https://doi.org/10.1038/nature25738); pmid: 29466337
37. P. Librado *et al.*, The origins and spread of domestic horses from the Western Eurasian steppes. *Nature* **598**, 634–640 (2021). doi: [10.1038/s41586-021-04018-9](https://doi.org/10.1038/s41586-021-04018-9); pmid: 34671162
38. O. Balanovsky *et al.*, Genetic differentiation between upland and lowland populations shapes the Y-chromosomal landscape of West Asia. *Hum. Genet.* **136**, 437–450 (2017). doi: [10.1007/s00439-017-1770-2](https://doi.org/10.1007/s00439-017-1770-2); pmid: 28281087
39. C. Cinnioglu *et al.*, Excavating Y-chromosome haplotype strata in Anatolia. *Hum. Genet.* **114**, 127–148 (2004). doi: [10.1007/s00439-003-1031-4](https://doi.org/10.1007/s00439-003-1031-4); pmid: 14586639
40. P. S. Bellwood, *The Encyclopedia of Global Human Migration* (Wiley, 2013), vol. 1, *Prehistory*.
41. J. P. Mallory, *In Search of the Indo-Europeans: Language, Archaeology, and Myth* (Thames and Hudson, 1989).
42. E. H. Sturtevant, On the position of Hittite among the Indo-European languages. *Language* **2**, 25–34 (1926). doi: [10.2307/408784](https://doi.org/10.2307/408784)
43. D. Ringe, T. Warnow, A. Taylor, Indo-European and computational cladistics. *Trans. Philol. Soc.* **100**, 59–129 (2002). doi: [10.1111/1467-968X.00091](https://doi.org/10.1111/1467-968X.00091)
44. M. J. Kümmel, "Wheel and chariot in early IE: What exactly can we conclude from the linguistic data?" paper presented at the 25th Annual Meeting of the European Association of Archaeologists, Bern, Switzerland, 4–7 September 2019 (abstract #374).
45. D. W. Anthony, *The Horse, the Wheel, and Language: How Bronze-Age Riders from the Eurasian Steppes Shaped the Modern World* (Princeton Univ. Press, 2010).
46. D. W. Anthony, D. Ringe, The Indo-European homeland from linguistic and archaeological perspectives. *Annu. Rev. Linguist.* **1**, 199–219 (2015). doi: [10.1146/annurev-linguist-030514-124812](https://doi.org/10.1146/annurev-linguist-030514-124812)
47. J. Nichols, The origin and dispersal of Uralic: Distributional typological view. *Annu. Rev. Linguist.* **7**, 351–369 (2021). doi: [10.1146/annurev-linguistics-011619-030405](https://doi.org/10.1146/annurev-linguistics-011619-030405)
48. T. V. Gamkrelidze, V. V. Ivanov, R. Jakobson, N. Johanna, *Indo-European and the Indo-Europeans, A Reconstruction and Historical Analysis of a Proto-Language and Proto-Culture* (Mouton de Gruyter, 1995).
49. S. A. Grigoriev, *Ancient Indo-Europeans* (RIFEI, 2002).
50. 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). doi: [10.1015/ling-2020-0060](https://doi.org/10.1015/ling-2020-0060)
51. W. Jones, *Discourses Delivered Before the Asiatic Society: And Miscellaneous Papers, on the Religion, Poetry, Literature, Etc., of the Nations of India* (C. S. Arnold, 1824).
52. S. Alpaslan-Roodenberg *et al.*, Ethics of DNA research on human remains: Five globally applicable guidelines. *Nature* **599**, 41–46 (2021). doi: [10.1038/s41586-021-04008-x](https://doi.org/10.1038/s41586-021-04008-x); pmid: 34671160
53. M. Binois, R. B. Gramacy, hetgp: Heteroskedastic Gaussian process modeling and sequential design in R. *J. Stat. Softw.* **98**, 1–44 (2021). doi: [10.18637/jss.v098.i13](https://doi.org/10.18637/jss.v098.i13)
54. 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. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 5652–5657 (2016). doi: [10.1073/pnas.1514696113](https://doi.org/10.1073/pnas.1514696113); pmid: 27140627

#### ACKNOWLEDGMENTS

This study was performed following the principles for ethical DNA research on human remains described in (52). We are grateful to the authorities and sample stewards, including museums, museum curators, and archaeologists, for providing written permission to sample each human remain. We acknowledge the ancient individuals whose genetic data we analyzed and whose permission we could not directly ask. We aimed to write a manuscript that was respectful of the ancient individuals, treating samples from them as derived from real people whose memories must be respected. We sought to reflect the perspectives of people from the diverse geographic regions and cultural contexts from which the sampled individuals came by having each sample be represented by at least one coauthor who was a sample steward and was part of a network engaged with local communities. We thank J. Bennett, V. Narasimhan, H. Ringbauer, J. Sedig, A. Shaus, L. Vokotopoulos, M. Wiener, and several anonymous reviewers for critical comments; D. Mitrevski and M. Pantelidou-Gofa for archaeological work; G. Rollefson for support for publishing additional data from 'Ain Ghazal and advising on archaeological contextualization; Y. Urasin for technical help with the yfll.com phylogeny; and N. Adamski for lab work. Most of the samples from Albania (sites of Podgorje, Tren Cave 2, Dukat, Çinamak, Kënetë, Bardhok, Shtikë, Barç, and Pazhok) were included in this study as part of the project "The paleogenetics of Southeastern Europeans, admixture, selection and transformations: The case of Albania," a joint collaboration between the Albanian Institute of Archaeology (Tirana) and the Anthropology Department of the University of Vienna (principal investigators R.P., R.K., and R.R.). We thank the National Museum of Bosnia and Herzegovina and its staff. We honor the memories of G. Areshian, M. Bilbija, and E. Peltenburg, who would have been our coauthors if they had not passed away during the course of this study. **Funding:** The computational analysis and ancient DNA data generation for this study were supported by the National Institutes of Health (National

Institute of General Medical Sciences grant GM100233 and National Human Genome Research Grant HG012287); the John Templeton Foundation (grant 61220); a private gift from Jean-Francois Clin; the Allen Discovery Center program, a Paul G. Allen Frontiers Group advised program of the Paul G. Allen Family Foundation; the Howard Hughes Medical Institute (D.R.); the Ministry of Science and Innovation of the Spanish Government (RYC2019-027909-1/AEI/10.13039/501100011033) and Ikerbasque-Basque Foundation of Science grants (I.O.). The archaeological work was supported by the NOMIS Foundation (D.B.); the European Research Council (ERC Starting Grant Project HIDDEN FOODS 639286 to E.Cr.); the Romanian Ministry of Research, Innovation, and Digitization (CNCS, CNFIS, CCCDI – UEFISCDI project numbers 351PED PN-III-P2-1-PED-2019-4171 and CNFIS-FDI-2021-0405 D6/ 2021 within PNCDI III to C.L.); the Face to Face: Meet an Ancient Cypriot project (FF-MAC project INTEGRATED/0609/29); the BioMERA project (Platform for Biosciences and Human Health in Cyprus: MicroCT and Synchrotron Radiation Enabled Analyses; grant INFRASTRUCTURES/1216/09) cofinanced by the European Regional Development Fund and the Republic of Cyprus through the Research and Innovation Foundation (KOL); the Hungarian Research, Development and Innovation Office (grant FK228013 to T.H., T.Sz., and K.K.); the Hungarian Academy of Science (Bolyai Scholarship to T.H.); the Croatian Science Foundation (grant HRZZ IP-2016-06-1450 to M.N., I.J., and J.B. and grant NCN 2015/17/B/HS3/01327 to P.W.); and the Bursa Uludağ University (Turkey) General Research Project (grant SGA-2021-389, project title "Early Christian martyriums in the light of the Basilica Church of the Lake of Iznik," to M.Şa.). **Author contributions:** Conceived of the study: I.La., S.A.-R., R.P., and D.R. Supervised the study: S.A.-R., D.J.K., N.Pat., N.R., R.P., and D.R. Assembled archaeological material and prepared the site descriptions: S.A.-R., A.Aca., A.Aci., A.Ag., L.A., U.A., D.And., G.A., D.Ant., I.A., A.At., P.A., A.I.A., K.Ba., R.Ba., J.B., L.B., A.Be., H.B., A.Bi., M.Bod., M.Bon., C.B., D.B., N.B., M.Ca., S.Cho., M.-E.C., S.Chr., I.C., N.C., M.Co., E.Cr., J.D., R.D., T.I.D., V.De., Z.D., S.Des., S.Dev., V.Dj., N.El., A.E., N.Er., S.E.-P., A.F., M.L.G., B.Gas., B.Gay., E.G., T.G., S.G., T.H., V.H., S.H., N.H., I.I., S.I., I.J., L.J., P.Ka., B.K.-K., E.H.K., S.D.K., A.K., K.K., S.Ki., P.Ki., S.K.B.N.V., S.Ko., M.K.-N., M.K.S., R.K., P.Ku., C.L., K.L., T.E.L., I.Li., K.O.L., S.L., K.M.-O., R.M., W.M., R.Mc., V.M., L.M., J.M.M., G.N., R.N., A.G.N., V.N., M.N., A.O., C.O., N.O., N.Papad., N.Papak., A.Pa., L.Pa., E.G.P., I.P., L.Pe., V.P., A.P.-T., A.Pi., N.P.K., H.P., B.P.-B., Z.P., T.D.P., S.Rad., K.R.A., P.R.S., K.R.R., S.Raz., A.R., J.R., R.R., V.R., M.Ş., A.Ş., E.S., A.Su., L.S., T.Se., A.S.-E., M.S.-P., H.M.S., A.Sid., A.Sim., H.S., V.S., G.S., M.S., A.Sof., B.S., A.Sof., Ç.S.-S., M.Sta., M.Ste., S.S., F.S.-A., A.S.-N., T.Sz., N.Te., S.Te., N.To., U.T., G.T., S.Tr., A.T., M.U., F.V., Z.V., C.V., S.W., P.W., A.Y., E.Y., A.Y.Y., H.Y., R.Ç., and R.P. Performed laboratory work: S.A.-R., G.B.M., K.Bu., K.C., F.C., B.J.C., E.Cu., K.S.D.C., L.R.E., D.M.F., M.F., S.F., B.Gam., L.I., D.K., A.M.L., K.Ma., M.Mi., J.O., K.T.Ö., L.Q., C.S., K.Si., K.St., A.W., J.N.W., F.Z., A.Z., and N.R. Performed population genetic analyses: I.La. and D.R. Analyzed data: I.La., S.A.-R., R.Be., O.C., M.Ma., S.M., A.Mic., A.Mit., I.O., Z.Z., N.R., and D.R. Wrote the manuscript and compiled the supplementary sections with the input of all other coauthors: I.La., S.A.-R., and D.R. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Genotype data for individuals included in this study can be obtained from the Harvard Dataverse repository through the following link (<https://doi.org/10.7910/DVN/3AR0CD>). BAM files of aligned reads can be obtained from the European Nucleotide Archive (accession no. PRJEB54831). All other data needed to evaluate the conclusions in this study are present in the main manuscript or the supplementary materials. **License information:** Copyright © 2022 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>

#### SUPPLEMENTARY MATERIALS

[science.org/doi/10.1126/science.abm4247](https://doi.org/10.1126/science.abm4247)

Materials and Methods

Supplementary Text S1 to S5

Figs. S1 to S78

Tables S1 to S34

References (55–483)

Data S1 to S5

MDAR Reproducibility Checklist

[View/request a protocol for this paper from Bio-protocol.](#)

Submitted 16 September 2021; resubmitted 17 March 2022

Accepted 21 July 2022

[10.1126/science.abm4247](https://doi.org/10.1126/science.abm4247)

## The genetic history of the Southern Arc: A bridge between West Asia and Europe

Iosif Lazaridis, Songül Alpaslan-Roodenberg, Ayşe Acar, Ayten Açıkkol, Anagnostis Agelarakis, Levon Aghikyan, Uğur Akyüz, Desislava Andreeva, Gojko Andrijašević, Dragana Antonović, Ian Armit, Alper Atmaca, Pavel Avetisyan, Ahmet Hsyan Aytekin, Krum Bacvarov, Ruben Badalyan, Stefan Bakardzhiev, Jacqueline Balen, Lorenc Bejko, Rebecca Bernardos, Andreas Bertatos, 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 Chrystoulaki, Ion Ciobanu, Natalija Čonđić, Mihai Constantinescu, Emanuela Cristiani, Brendan J. Culleton, Elizabeth Curtis, Jack Davis, Ruben Davtyan, Tatiana I. Demcenco, Valentin Dergachev, Zafer Derin, Sylvia Deskaj, Seda Devejian, Vojislav Djordjević, Kellie Sara Duffett Carlson, Laurie R. Eccles, Nedko Elenski, Atilla Engin, Nihat Erdoğan, Sabiha Erir-Pazarç, 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, İlyay İvgin, Ivor Janković, Lence Jovanova, Panagiotis Karkanias, Berna Kavaz-Kandemir, Esra Hilal Kaya, Denise Keating, Douglas J. Kennett, Seda Deniz Kesici, Anahit Khudaverdyan, Krisztián Kiss, Sinan Kömürcü, Paul Klostermann, Sinem Kostak Boca Negra Valdes, Saša Kovačević, Marta Krenz-Niedbala, Maja Krznarić Škrivanko, Rovena Kurti, Pasko Kuzman, Ann Marie Lawson, Catalin Lazar, Krassimir Leshtakov, Thomas E. Levy, Ioannis Liritzis, Kiri 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, Lidija Milašinović, 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 Papatou, Lujana Paraman, Evgeny G. Paskary, Nick Patterson, Ilian Petrakiev, Levon Petrosyan, Vanya Petrova, Anna Philippa-Touchais, Ashot Pilibosyan, Nada Pocuca Kuzman, Hrvoje Potrebica, Bianca Preda-Bonić, 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 Schatke, Lynne Schepartz, Tayfun Selçuk, Ayla Sevim-Erol, Michel Shammoun-Pour, Henry M. Shephard, Athanasios Sideris, Angela Simalcsik, Hakob Simonyan, Vitalij Sinika, Kendra Sirak, Ghénadie 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 Włodarczyk, J. Noah Workman, Aram Yardumian, Evgenii Yarovoy, Alper Yener Yavuz, Hakan Yılmaz, Fatma Zalzal, Anna Zettl, Zhao Zhang, Rafet Çavuşoğlu, Nadin Rohland, Ron Pinhasi, and David Reich

*Science*, 377 (6609), eabm4247.

DOI: 10.1126/science.abm4247

### Connecting genes and history

Stories about the peopling—and people—of Southern Europe and West Asia have been passed down for thousands of years, and these stories have contributed to our historical understanding of populations. Genomic data provide the opportunity to truly understand these patterns independently from written history. In a trio of papers, Lazaridis *et al.* examined more than 700 ancient genomes from across this region, the Southern Arc, spanning 11,000 years, from the earliest farming cultures to post-Medieval times (see the Perspective by Arbuckle and Schwandt). On the basis of these results, the authors suggest that earlier reliance on modern phenotypes and ancient writings and artistic depictions provided an inaccurate picture of early Indo-Europeans, and they provide a revised history of the complex migrations and population integrations that shaped these cultures. —SNV

Use of this article is subject to the [Terms of service](#)

*Science* (ISSN ) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2022 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works



**View the article online**

<https://www.science.org/doi/10.1126/science.abm4247>

**Permissions**

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

---

*Science* (ISSN ) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2022 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works