

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

Article

Accepted Version

Lazaridis, I. ORCID: <https://orcid.org/0000-0002-4094-9347>, Alpaslan-Roodenberg, S. ORCID: <https://orcid.org/0000-0002-1536-2931>, Acar, A. ORCID: <https://orcid.org/0000-0002-1327-6378>, Açikkol, A. ORCID: <https://orcid.org/0000-0003-1086-9749>, Agelarakis, A., Aghikyan, L. ORCID: <https://orcid.org/0000-0002-5537-4144>, Akyüz, U., Andreeva, D. ORCID: <https://orcid.org/0000-0003-2129-9016>, Andrijašević, G. ORCID: <https://orcid.org/0000-0002-9075-2112>, Antonović, D. ORCID: <https://orcid.org/0000-0002-6534-1258>, Armit, I. ORCID: <https://orcid.org/0000-0001-8669-3810>, Atmaca, A., Avetisyan, P. ORCID: <https://orcid.org/0000-0002-4469-9517>, Aytek, A. İ. ORCID: <https://orcid.org/0000-0002-7620-2333>, Bacvarov, K., Badalyan, R. ORCID: <https://orcid.org/0000-0002-0641-5341>, Bakardzhiev, S., Balen, J. ORCID: <https://orcid.org/0000-0001-7770-184X>, Bejko, L. ORCID: <https://orcid.org/0000-0003-3821-6678>, Bernardos, R. ORCID: <https://orcid.org/0000-0003-4625-3727>, Bertsatos, A. ORCID: <https://orcid.org/0000-0001-8734-5236>, Biber, H. ORCID: <https://orcid.org/0000-0003-3719-4120>, Bilir, A. ORCID: <https://orcid.org/0000-0001-7493-383X>, Bodružić, M. ORCID: <https://orcid.org/0000-0002-8112-5918>, Bonogofsky, M., Bonsall, C. ORCID: <https://orcid.org/0000-0003-0243-3232>, Borić, D. ORCID: <https://orcid.org/0000-0003-0166-627X>, Borovinić, N. ORCID: <https://orcid.org/0000->

0001-6542-6334, Bravo Morante, G., Buttinger, K., Callan, K.  
ORCID: <https://orcid.org/0000-0003-3170-8514>, Candilio, F.  
ORCID: <https://orcid.org/0000-0002-4668-1361>, Carić, M.  
ORCID: <https://orcid.org/0000-0003-4263-6528>, Cheronet, O.  
ORCID: <https://orcid.org/0000-0001-6760-1204>, Chohadzhiev,  
S., Chovalopoulou, M.-E. ORCID: <https://orcid.org/0000-0001-8187-8031>, Chryssoulaki, S., Ciobanu, I. ORCID:  
<https://orcid.org/0000-0002-9890-4010>, Čondić, N.,  
Constantinescu, M. ORCID: <https://orcid.org/0000-0002-7875-9119>, Cristiani, E. ORCID: <https://orcid.org/0000-0002-2748-9171>, Culleton, B. J. ORCID: <https://orcid.org/0000-0003-1200-9531>, Curtis, E. ORCID: <https://orcid.org/0000-0002-0263-0765>, Davis, J., Demcenco, T. I., Dergachev, V. ORCID:  
<https://orcid.org/0000-0002-6095-2199>, Derin, Z. ORCID:  
<https://orcid.org/0000-0002-0282-0593>, Deskaj, S., Devejyan,  
S., Djordjević, V., Duffett Carlson, K. S. ORCID:  
<https://orcid.org/0000-0002-7194-6608>, Eccles, L. R., Elenski,  
N., Engin, A. ORCID: <https://orcid.org/0000-0003-3165-1562>,  
Erdoğan, N., Erir-Pazarcı, S. ORCID: <https://orcid.org/0000-0003-2806-3874>,  
Fernandes, D. M. ORCID:  
<https://orcid.org/0000-0002-7434-6552>, Ferry, M. ORCID:  
<https://orcid.org/0000-0001-7319-0085>, Freilich, S. ORCID:  
<https://orcid.org/0000-0003-0472-6311>, Frînculeasa, A.  
ORCID: <https://orcid.org/0000-0002-1694-911X>, Galaty, M. L.  
ORCID: <https://orcid.org/0000-0001-7754-8347>, Gamarra, B.  
ORCID: <https://orcid.org/0000-0003-3764-497X>, Gasparyan,  
B., Gaydarska, B. ORCID: <https://orcid.org/0000-0002-6236-872X>, Genç, E. ORCID: <https://orcid.org/0000-0003-0508-7267>, Gültekin, T. ORCID: <https://orcid.org/0000-0003-3520-5308>, Gündüz, S. ORCID: <https://orcid.org/0000-0002-8923-8888>, Hajdu, T. ORCID: <https://orcid.org/0000-0002-3604-1125>, Heyd, V., Hobosyan, S., Hovhannisyan, N. ORCID:  
<https://orcid.org/0000-0002-9985-1547>, Iliev, I., Iliev, L.  
ORCID: <https://orcid.org/0000-0003-4793-9421>, İvgin,  
İ. ORCID: <https://orcid.org/0000-0001-8284-5346>, Janković, I.  
ORCID: <https://orcid.org/0000-0001-5175-6885>, Jovanova, L.,  
Karkanas, P. ORCID: <https://orcid.org/0000-0002-7156-671X>,

Kavaz-Kındığılı, B. ORCID: <https://orcid.org/0000-0002-6643-8852>, Kaya, E. H. ORCID: <https://orcid.org/0000-0001-8853-8072>, Keating, D. ORCID: <https://orcid.org/0000-0003-0900-0245>, Kennett, D. J. ORCID: <https://orcid.org/0000-0001-6144-7365>, Deniz Kesici, S. ORCID: <https://orcid.org/0000-0003-1594-3047>, Khudaverdyan, A. ORCID: <https://orcid.org/0000-0002-1458-783X>, Kiss, K. ORCID: <https://orcid.org/0000-0002-9371-3397>, Kılıç, S. ORCID: <https://orcid.org/0000-0003-3424-4081>, Klostermann, P. ORCID: <https://orcid.org/0000-0002-4668-8520>, Kostak Boca Negra Valdes, S. ORCID: <https://orcid.org/0000-0001-8333-6121>, Kovačević, S. ORCID: <https://orcid.org/0000-0002-7197-5994>, Krenz-Niedbała, M. ORCID: <https://orcid.org/0000-0001-7420-5621>, Krznarić Škrivanko, M. ORCID: <https://orcid.org/0000-0002-6409-8663>, Kurti, R., Kuzman, P. ORCID: <https://orcid.org/0000-0003-4477-289X>, Lawson, A. M. ORCID: <https://orcid.org/0000-0003-0990-2329>, Lazar, C. ORCID: <https://orcid.org/0000-0001-7559-1843>, Leshtakov, K., Levy, T. E. ORCID: <https://orcid.org/0000-0002-8618-6672>, Liritzis, I. ORCID: <https://orcid.org/0000-0003-3079-1336>, Lorentz, K. O., Łukasik, S. ORCID: <https://orcid.org/0000-0003-2210-7434>, Mah, M. ORCID: <https://orcid.org/0000-0001-8987-6436>, Mallick, S., Mandl, K. ORCID: <https://orcid.org/0000-0002-2078-0775>, Martirosyan-Olshansky, K. ORCID: <https://orcid.org/0000-0002-1896-9085>, Matthews, R. ORCID: <https://orcid.org/0000-0002-8316-4312>, Matthews, W. ORCID: <https://orcid.org/0000-0002-7408-6885>, McSweeney, K. ORCID: <https://orcid.org/0000-0001-6488-2928>, Melikyan, V., Micco, A. ORCID: <https://orcid.org/0000-0001-6016-964X>, Michel, M. ORCID: <https://orcid.org/0000-0002-5484-7974>, Milašinović, L., Mitnik, A. ORCID: <https://orcid.org/0000-0002-6963-4824>, Monge, J. M. ORCID: <https://orcid.org/0000-0002-4364-6296>, Nekhrizov, G. ORCID: <https://orcid.org/0000-0003-0933-7980>, Nicholls, R., Nikitin, A. G. ORCID: <https://orcid.org/0000-0002-3897-4607>, Nikolov, V., Novak, M. ORCID: <https://orcid.org/0000-0002-4567-8742>, Olalde, I. ORCID: <https://orcid.org/0000-0002-2660-6807>, Oppenheimer,

J. ORCID: <https://orcid.org/0000-0001-7973-6173>, Osterholtz, A. ORCID: <https://orcid.org/0000-0002-8437-9147>, Özdemir, C., Özdoğan, K. T., Öztürk, N., Papadimitriou, N. ORCID: <https://orcid.org/0000-0002-2894-7250>, Papakonstantinou, N. ORCID: <https://orcid.org/0000-0002-6869-645X>, Papathanasiou, A., Paraman, L. ORCID: <https://orcid.org/0000-0001-8780-5845>, Paskary, E. G., Patterson, N. ORCID: <https://orcid.org/0000-0003-0302-684X>, Petrakiev, I., Petrosyan, L., Petrova, V. ORCID: <https://orcid.org/0000-0001-7844-1058>, Philippa-Touchais, A., Piliposyan, A. ORCID: <https://orcid.org/0000-0001-7367-5250>, Pocuca Kuzman, N. ORCID: <https://orcid.org/0000-0003-4477-289X>, Potrelica, H. ORCID: <https://orcid.org/0000-0002-5474-4254>, Preda-Bălănică, B. ORCID: <https://orcid.org/0000-0002-2956-2336>, Premužić, Z., Price, T. D. ORCID: <https://orcid.org/0000-0001-5951-5621>, Qiu, L. ORCID: <https://orcid.org/0000-0003-4377-8659>, Radović, S. ORCID: <https://orcid.org/0000-0001-7838-7952>, Raeuf Aziz, K., Rajić Šikanjić, P. ORCID: <https://orcid.org/0000-0002-7760-5193>, Rasheed Raheem, K., Razumov, S. ORCID: <https://orcid.org/0000-0001-6030-9390>, Richardson, A. ORCID: <https://orcid.org/0000-0002-7376-8948>, Roodenberg, J., Ruka, R., Russeva, V. ORCID: <https://orcid.org/0000-0001-8998-6782>, Şahin, M. ORCID: <https://orcid.org/0000-0001-5153-1918>, Şarbak, A. ORCID: <https://orcid.org/0000-0003-3199-1612>, Savaş, E. ORCID: <https://orcid.org/0000-0002-3806-8399>, Schattke, C., Schepartz, L. ORCID: <https://orcid.org/0000-0003-1379-0817>, Selçuk, T. ORCID: <https://orcid.org/0000-0003-0322-2015>, Sevim-Erol, A. ORCID: <https://orcid.org/0000-0001-7776-3864>, Shamooun-Pour, M. ORCID: <https://orcid.org/0000-0002-2094-002X>, Shephard, H. M. ORCID: <https://orcid.org/0000-0003-0168-5462>, Sideris, A. ORCID: <https://orcid.org/0000-0003-0635-4718>, Simalcsik, A. ORCID: <https://orcid.org/0000-0001-8639-3535>, Simonyan, H., Sinika, V. ORCID: <https://orcid.org/0000-0002-1621-9205>, Sirak, K. ORCID: <https://orcid.org/0000-0003-2347-3479>, Sirbu, G. ORCID: <https://orcid.org/0000-0002-4525-1239>,

Šlaus, M. ORCID: <https://orcid.org/0000-0002-4941-2212>,  
Soficaru, A. ORCID: <https://orcid.org/0000-0001-8658-6695>,  
Söğüt, B. ORCID: <https://orcid.org/0000-0002-0171-318X>,  
Sołtysiak, A. ORCID: <https://orcid.org/0000-0002-9040-5022>,  
Sönmez-Sözer, Ç. ORCID: <https://orcid.org/0000-0002-2818-5364>,  
Stathi, M. ORCID: <https://orcid.org/0000-0002-6698-8977>,  
Steskal, M. ORCID: <https://orcid.org/0000-0002-2964-3953>,  
Stewardson, K., Stocker, S., Suata-Alpaslan, F. ORCID:  
<https://orcid.org/0000-0003-0404-9381>,  
Suvorov, A.,  
Szécsényi-Nagy, A. ORCID: <https://orcid.org/0000-0003-2095-738X>,  
Szeniczey, T. ORCID: <https://orcid.org/0000-0003-1546-7140>,  
Telnov, N. ORCID: <https://orcid.org/0000-0001-6878-0216>,  
Temov, S. ORCID: <https://orcid.org/0000-0003-0133-5526>,  
Todorova, N. ORCID: <https://orcid.org/0000-0002-7117-5452>,  
Tota, U., Touchais, G. ORCID: <https://orcid.org/0000-0002-7225-2157>,  
Triantaphyllou, S. ORCID:  
<https://orcid.org/0000-0003-1231-6515>,  
Türker, A. ORCID:  
<https://orcid.org/0000-0003-3893-4135>,  
Ugarković, M. ORCID:  
<https://orcid.org/0000-0002-1134-7531>,  
Valchev, T.,  
Veljanovska, F., Videvski, Z., Virag, C. ORCID:  
<https://orcid.org/0000-0002-7867-7498>,  
Wagner, A. ORCID:  
<https://orcid.org/0000-0003-1859-2297>,  
Walsh, S. ORCID:  
<https://orcid.org/0000-0002-6820-8516>,  
Włodarczak, P.  
ORCID: <https://orcid.org/0000-0003-0359-7386>,  
Workman, J.  
N. ORCID: <https://orcid.org/0000-0002-0921-1803>,  
Yardumian,  
A., Yarovoy, E., Yavuz, A. Y. ORCID: <https://orcid.org/0000-0002-4959-5581>,  
Yılmaz, H. ORCID: <https://orcid.org/0000-0002-2399-5085>,  
Zalzala, F. ORCID: <https://orcid.org/0000-0002-8981-1277>,  
Zettl, A. ORCID: <https://orcid.org/0000-0002-8201-3993>,  
Zhang, Z. ORCID: <https://orcid.org/0000-0002-7562-7559>,  
Çavuşoğlu, R. ORCID: <https://orcid.org/0000-0002-0049-902X>,  
Rohland, N., Pinhasi, R. ORCID:  
<https://orcid.org/0000-0003-1629-8131> and Reich, D. ORCID:  
<https://orcid.org/0000-0002-7037-5292> (2022) The genetic  
history of the Southern Arc: a bridge between West Asia and  
Europe. *Science*, 377 (6609). ISSN 1095-9203 doi:  
<https://doi.org/10.1126/science.abm4247> Available at

<https://centaur.reading.ac.uk/107141/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1126/science.abm4247>

Publisher: American Association for the Advancement of Science

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

## **CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

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

**Authors:** 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 Andrijasevic<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 AYTEK<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 Bertatos<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 Borčić<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 Chrysoulaki<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>, Tatiana I. Demcenco<sup>39</sup>, Valentin Dergachev<sup>40</sup>, Zafer Derin<sup>41</sup>, Sylvia Deskaj<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>, Atilla Engin<sup>46</sup>, Nihat Erdoğan<sup>47</sup>, Sabiha Erir-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>, Bisserka Gaydarska<sup>54</sup>, Elif Genç<sup>55</sup>, Timur Gültekin<sup>56</sup>, Serkan Gündüz<sup>57</sup>, Tamás Hajdu<sup>58,59</sup>, Volker Heyd<sup>60</sup>, Suren Hobosyan<sup>7</sup>, Nelli Hovhannisyan<sup>61</sup>, Iliya Iliev<sup>16</sup>, Lora Iliev<sup>2,27</sup>, Stanislav Iliev<sup>62</sup>, İlkey İvgin<sup>63</sup>, Ivor Janković<sup>29</sup>, Lence Jovanova<sup>64</sup>, Panagiotis Karkanias<sup>65</sup>, Berna Kavaz-Kındıgılı<sup>66</sup>, Esra Hilal Kaya<sup>67</sup>, Denise Keating<sup>3</sup>, Douglas Kennett<sup>37,68</sup>, Seda Deniz Kesici<sup>69</sup>, Anahit Khudaverdyan<sup>7</sup>, Krisztián Kiss<sup>58,59</sup>, Sinan Kılıç<sup>20</sup>, Paul Klostermann<sup>70</sup>, Sinem Kostak Boca Negra Valdes<sup>69</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>, Lidjia Milasinovic<sup>85</sup>, Alissa Mittnik<sup>1,2</sup>, Janet M. Monge<sup>86</sup>, Georgi Nekhrizov<sup>15</sup>, Rebecca Nicholls<sup>87</sup>, Alexey G. Nikitin<sup>88</sup>, Vassil Nikolov<sup>15</sup>, Mario Novak<sup>29</sup>, Iñigo Olalde<sup>2,89</sup>, Jonas Oppenheimer<sup>2,27</sup>, Anna Osterholtz<sup>90</sup>, Celal Özdemir<sup>13</sup>, Kadir Toykan Özdoğan<sup>3</sup>, Nurettin Öztürk<sup>66</sup>, Nikos Papadimitriou<sup>91</sup>, Niki Papakonstantinou<sup>92</sup>, Anastasia Papathanasiou<sup>93</sup>, Lujana Paraman<sup>94</sup>, Evgeny G. Paskary<sup>95</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>96</sup>, Ashot Piliposyan<sup>97</sup>, Nada Pocuca Kuzman<sup>75</sup>, Hrvoje Potrebica<sup>98</sup>, Bianca Preda-Bălănică<sup>60</sup>, Zrinka Premužić<sup>99</sup>, T. Douglas Price<sup>100</sup>, Lijun Qiu<sup>2,27</sup>, Siniša Radović<sup>101</sup>, Kamal Raeuf Aziz<sup>102</sup>, Petra Rajić ikanjić<sup>29</sup>, Kamal Rasheed Raheem<sup>102</sup>, Sergei Razumov<sup>103</sup>, Amy Richardson<sup>84</sup>, Jacob Roodenberg<sup>104</sup>, Rudenc Ruka<sup>74</sup>, Victoria Russeva<sup>105</sup>, Mustafa Şahin<sup>57</sup>, Ayşegül Şarbak<sup>106</sup>, Emre Savaş<sup>69</sup>, Constanze Schattke<sup>3</sup>, Lynne Schepartz<sup>107</sup>, Tayfun Selçuk<sup>69</sup>, Ayla Sevim-Erol<sup>108</sup>, Michel Shamoon-Pour<sup>109</sup>, Henry M. Shephard<sup>110</sup>, Athanasios Sideris<sup>111</sup>, Angela Simalcsik<sup>32,112</sup>, Hakob Simonyan<sup>113</sup>, Vitalij Sinika<sup>103</sup>, Kendra Sirak<sup>2</sup>, Ghenadie Sirbu<sup>114</sup>, Mario Šlaus<sup>115</sup>, Andrei Soficaru<sup>35</sup>, Bilal Söğüt<sup>116</sup>, Arkadiusz Sołtysiak<sup>117</sup>, Çilem Sönmez-Sözer<sup>108</sup>, Maria Stathi<sup>118</sup>, Martin Steskal<sup>119</sup>, Kristin Stewardson<sup>2,27</sup>, Sharon Stocker<sup>38</sup>, Fadime Suata-Alpaslan<sup>120</sup>, Alexander Suvorov<sup>60</sup>, Anna Szécsényi-Nagy<sup>121</sup>, Tamás Szeniczey<sup>58,59</sup>, Nikolai Telnov<sup>103</sup>, Strahil Temov<sup>122</sup>, Nadezhda Todorova<sup>77</sup>, Ulsi Tota<sup>74,123</sup>, Gilles Touchais<sup>124</sup>, Sevi Triantaphyllou<sup>92</sup>, Atila Türker<sup>125</sup>, Marina Ugarković<sup>71</sup>, Todor Valchev<sup>16</sup>, Fanica Veljanovska<sup>122</sup>, Zlatko Videvski<sup>122</sup>, Cristian Virag<sup>126</sup>, Anna Wagner<sup>3,127</sup>, Sam Walsh<sup>84</sup>, Piotr Włodarczak<sup>128</sup>, J. Noah Workman<sup>2</sup>, Aram Yardumian<sup>129,130</sup>, Evgenii Yarovoy<sup>131</sup>,

Alper Yener Yavuz<sup>1,3,2</sup>, Hakan Yılmaz<sup>2,0</sup>, Fatma Zalzal<sup>2,27</sup>, Anna Zettl<sup>3</sup>, Zhao Zhang<sup>2</sup>, Rafet Çavuşoğlu<sup>2,0</sup>, Nadin Rohland<sup>2</sup>, Ron Pinhasi<sup>3\*</sup>, David Reich<sup>1,2,27,82\*</sup>

**Affiliations:**

<sup>1</sup>Department for 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, 1090 Vienna, Austria

<sup>4</sup>Mardin Artuklu University, Faculty of Letters, Department of Anthropology, Artuklu, 47510, Mardin, Turkey

<sup>5</sup>Sivas Cumhuriyet University, Faculty of Letters, Department of Anthropology, 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, İlkadım, 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>Mehmet Akif University, Faculty of Arts and Science, Department of Anthropology, 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, Tirana 1010, 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>Van Yüzüncü Yıl University, Faculty of Humanities, Department of Archaeology, 65090 Tuşba, Van, Turkey

<sup>21</sup>Düzce University, Faculty of Science and Letters, Department of Archaeology, 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 Kotor, 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>University of Veliko Tarnovo "St. St. Cyril and Methodius", 5003 Veliko Tarnovo, Bulgaria
- 5 <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
- 10 <sup>34</sup>Archaeological Museum in Zadar, 23000 Zadar, Croatia
- <sup>35</sup>Fr. I. Rainer" Institute of Anthropology, 76241 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
- 15 <sup>38</sup>University of Cincinnati, Department of Classics, 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>Ege University, Faculty of Letters, Department of Archaeology, 35100 Bornova-Izmir, Turkey
- 20 <sup>42</sup>University of Michigan, Museum of Anthropological Archaeology, Ann Arbor, MI 48109, USA
- <sup>43</sup>Narodni muzej Pančevo, 26101 Pančevo, Serbia
- <sup>44</sup>Human Paleoecology 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
- 25 <sup>46</sup>Gaziantep University, Faculty of Science and Letters, Department of Archaeology, 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>CIAS, Department of Life Sciences, University of Coimbra, 3000-456 Coimbra, Portugal
- 30 <sup>50</sup>Prahova County Museum of History and Archaeology, 100042 Ploieşti, Romania
- <sup>51</sup>Institut Català de Paleoecologia Humana i Evolució Social, 43007 Tarragona, Spain
- <sup>52</sup>Universitat Rovira i Virgili, Departament d'Història i Història de l'Art, 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
- 35 <sup>55</sup>Çukurova University, Faculty of Science and Letters, Department of Archaeology, 01330 Balçalı-Sarıçam-Adana, Turkey
- <sup>56</sup>Ankara University, Faculty of Humanities, Department of Anthropology, 06100 Sıhhiye, Ankara, Turkey

- <sup>57</sup>Uludağ University, Faculty of Science and Letters, Department of Archaeology, 16059 Görükle, Bursa, Turkey
- <sup>58</sup>Department of Biological Anthropology, Institute of Biology, Eötvös Loránd University, Budapest, Hungary
- 5 <sup>59</sup>Department of Anthropology, Hungarian Natural History Museum, 1117 Budapest, Hungary
- <sup>60</sup>Department of Cultures, University of Helsinki, 00100 Helsinki, Finland
- <sup>61</sup>Yerevan State University, 0025 Yerevan, Armenia
- <sup>62</sup>Regional Museum of History, 6300 Haskovo, Bulgaria
- <sup>63</sup>Ministry of Culture and Tourism, İsmet İnönü Bulvarı, 06100 Emek, Ankara, Turkey
- 10 <sup>64</sup>Museum of the City of Skopje, 1000 Skopje, North Macedonia
- <sup>65</sup>Malcolm H. Wiener Laboratory, American School of Classical Studies at Athens, 10676 Athens, Greece
- <sup>66</sup>Atatürk University, Faculty of Letters, Department of Archaeology, 25100 Erzurum, Turkey
- <sup>67</sup>Muğla Archaeological Museum and Yatağan Thermal Power Generation Company, Rescue Excavations, 48000 Muğla, Turkey
- 15 <sup>68</sup>Department of Anthropology, University of California Santa Barbara, Santa Barbara, CA 93106, USA
- <sup>69</sup>Bodrum Museum of Underwater Archeology, Çarşı Neighbourhood, 48400 Bodrum, Muğla, Turkey
- <sup>70</sup>Natural History Museum Vienna, Department of Anthropology, 1010 Vienna, Austria
- 20 <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
- 25 <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, UC San Diego, La Jolla, CA 92093, USA
- 30 <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>Department of Mediterranean Studies, Lab of Archaeometry, University of the Aegean, 85100 Rhodes, Greece
- 35 <sup>81</sup>The Cyprus Institute, Science and Technology in Archaeology and Culture Research Center, 2121 Aglantzia, Nicosia, Cyprus
- <sup>82</sup>Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
- <sup>83</sup>Cotsen Institute of Archaeology, UCLA, 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>University of Pennsylvania Museum of Archaeology and Anthropology, Philadelphia, PA 19104, USA

<sup>87</sup>School of Archaeological and Forensic Sciences, Faculty of Life Sciences, University of Bradford, Bradford, BD7 1DP, UK

<sup>88</sup>Department of Biology, Grand Valley State University, Allendale, MI 49401, USA

<sup>89</sup>BIOMICs Research Group, University of the Basque Country UPV/EHU, 01006 Vitoria-Gasteiz, Spain

<sup>90</sup>Department of Anthropology and Middle Eastern Cultures, Mississippi State University, MS 39762, USA

<sup>91</sup>Paul and Alexandra Canellopoulos Museum, 105-55 Athens, Greece

<sup>92</sup>Faculty of Philosophy, School of History and Archaeology, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

<sup>93</sup>Ephorate of Paleontology and Speleology, Greek Ministry of Culture, 11636 Athens, Greece

<sup>94</sup>Muzej grada Trogira, 21220 Trogir, Croatia

<sup>95</sup>Moldovan Historic - Geographical Society, 2044 Chişinău, Moldova

<sup>96</sup>French School of Archaeology at Athens, 10680 Athens, Greece

<sup>97</sup>Armenian State Pedagogical University After Khachatur Abovyan, 0010 Yerevan, Armenia

<sup>98</sup>Department of Archaeology, Faculty of Humanities and Social Sciences, University of Zagreb, 10000 Zagreb, Croatia

<sup>99</sup>Independent Researcher, 10000 Zagreb, Croatia

<sup>100</sup>University of Wisconsin-Madison, Laboratory for Archaeological Chemistry, Madison, WI 53706, USA

<sup>101</sup>Institute for Quaternary Palaeontology and Geology, Croatian Academy of Sciences and Arts, 10000 Zagreb, Croatia

<sup>102</sup>Sulaimaniyah Directorate of Antiquities and Heritage, Sulaimaniyah, Iraq

<sup>103</sup>Pridnestrovian University named after Taras Shevchenko, 3300 Tiraspol, Moldova

<sup>104</sup>The Netherlands Institute for the Near East, 2311 Leiden, Netherlands

<sup>105</sup>Bulgarian Academy of Science, Institute of Experimental Morphology, Pathology and Archeology with Museum, 1113 Sofia, Bulgaria

<sup>106</sup>Hitit University, Faculty of Science and Letters, Department of Anthropology, 19040 Çorum, Turkey

<sup>107</sup>School of Anatomical Sciences, The University of the Witwaters, 2193 Johannesburg, South Africa

<sup>108</sup>Ankara University, Faculty of Language and History - Geography, Department of Anthropology, 06100 Sıhhiye, Ankara, Turkey

<sup>109</sup>Binghamton University, Department of Anthropology, Binghamton University, Binghamton, NY 13902, USA

<sup>110</sup>Archaeological Institute of America, Boston, MA 02108, USA

<sup>111</sup>Institute of Classical Archaeology, Charles University, 11636 Prague, Czechia

<sup>112</sup>"Olga Necrasov" Centre of Anthropological Research, Romanian Academy Iași Branch, 2012 Iași Romania

<sup>113</sup>Scientific Research Center of The Historical And Cultural Heritage, 0010, Yerevan, Armenia

<sup>114</sup>Thracology Scientific Research Laboratory of the State University of Moldova, Department of Academic Management, Academy of Science of Moldova, 2009 Chișinău, Moldova

<sup>115</sup>Anthropological Center of the Croatian Academy of Sciences and Arts, 10000 Zagreb, Croatia

<sup>116</sup>Pamukkale University, Faculty of Science and Arts, Department of Archaeology, 20070 Denizli, Turkey

<sup>117</sup>Faculty of Archaeology, University of Warsaw, 00-927 Warszawa, Poland

<sup>118</sup>Ephorate of Antiquities of East Attica, Ministry of Culture and Sports, 10682 Athens, Greece

<sup>119</sup>Austrian Archaeological Institute at the Austrian Academy of Sciences, 1190 Vienna, Austria

<sup>120</sup>Istanbul University, Faculty of Letters, Department of Anthropology, 34134 Istanbul, Turkey

<sup>121</sup>Laboratory of Archaeogenetics, Institute of Archaeology, Research Centre for the Humanities, Hungarian Academy of Sciences, 1097 Budapest, Hungary

<sup>122</sup>Archaeology Museum of North Macedonia, 1000 Skopje, North Macedonia

<sup>123</sup>University of Avignon, Avignon, France

<sup>124</sup>Université Paris 1 Panthéon-Sorbonne, F-75006 Paris, France

<sup>125</sup>Ondokuz Mayıs University, Faculty of Science and Letters, Department of Archaeology, 55139 Atakum-Samsun, Turkey

<sup>126</sup>Satu Mare County Museum, 440031 Satu Mare, Romania

<sup>127</sup>Laboratory of Anthropology, University of Granada, 18011 Granada, Spain

<sup>128</sup>Institute of Archaeology and Ethnology, Polish Academy of Sciences, 50-118 Krakow, Poland

<sup>129</sup>Department of History-Social Sciences, Bryn Athyn College, Bryn Athyn, PA 19009, USA

<sup>130</sup>University of Pennsylvania, Penn Museum, PA 19104, USA

<sup>131</sup>Moscow Region State University, Moscow Region, 141014 Mytishi, Russia

<sup>132</sup>Burdur Mehmet Akif Ersoy University, Istiklal Campus, Department of Anthropology, 15100 Burdur, Turkey

† Co-lead authors

\* Corresponding authors. E-mail: Iosif Lazaridis (lazaridis@genetics.med.harvard.edu), Songül Alpaslan-Roodenberg (MuyesserSongul\_Alpaslan@hms.harvard.edu), Ron Pinhasi (ron.pinhasi@univie.ac.at), David Reich (reich@genetics.med.harvard.edu)

**Abstract:** We present an integrative genetic history of the Southern Arc, an area divided geographically between West Asia and Europe, but which we define as spanning the culturally entangled regions of Anatolia and its neighbors, in both Europe (Aegean and the Balkans), and in West Asia (Cyprus, Armenia, the Levant, Iraq and Iran). We employ a new analytical framework to analyze genome-wide data at the individual level from a total of 1,320 ancient individuals, 731 of which are newly reported and address major gaps in the archaeogenetic record. We report the first ancient DNA from the world's earliest farming cultures of southeastern Anatolia and northern Mesopotamia, as well as the first Neolithic period data from Cyprus and Armenia, and discover that it was admixture of Natufian-related ancestry from the Levant—mediated by Mesopotamian and Levantine farmers, and marked by at least two expansions associated with dispersal of pre-pottery and pottery cultures—that generated a pan-West Asian Neolithic continuum. Our comprehensive sampling shows that Anatolia received hardly any genetic input from Europe or the Eurasian steppe from the Chalcolithic to the Iron Age; this contrasts with Southeastern Europe and Armenia that were impacted by major gene flow from Yamnaya steppe pastoralists. In the Balkans, we reveal a patchwork of Bronze Age populations with diverse proportions of steppe ancestry in the aftermath of the ~3000 BCE Yamnaya migrations, paralleling the linguistic diversity of Paleo-Balkan speakers. We provide insights into the Mycenaean period of the Aegean by documenting variation in the proportion of steppe ancestry (including some individuals who lack it altogether), and finding no evidence for systematic differences in steppe ancestry among social strata, such as those of the elite buried at the Palace of Nestor in Pylos. A striking signal of steppe migration into the Southern Arc is evident in Armenia and northwest Iran where admixture with Yamnaya patrilineal descendants occurred, coinciding with their 3<sup>rd</sup> millennium BCE displacement from the steppe itself. This ancestry, pervasive across numerous sites of Armenia of ~2000-600 BCE, was diluted during the ensuing centuries to only a third of its peak value, making no further western inroads from there into any part of Anatolia, including the geographically adjacent Lake Van center of the Iron Age Kingdom of Urartu. The impermeability of Anatolia to exogenous migration contrasts with our finding that the Yamnaya had two distinct gene flows, both from West Asia, suggesting that the Indo-Anatolian language family originated in the eastern wing of the Southern Arc and that the steppe served only as a secondary staging area of Indo-European language dispersal. The demographic significance of Anatolia on a Mediterranean-wide scale is further documented by our finding that following the Roman conquest, the Anatolian population remained stable and became the geographic source for much of the ancestry of Imperial Rome itself.

**One-Sentence Summary:** A genetic history of the first farmers, early Indo-Europeans, Greco-Romans, Turks and Armenians between Europe and West Asia.

**Main Text:** The Histories of Herodotus of Halicarnassus—a native of a Greek colony on the Aegean coast of Anatolia—begin (Hdt. 1.4) with an account of the conflicts between Europe and Asia that anticipated his main topic of the Greco-Persian Wars of the 5<sup>th</sup> century BCE. The sharp distinction made in Herodotus’s work, between Europe and Asia, has survived in the worldview of succeeding generations despite its being a convenient narrative tool rather than a topographical or geological reality. From the vantage of this binary distinction, the Balkans and Anatolia are seen as peripheral to Europe and Asia respectively, rather than 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” (SA), a region centered on the large Anatolian peninsula (Turkey), including in the west Southeastern Europe (the Balkans and the Aegean), and in the east and south Cyprus, Mesopotamia, the Levant, Armenia, and Iran. We present new genome-wide DNA data from 783 individuals from the SA: 731 previously unsampled and 52 previously published for which we report new data (Data S1) from 1,398 new libraries (Data S2). For 239 individuals we provide new radiocarbon dates (Data S3). We study these along with previously published individuals for a total sample size of 1,320 (Fig. 1).

Our sampling includes all major geographical regions of Turkey, with particular density from the western (Aegean, Marmara), northern (Black Sea) and eastern (Eastern Anatolia, Southeastern Anatolia) regions connecting it with the rest of the SA. Our data allow us to study Anatolia across all time periods from the previously unsampled Pre-Pottery Neolithic (PPN) of its southeastern extreme (northern Mesopotamia), through post-medieval times. We also report on another set of PPN period individuals from the island of Cyprus, resolving questions around the ancestral origins of some of the world’s first farmers who lived in a location geographically intermediate between Europe, Anatolia, and the Levant. An area of emphasis is Armenia with substantial sampling of the Bronze and Iron Ages representing an order of magnitude more individuals than previously available, and also including the first Neolithic-era individuals from this region. Many individuals of Bronze-to-Iron Age timeframe are also sampled from the Iranian highlands at Hasanlu, where only a single individual has previously been studied (1), and Dinkha Tepe, neighboring Anatolia, Mesopotamia, Armenia and the Caucasus. In the southern part of southeastern Europe we sample Mycenaean-era individuals from several regions of the Aegean, including Attica, Kastrouli near Delphi in Phokis, and the “Palace of Nestor” at Pylos, including the “Griffin Warrior”(2), and a previously unreported individual of the Middle Minoan period from Zakros in Crete. From the southern Balkans, we present a novel full time transect of Albania, numerous individuals from North Macedonia where previously data from only a single Neolithic individual had been published,(3) as well as more than doubling the previously available individuals from Bulgaria. Further north, at the western wing of the SA, 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: >100 Bronze Age individuals, including many from Cetina Valley and Bezdanjača Cave in Croatia are added to only 5 previously published from the entire area.(3, 4) Some of the Balkan individuals include Yamnaya culture individuals from Serbia and Bulgaria, allowing us to compare them with those of the Eurasian steppe. Details of all studied individuals can be found in(5).

To discuss the geographic distribution of these individuals, we take a flexible approach, in some case 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 be more precise(6). In the interests of having a

uniform nomenclature that is easily accessible to readers familiar with the current political map of the SA, we also refer to groups of individuals with namings prefixed with 3-letter International Standards Organization (ISO) codes for countries (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 archaeological designations that are conventional for each region, e.g., Eneolithic and Chalcolithic both denote copper-using cultures in different parts of the archaeological literature, and the transition between them and the Bronze Age did not occur simultaneously in different parts of the Southern Arc. Any ambiguities may be resolved by referring to the precise geographical and temporal information or archaeological context for all individuals provided in Data S1, which also includes additional information about the individuals' genetic characteristics.

### Overview of genetic variation in the Southern Arc

To understand genetic variation in the SA, we began with ADMIXTURE (Fig. 1C, Fig. S1) analysis, which allowed us to detect individuals with non-West Eurasian-related ancestry and to outline the broad pattern of ancestry in terms of "Balkan hunter-gatherer" (BHG), "Eastern European hunter-gatherer" (EHG), Anatolian-Levantine-related, and Iran/Caucasus-related components. Principal components analysis (Fig. 1D) of SA individuals together with other West Eurasian individuals demonstrate the central position of the SA 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 SA individuals, we developed a 5-source modeling framework (using qpAdm and F4admix) that is able to both model the SA population as a whole and as individuals. Our framework aims to maximize the quality of the statistical fit of the model while minimizing the standard errors in inferences of ancestry proportions for as large a number of individuals as possible: we emphasize that the inferred source populations are only surrogates for the true sources and may not be directly ancestral to later groups. One advantage of the approach is that it does not group individuals arbitrarily into fixed populations, but allows one to study these flexibly at any desired level of granularity. The five sources of ancestry in our model are: Caucasus hunter-gatherers (CHG), EHG, Levantine PPN, BHG from the Iron Gates in Serbia, and Northwestern Anatolian Neolithic from Barcin; the inferred proportions of ancestry for individuals (Data S4) are summarized in Fig. S2 and discussed in.

We summarize key genetic inferences drawn by our study in Table 1 and in what follows.

### The Neolithic continuum

Previous studies have shown that the Neolithic population of northwestern Anatolia was related to the Neolithic and Epipaleolithic population of Central Anatolia and this "Anatolian Neolithic" cluster was the main ancestry source for the Neolithic populations of southeastern Europe which then spread more broadly across Europe. At the eastern end of the SA, a highly differentiated population of farmers lived in Iran, related to the CHG. The Levantine PPN was related to the NW Anatolian Neolithic population, but could be derived to a great degree from local Epipaleolithic Natufian hunter-gatherers.

Our study fills the vacuum in the heartland of the SA, between the Anatolian/Zagros/Levant edges, and reveals the existence of hitherto unknown intermediate populations both by the 5-way

admixture model and the co-analysis of Neolithic populations without assumptions of which one is admixed or not(5), thus avoiding “publication order bias”, i.e., the tendency to update published models to accommodate new data rather than always infer models anew. PPN farmers from northern Mesopotamia (Boncuklu Tarla near Mardin in southeastern Turkey and Nemrik 9 in northern Iraq) were admixed between the Iran/Caucasus northern end and Anatolian/Levantine southern end of the West Asian cline, and so were Neolithic individuals from Armenia and Azerbaijan (Fig 1). The two sampled 6<sup>th</sup> millennium BCE individuals from Aknashen and Masis Blur (Armenia) differed in being more Caucasus- and Levant-like respectively despite being sampled only ~20km and a few centuries apart; thus, the Neolithic of Armenia was not homogeneous but exhibited variation which also encompassed its neighbors in Azerbaijan.(15) The Armenia and Azerbaijan Neolithic individuals differed from the much earlier 9<sup>th</sup> and 8<sup>th</sup> millennia BCE pre-pottery farmers from northern Mesopotamia in having more Anatolian ancestry whereas the contemporary pre-pottery farmers from the north Zagros at Bestansur and Zawi Chemi Shanidar in Iraq had minimal Anatolian ancestry and were genetically intermediate between the Levant and early farmers of the central Zagros from Ganj Dareh.(10) Finally, we report the first PPN individuals from the island of Cyprus: these individuals, whose fragmentary remains were disposed of within a disused water well at from Kissonerga-Myliouthkia, were geographically intermediate between Anatolia and the Levant within the Eastern Mediterranean, but resembled the Neolithic of Central Anatolia more than that of the Levant.

The newly reported Neolithic individuals from West Asia prove that genetically intermediate populations had already existed in this archaeologically interconnected region(16) in the earliest Neolithic in what must have been a continuum of ancestry encompassed by the highly differentiated previously known Paleolithic West Asian sources of Caucasus-Iran, Anatolia, and the Levant(7, 9, 10) (Fig. 1E). This continuum was generated by admixture from a Natufian-related source(5) via intermediaries: maximally in the formation of the Levantine PPN population (~80%) and ~44% of the ancestry of the North Mesopotamian PPN population, which then contributed variably (~30-70%) to the PPN populations of Central Anatolia(13, 14), complementing the local hunter-gatherer ancestry represented by the Epipaleolithic sample from Pınarbaşı. The ancestry in the pottery Neolithic populations of Anatolia at Barcın(9) and Çatalhöyük,(14) also included a 3<sup>rd</sup> source of ~13-17% Levantine PPN ancestry, differentiating them from the Anatolian PPN populations which lacked this component. Thus, at least two migration waves transmitted Natufian ancestry northwestward: one well-modelled as Mesopotamian-only, mediating the spread of the PPN, and one more Levantine in ancestry mediating the spread of the pottery Neolithic.

The rich history of admixture that formed West Asian Neolithic populations presents an important reminder that the modeling of later populations in terms of the 5-way model, to which we now turn, should not be interpreted as indicating recent admixture, but rather trends of ancestral composition across different individuals using a common measuring instrument.

### **The Caucasus and Anatolian bridges**

We discuss our results on individuals from the Chalcolithic and Bronze Age in terms of two geographical bridges connecting West Asia and Europe: the Caucasus bridge linking West Asia via the South and North Caucasus to the Eurasian steppe in the east (Fig. 2A; Figs. S3, S4), and the Anatolian bridge linking West Asia via Anatolia and the Aegean to the Balkans in the west (Fig 2b). Tracing the Caucasus bridge from the steppe southwards, we see that EHG ancestry predominates in the north, is drastically reduced in the Maykop cluster of the North



Caucasus(17), is present in Armenia, is minimal in Iran, and is virtually non-existent in Anatolia and the Levant. Conversely, CHG ancestry is ubiquitous across the Caucasus bridge, higher in the south, but also comprises nearly half of the ancestry in the steppe Yamnaya cluster and the piedmont of the North Caucasus at Progress-2(17) and far beyond it in the Eneolithic at Khvalynsk.(9) CHG ancestry became widespread in the steppe at the earliest in the 6<sup>th</sup> millennium BCE, given the presence of an EHG hunter-gatherer at Lebyazhinka IV in the Samara region(8) in the first half of that millennium and the appearance of the CHG-admixed Eneolithic individuals at its end and into the 5<sup>th</sup> millennium BCE. CHG ancestry was present in the Neolithic of Armenia, Azerbaijan, and Mesopotamia discussed above, and made its appearance in the Levant by the Chalcolithic of the late 5<sup>th</sup> millennium BCE.(18)

Anatolia was inhabited by contrasting Northwest (Marmara region) and Mesopotamia (Southeast region) populations sampled for the first time in our study for the Neolithic, with Central Anatolia being similar overall to the Northwest.(12, 13) During the Chalcolithic and Bronze Age, SE Anatolia was composed of a CHG/Levantine mix, but with some NW Anatolian ancestry. Conversely, in the Central and NW regions, CHG ancestry became the second most prevalent component. The Chalcolithic was a period of homogenization within Anatolia in which “eastern” ancestry flowed westward across the peninsula, reaching the very west by the early 4<sup>th</sup> millennium BCE. CHG ancestry also crossed into Europe by the Bronze Age, when it was present in the Aegean and the Balkans. Southeastern Europe experienced an influx of EHG ancestry southward from the steppe, penetrating across the Balkans and—to a lesser extent—the Aegean, but it did not cross into Anatolia (from the west) to the limits of our resolution.

The two bridges linking West Asia and Europe were not bidirectional with regard to their demographic impact but instead were both characterized by the impermeability of Anatolia to exogenous gene flow, which could be explained by either a large and stable population base that attenuated the demographic impact of external migration, or cultural factors impeding migration altogether. The asymmetry of migration between Anatolia and its neighbors is evident for example in the fact that CHG-related ancestry flowed westward across Anatolia into the Balkans and northward into the Eurasian steppe, but BHG ancestry did not flow into Anatolia or further eastward, and EHG ancestry entered West Asia only as far south as Armenia and to a lesser extent Iran, but not into Anatolia (from the east) or the Levant. The source of CHG ancestry in the Chalcolithic and afterward need not have been a recently expanding population, considering that CHG ancestry was present in the Caucasus since at least the Upper Paleolithic, as well as in northern Mesopotamia since at least the PPN period. At most, we can speak of expansions out of the general area of the Caucasus-Mesopotamia-Zagros northwards (into the steppe) and westwards (across Anatolia and into the Levant).

The steppe itself did not receive only CHG ancestry, as Yamnaya-cluster individuals ~3000BCE there have an excess of Anatolian ( $3\pm 1\%$ ) and Levantine ( $3.5\pm 1\%$ )-related ancestry (Fig. 2A,D) not present in the aforementioned 5<sup>th</sup> millennium Eneolithic individuals in our general 5-source model. This inference is further supported by our analysis of Yamnaya origins(5) which indicates that they derived ancestry from at least two southern sources: an early Eneolithic one of CHG ancestry only, and a later (some time prior to the formation of the Yamnaya cluster) that also included Anatolian/Levantine-related ancestry (distally), ancestry related to the Neolithic of Armenia, which included both of these components (more proximally), and Chalcolithic populations of the Caucasus to SE Anatolia (even more proximally). Archaeological evidence documents how steppe populations interacted culturally with European farmer groups in the west

such as the Tripolye-Cucuteni and Globular Amphora, and it was previously suggested that ancestry from such groups contributed to the ancestry of the Yamnaya.(17) Our genetic results contradict this scenario, as European farmers were themselves a mixture of Anatolian Neolithic and European hunter-gatherer (BHG or EHG) ancestry, but the Yamnaya: (i) lacked the European hunter-gatherer ancestry differentiating European from West Asian farmers, and (ii) had an approximately 1:1 ratio of Levantine-to-Anatolian ancestry in our 5-way model, contrasting with the near absence of Levantine ancestry in European farmers. The CHG/EHG/WHG/Anatolian Neolithic model of (17) fails ( $p < 1e-10$ ) 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 the Yamnaya firmly to the south of the steppe and the eastern wing of the SA; 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 CHG infiltration northwards and identify the likely geographical region for the emergence of the distinctive Yamnaya cluster which we show has an autosomal signal of admixture(8) dating to the mid-5<sup>th</sup> millennium BCE (Fig. S5), coinciding with the direct evidence of the first southern influence provided by the Eneolithic individuals of the steppe.

CHG and EHG ancestry were present in roughly equal proportions in the Yamnaya cluster and Corded Ware/Bell Beaker-related individuals of central-northern Europe, but in Southeastern Europe CHG ancestry exceeded EHG ancestry (Fig. 2B,D). Thus, while CHG ancestry in central-northern Europe can be parsimoniously explained by migrations of steppe pastoralists introducing an approximately equal CHG/EHG mix,(8) this was not the case in southeastern Europe where both migration from the steppe (accounting for the EHG component), and across the Anatolian bridge (accounting for the excess CHG one) must have taken place.

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

While the analysis of genome-wide ancestry provides evidence for continuity and discontinuity in a particular region, it cannot prove a recent genetic link as the five sources of our model represent “deep” ancestry. Y-chromosome variation ((5); Data S4) can be used to document the timing of shared ancestry between populations, and has the potential to provide insight into the social processes of patrilineal societies. Subclades of Y-chromosome haplogroup R-L389 are informative for tracing connections between the SA and the Eurasian steppe (Fig. 3). First, haplogroup R-V1636 with an inferred common ancestor in the 5<sup>th</sup> millennium BCE documents gene flow between the steppe and the SA in the Eneolithic/Chalcolithic period (Fig. 3B): it is present in two individuals from the Late Chalcolithic at Arslantepe (Turkey)(15) and the Early Bronze Age in Armenia at Kalavan(10), and 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).(19) Importantly, the individuals from Armenia and Arslantepe lack any detectable EHG autosomal ancestry (Fig. 3C) 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 SA 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 CHG ancestry, suggesting an eastern European origin of this clade which would eventually give rise to the R-M269 clade that spread extremely widely in the Bronze Age.

5 Haplogroup R-M269, which is inferred to have a shared common ancestor in the 5<sup>th</sup> millennium BCE, and its subclades are crucial for understanding steppe expansions as it was the dominant lineage of the Yamnaya-Afanasievo group,(4, 8, 20) in its 4<sup>th</sup> millennium BCE R-Z2103→R-M12149 sub-lineage. Our newly reported data reveals that a large proportion of individuals in Armenia and NW Iran belonged to this haplogroup during the 2<sup>nd</sup> and early 1<sup>st</sup> millennium BCE, 10 providing a genetic link with the Yamnaya in these regions where no archaeological presence of the Yamnaya culture is attested. It represents a more direct link than either R-V1636 or the early appearance of EHG ancestry at Areni1 cave in Armenia(10) during the Chalcolithic at the end of the 5<sup>th</sup> millennium BCE and the converse movement of CHG ancestry into the steppe Eneolithic. Steppe ancestry is not detected in the Early Bronze Age sites in Armenia associated with the 15 Kura-Araxes culture, and only re-appeared during the Middle/Late Bronze Age and onwards. The actual date of steppe admixture in the ancestry of Bronze-to-Iron Age people from Armenia is estimated to the mid-3<sup>rd</sup> millennium BCE on average based on the decay of admixture linkage disequilibrium (Fig. S5) indicating that the mixture started in an earlier group before becoming geographically widespread after the decline of the Kura-Araxes culture.

20 Despite the Y-chromosome movement southward attested by our data, any association between R-haplogroup bearers and EHG ancestry was lost south of the steppe as these had similar proportions of EHG ancestry as I-Y16419 bearers (the second most prevalent lineage in Armenia). Two Bronze-to-Iron Age sites with substantial sample sizes (Bagheri Tchala, n=7 and 25 Noratus, n=12 unrelated males) have contrasting haplogroup distributions dominated by R-M12149 and I-Y16419 respectively (Fisher's exact test  $p < 0.001$ ), suggesting the existence of a patrilocal mating system ~1000BCE in Armenia. During the same period at Hasanlu in NW Iran many individuals have no trace of EHG ancestry at all despite the presence of R-M12149 there, suggesting that the initial association of this lineage with EHG ancestry on the steppe had 30 vanished as R-M12149 bearers intermarried with SA groups without EHG ancestry (Fig. 3C). We hypothesize that the major steppe takeover that replaced R-M12149 (within haplogroup R1b) with R-Z93 Y-chromosomes (within haplogroup R1a) sometime during the 3<sup>rd</sup> millennium BCE—as the Corded Ware/Fatianovo(21) descendants spawned later steppe cultures such as Sintashta(20)—may have either caused the southward migration of Yamnaya-descended groups, 35 or followed it to fill the niche they had vacated. Where the two lineages lived until the Bronze Age and whence their dramatic demographic expansion began may shed light on their possible rivalry. The Yamnaya descended R-M12149 patrilineages survived in Armenia, but their autosomal EHG legacy was diminished first by a factor of 4, as Yamnaya descendants mixed with local inhabitants to form the Bronze/Iron Age population we sample in our study, and then 40 by another factor of 3 during the 1<sup>st</sup> millennium BCE(5). In terms of deep ancestry, the population of Armenia has been resilient to change, composed primarily of the same CHG/Anatolian/Levantine mix that has prevailed in the region since the Neolithic. However, this picture of continuity can be deceptive, as the 3× reduction of Steppe ancestry in the 1<sup>st</sup> millennium BCE implies a massive replacement of 2/3 of the local population as steppe 45 descendants admixed with local and neighboring groups of West Asia of similar deep ancestral mix but lacking the EHG ancestry. This process was largely complete by ~2,000 years ago, when people very similar ancestrally to present-day people were present at sites like Aghitu and ~1,000 years ago at Agarak.

5 Steppe migration into mainland Europe and South Asia (unlike into Armenia) was carried out by  
people with steppe ancestry who were not, however, patrilineal descendants of the dominant R-  
M12149 of the Yamnaya population. Instead, they belonged to different descent groups that had  
received autosomal steppe admixture while replacing their Y-chromosome lineages. Armenia  
10 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%  
C.I.: 0-4.5%) and in which haplogroups J (36 individuals) and G (17 individuals) are most  
common with the former achieving such prominence despite occurring in only 1/26 Neolithic  
individuals from Barcin during the pre-Chalcolithic period.

### **The Urartian Kingdom**

15 The linguistic landscape in the mountainous and geographically fragmented regions of eastern  
Turkey and Armenia must have been complex in the Bronze and Iron Ages; a clue to this is the  
genetic differentiation existing within the people of the Iron Age kingdom of Urartu. The people  
at the center of this kingdom in the Lake Van region of Turkey (Çavuştepe) and its northern  
extension in Armenia, were strongly connected by material culture, and were buried only  
~200km apart, yet formed distinct genetic clusters with little overlap during the kingdom's early  
20 (9<sup>th</sup>-8<sup>th</sup> c. BCE) period. The Van cluster is in continuity with the pre-Urartian population  
(~1300BCE) at neighboring Muradiye also in the Van region, and is characterized by more  
Levantine ancestry and the absence of steppe ancestry, contrasting it with the cluster of Urartian  
period individuals from Armenia which have less Levantine and some steppe ancestry like the  
pre-Urartian individuals of the Early Iron Age(5). Our genetic results help explain the formation  
25 of linguistic relationships in the region. Population continuity of the Lake Van core population  
with greater "Levantine" ancestry may well correspond to the Hurro-Urartian language  
family(22) that linked the non-Indo-European Urartian language of the kingdom with the earlier  
Bronze Age Hurrian language whose more southern distribution encompassed parts of Syria and  
North Mesopotamia. Into the periphery of this Hurro-Urartian linguistic sphere came a steppe-  
30 admixed population from the north, whose presence marks the southern edge of steppe  
expansion we discussed above and whose close proximity to the Urartian speakers would  
provide a mechanism for the incorporation of Urartian words into the Armenian lexicon.

### **Interplay of Local, Steppe, and West Asian ancestries in southeastern Europe**

35 Southeastern Europe preserved the Anatolian Neolithic ancestry of its Neolithic cultural  
forebears during the Chalcolithic and Bronze Age (Fig. 2A), and also—to a lesser degree—of the  
BHG ancestry local to the region. BHG ancestry was detected only in the Balkans, thus  
precluding any substantial migration from the area to the rest of the SA, and even within the  
Balkans, it diminishes from the north to the Aegean where it is virtually non-existent. This raises  
40 the question of the genetic composition of pre-Neolithic Aegean populations: were they  
genetically like those of the Iron Gates to the north or Anatolia to the east? BHG ancestry was  
variable during the Bronze Age and related to geography: a striking contrast is found within  
Romania where our new data shows that it makes up ~12% of the ancestry of 42 individuals  
from the Bodrogkeresztúr Chalcolithic, and ~24-30% in 10 Bronze Age individuals from Arman,  
45 and in Ploiești and Târgșoru Vechi south of the Carpathian Mountains. Together with another  
Bronze Age individual from Padina in Serbia (2460-2296 calBCE) near the Iron Gates, whose  
BHG ancestry was a maximum of ~37%, these results prove a remarkable preservation of  
hunter-gatherer ancestry in the North Balkans postdating the arrival of both Anatolian Neolithic  
and steppe ancestry in the region.

EHG ancestry albeit in mixed form became ubiquitous in Southeastern Europe during the Bronze Age after its sporadic Chalcolithic appearance.<sup>(3)</sup> This is most evident (~31-44%) in Moldova and individuals from Romania (Trestiana and Smeeni) on the eastern/southeastern slopes of the Carpathians which contrast with the high-BHG group from Arman. For the rest of the Balkans, the amount of EHG ancestry is ~15% and drops to ~4% in Mycenaean Greece from which we report data from multiple new sites at Kastrouli near Delphi, Attica, and the Palace of Nestor in Pylos that complement previously published ones from the Peloponnese and Salamis.<sup>(23)</sup> EHG ancestry is absent in a newly reported Middle Minoan individual from Zakros in Crete, generally similar to those previously published,<sup>(24)</sup> but with significant Levantine ancestry (~32%) which is consistent with her either being a migrant to the island from the east or part of a structured Cretan population whose past ethnic diversity was noted as early as the *Odyssey* of Homer (Hom. Od. 19.172-177). We show for the first time, that EHG ancestry was also absent in some Mycenaean individuals, suggesting that while the contrast between the mainland and Crete was real (Fig. S6), the penetration of EHG ancestry did not reach the totality of the mainland population during the Late Bronze Age and was even significantly variable within Mycenaean sites. The Griffin Warrior<sup>(2)</sup>, the earliest individual (~1450 BCE) from the Palace of Nestor in Pylos, is right in the middle of the general population of the Aegean, and was thus plausibly of local Aegean origin. However, he was at the edges of the Mycenaean and Minoan clusters, and more specifically, had no detectible EHG ancestry ( $5.9 \pm 1.1\%$  less than the rest of the individuals sampled at the Palace; Fig. S6). This finding could be consistent with a Cretan origin of this individual in accordance with the Minoan iconography of the rich burial goods of his shaft grave; alternatively he could be drawn from a mainland population that had not experienced EHG admixture, as could two other individuals from Pylos buried close to the Palace in a chamber tomb and outside the limits of the town in a cist grave. Regardless of the Griffin Warrior's proximate origins, data from this individual and the others from Pylos shows that there was no general association of elite burial status with ancestry. Variation in EHG ancestry is observed at short geographical distance scales, as we observe that 4 individuals of the population of Attica buried at Marathon had only  $2 \pm 1\%$  EHG ancestry that was significantly less (by more than 2 standard errors) than those of the neighboring island of Salamis and all sampling locations in the Peloponnese, provisionally suggesting that the classical Athenian claim (e.g., Plat. Menex. 237b) of having received fewer migrants than other Greek *poleis* in the remote past may not have been altogether invented. But migrants did come to Greece as a whole, as shown e.g., by a Y-chromosome match of the rare R-PF7562 haplogroup between a pair of patrilineal relatives from the Palace of Nestor (Fig. 3B) links Late Bronze Age Mycenaean Greece with an Early Bronze Age individual of the North Caucasus at Lysogorskyja that is genetically similar to the Yamnaya.<sup>(17)</sup> One of these two relatives at Pylos (I13518) was almost certainly the offspring of first cousins; we document such close-kin unions not only in elite Mycenaean society but also in different localities of the Bronze Age SA (Fig. S7 and (5), including an individual from Bezdanjača in Croatia (I18717) who was likely the offspring of an uncle/niece pairing, documenting for the first time the later persistence of the practice of close kin marriage that had started with the Neolithic<sup>(14, 25)</sup>. Did descriptions of such unions in classical mythological accounts of the "Heroic Age" reflect a memory of real Bronze Age events or were they based on practices that had persisted to the authors' own time? Future ancient DNA studies will complement the written sources in documenting the later history of marriage norms in the region.

Several other individuals from the Balkans could be assigned to haplogroup R-M269, whose earliest single example in Chalcolithic Smyadovo in Bulgaria<sup>(3)</sup> in the mid-5<sup>th</sup> millennium BCE

should be treated with some caution as it is based on a single SNP and is also an ancestry outlier for its population. All other known individuals carrying R-M269 are from the Bronze Age or later; a group of six (Fig. 3B) are associated with high steppe ancestry (>30% EHG) and this includes not only individuals from Moldova (Catacomb and Multi Cordoned Ware), adjacent to the steppe, but also from further south. This group furthermore includes two early Bronze Age individuals 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). A female from Boyanovo and a Yamnaya-associated individual from Vojlovica-Humka in Serbia also have high steppe ancestry. Thus, in the Balkans, unlike Armenia, the gamut of steppe ancestry extended from zero to steppe-like levels at the dawn of the Bronze Age. Inferred dates of admixture between Steppe pastoralist and European farmer ancestry in Southeastern European individuals from the Bronze Age to the 1<sup>st</sup> millennium BCE are inversely correlated with the absolute dates of the individuals (Fig. S8), and lead to the inference of major admixture between steppe and local Neolithic populations ~4,850 years ago, soon after the earliest expansion of the Yamnaya culture.(26) 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, beyond both the cultural “frontier” between steppe pastoralist and settled populations, and the geographical transition from the eastern European flatlands into mountainous areas. By the Late Bronze Age and later, no high-steppe ancestry individuals are observed, but steppe-related Y-chromosomes persist, including R-Z2106, a lineage that links North Macedonia (Ulanci-Veles), Albania (Çinamak), the steppe, and Armenia.

Southeastern Europe contrasts strongly with the Afanasievo,(4, 20) Corded Ware/Fatyanovo/Sintashta(4, 8, 20, 21) and Beaker(27) steppe expansion that have strong associations with R-M417 and R-L51 Y-chromosome haplogroups that collectively make up only ~5% of SA individuals from all periods in the SA, and with the remarkable presence of R-M12419 Yamnaya descendants observed in Armenia discussed above. Together with the extraordinary heterogeneity in autosomal ancestry in the Balkans, a picture emerges of a fragmented genetic landscape in the Balkans that may well parallel the poorly understood linguistic diversity of ancient IE Paleo-Balkan speakers prior to the spread of Latin and Slavic languages of which Albanian is the only surviving representative.

### **The Southern Arc after the Bronze Age: source and recipient of migrants**

We investigated the effects of Greek colonization by identifying individuals from both the SA and outside of it that were genetically similar to the Bronze Age Mycenaeans (see (5); Fig. 4A). This identifies an Archaic individual from Kastrouli in Phokis and individuals at Empúries in Spain(28), an outpost colonized by Phocaeans from western Anatolia who were themselves said to be colonists from Phokis (Paus. 7.3.10), thus capturing the end points of a long chain of transmission, with little admixture, across the Mediterranean; could the ancestry of the Empúries individuals be traced back to the beginning of this chain or was it drawn from another, genetically similar source? In the future, larger sample sizes from Greek cities may test specific metropolis-colony connections and document the extent to which migration and admixture played a role in the era of Greek colonization. Mycenaean-like ancestry spread also to the eastern Mediterranean as in the case of a “Philistine” individual from Ashkelon(29). We also show the similarity of at least some individuals from inland Thrace (at Kapitan Andreevo) with the

5 Mycenaean predating the Greek colonization, suggesting that the Mycenaean were genetically similar to some Thracians from the east Balkans, a region outside the sphere of the Late Bronze Age Aegean, cautioning us not to conflate genetic with cultural similarity. By contrast, Bronze-  
10 to-Iron Age individuals from Çinamak in Albania, and from numerous sites of North Macedonia, are genetically differentiated from the Mycenaean populations and do not fall within its variation; these areas to the north of Greece are differentiated by having  $\sim 3\times$  more EHG ancestry than the Mycenaean average and are similar to the Middle Bronze Age population from Elati-Logkas(30) in northern Greece(5). The coastal regions of Anatolia formed another area of Greek settlement during the Iron Age and much of the Anatolian peninsula was incorporated into the  
15 Hellenistic kingdoms established by the successors of Alexander the Great, providing opportunity for population transfer from Southeastern Europe to Anatolia. Yet, we do not find Mycenaean-like individuals either at 1<sup>st</sup> millennium BCE Greek colony sites such as Halicarnassus (modern Bodrum) or Amisos (modern Samsun) in the Aegean and Black Sea regions respectively, consistent, perhaps, with the account of Herodotus that early Greek colonists of Anatolia married indigenous Carian women of Anatolia when they first settled there (Hdt. 1.146), or across the peninsula in general reminiscent of the policy of intermarriage between Alexander himself and his companions with local women of the conquered Persian Empire (Arr. An. 7.4.4ff).

20 We were intrigued by the recently observed shift in population in Imperial Rome(31) towards the general Near East, and sought to identify the precise geographic sources of it by co-analyzing the data from Italy with the extensive new sampling from the SA. We were surprised to find that Romans of the Imperial period were on average almost identical to Roman/Byzantine individuals from Anatolia(5). Given the considerable diversity within both sets of populations, we clustered  
25 diverse Roman/Byzantine/Medieval individuals and immediate predecessors without any knowledge of their population labels and found that the Italian and Anatolian individuals clustered together and with those of pre-Roman Anatolia (Fig. 4B). These results suggest that the Roman Empire in both its short-lived western part and the longer-lasting eastern Empire centered on Anatolia had a diverse but similar population plausibly drawn to a substantial extent from  
30 Anatolian sources. It is an irony of history that while the Roman Republic faced an existential crisis during the Mithridatic Wars of the 1<sup>st</sup> century BCE, it was its victory in this struggle that set the stage for the history of its later Empire whose population was drawn heavily from the same Anatolian population that had so recently challenged the Roman expansion to the east. In the Augustan period at the beginning of the Roman Empire, Horace, the Latin poet, remarked on the prestige of Greek high culture when he remarked that “captured Greece captured her rude conqueror,” and a century later the satirist Juvenal remarked on the population influx from the  
35 East when he noted that “Syrian Orontes has long since flowed into the Tiber” but with regard to people it was Anatolia that became the demographic engine of Imperial Rome, realizing, in historical time, the mythical journey of Aeneas and his Trojan exiles from Anatolia to the shores  
40 of Italy.

45 The SA was also a recipient of people from outside. We detect the presence of two likely Sarmatian individuals in Samsun in the Black Sea region which contrast with the local population that was stable since the Chalcolithic(15), across the Early Bronze Age transition at Amasya, and down to the time of the Kingdom of Pontus which struggled with Rome during the Mithridatic wars. Overall population stability of Anatolia during the Roman/Byzantine period did not mean isolation, as outlier individuals of likely Levantine, northern European/Germanic, and Iberian origin are detected in the Marmara region (in the Basilica of Nicaea/present-day

Iznik and the Virgin Mary Monastery at Zeytinliada) close to the Imperial capital of Constantinople (present-day Istanbul) which may have attracted a more diverse set of foreigners. Other outliers are found at the periphery of the Southern Arc, in Moldova and Romania, in the Iron Age and long after the early steppe migrants previously discussed, distinctive because of the East Eurasian admixture of Central Asian Scythian individuals.(32-34)

East Eurasian ancestry also helps identify an intriguing set of outliers observed at Çapalıbağ in the Aegean coast of Turkey dating from the 14<sup>th</sup>-17<sup>th</sup> centuries (Fig. S1E and (5)); these have ~18% such ancestry unlike Byzantine-era individuals from Turkey (Fig. S9), suggesting a Central Asian influence. An admixture date estimate of  $12.2 \pm 1.4$  generations prior to their time using Roman/Byzantine and Central Asian sources places the admixture in the period surrounding the 11<sup>th</sup> century arrival and subsequent expansion of Seljuq Turks to Anatolia; present-day Turkish individuals have an admixture date estimate of  $30.6 \pm 1.9$  generations, and thus from the same early centuries of the 1000s CE which coincided with the transfer of control of Anatolia from the Romans to the Seljuqs and eventually Ottomans. The genetic contribution of Central Asian Turkic speakers to present-day people is difficult to gauge but can be provisionally placed at ~9-22% by comparison of Central Asian ancestry in present-day Turks and sampled ancient Central Asians. In the future, denser and targeted sampling from both Turkey and areas to its east may clarify the details of this process and more accurately quantify its extent.

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

The discoveries of massive migrations from the steppe both westwards into mainland Europe(4, 8) and eastwards into South Siberia(4) and Central/South Asia(20) have provided powerful evidence for the theory of steppe Indo-European origins by linking populations all the way from northwest Europe(27) to India via common steppe ancestry. The present paper 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(35) 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 during the Mycenaean period when the Greek language is first attested, albeit at lower levels.(23) 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 due to the absence of EHG ancestry in Anatolia(4, 10, 15, 23) which our study reinforces in three ways: (i) first, by documenting its paucity in ~100 new Anatolian samples from the Chalcolithic to pre-Roman antiquity, (ii) by contrasting western parts of Turkey with its immediate Aegean-Balkan neighbors to the west, and (iii) by contrasting eastern/northern parts of Turkey with its neighbors in Armenia in the east. The absence of EHG ancestry in Anatolia can never be categorically proven (as more sampling can always disclose *some* such ancestry), but at present, 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,(36) has been supported by more modern phylolinguistic analyses that indicate that Anatolian languages such as Hittite



are basal to the rest of the IE family tree(37), suggesting an early split between the two. We have shown that Anatolia was indeed transformed during the Chalcolithic by the spread of CHG-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 populations related to the West Asian highlanders into the steppe, which transformed its 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 large genetic impact on the population of Anatolia. The evidence is consistent with two hypotheses: Hypothesis A postulates that Proto-Indo-Anatolian (PIA; including both Anatolian languages and Proto-Indo-European) was spoken by a population with high EHG ancestry which 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, but this is a period when Anatolia is influenced by numerous linguistically non-Anatolian Indo-European populations, including Phrygians, Greeks, Persians, Galatians, Romans, to name only a few. Yet 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 amount of EHG ancestry is only ~2%, a tiny proportion 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 to the population, which 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 EHG 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 Indo-European,(38) thus potentially removing a technological advantage regarded as potentially crucial in the dissemination of Indo-European languages.(39) Hypothesis B postulates that Proto-Indo-Anatolian was spoken by a population of West Asia and the Caucasus, with low or no EHG ancestry which affected both Anatolia and the steppe. At least two migratory movements 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 1000-2000 years later by the dispersal of Indo-European languages from the steppe with the expansion of the Yamnaya culture. Linguistic borrowings(40) between Proto-Indo-European and other language families such as Kartvelian (spoken primarily in Georgia) could be useful for localizing the PIA homeland, but these may have alternatively come about by long-range mobility since the Chalcolithic, proven by such evidence as the presence of R-V1636 descendants ~3,000km apart from Khvalynsk to Anatolia during this period. Contributions of Indo-European to Uralic (spoken in the forest-zone of eastern Europe and Siberia) appears to have involved only Indo-Iranian speakers around 4,200 years ago(41), thus constraining the migratory history of Proto-Indo-Iranian, but not the deeper question of early Indo-Anatolian origins. A challenge for the theory that PIA was formed in the south in a CHG-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 PIA homeland(42, 43)) and to identify the place where the R-M269 ancestral lineage expanded from, as this will be a most plausible secondary homeland of IE expansion outside Anatolia.

5 The scenario of a West Asian source of PIA is consistent with a linguistic analysis(44) which places the split of Tocharian from the remaining (Inner Indo-European) languages ~3000BCE associated with the Yamnaya expansion and the disintegration of the remaining languages during the 3<sup>rd</sup> millennium BCE, in line 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  
10 placed by that study at ~3700BCE (4314–3450 BC, 95% Highest Posterior Density interval), a period during which the CHG ancestry first appears as far west as the Chalcolithic individuals from NW Anatolia (at Ilıncık) sampled in our study and during which the flow of CHG ancestry into the steppe had already commenced.

15 Indo-Europeans became a source of fascination during the 19<sup>th</sup> and 20<sup>th</sup> centuries after their discovery by linguists in the 18<sup>th</sup>, leading to elaborate reconstructions of their culture and a fascination with their physical appearance.(45) Our survey of ancient phenotypes from the Southern Arc and West Eurasia in general ((5); Data S4) shows that the past use of modern phenotypes, physical anthropology, and ancient literary and artistic depictions to infer a “Proto-Indo-European” physique characterized by depigmentation led to factually wrong conclusions.  
20 Early populations of both the Southern Arc and the steppe had predominantly dark hair and eyes and intermediate complexion, suggesting that this was also the most likely modal phenotype of Proto-Indo-Anatolians (and the Late Proto-Indo-European Yamnaya) as well, regardless of the solution to the question of their homeland. The SA population and the rest of West Eurasia were phenotypically diverse long prior to any Indo-European migrations (Fig. S10), but lacked the  
25 “northern European” constellation of traits (light hair, eyes, and skin pigmentation) once thought to be associated with the PIE community, an argument that sometimes continues to be made.(46) Thus, the 19<sup>th</sup> /20<sup>th</sup> century notion that the appearance of modern northern Europeans provided a clue to the ancestry of idealized “Aryans” was as misguided as the ancient idea, uninformed by a  
30 genetic theory of adaptation, that these same phenotypes were the products of the direct influence of climate.

In concluding our genetic history of the Southern Arc, we admit the limitations of the concept: just as populations at the border between “Europe” and “Asia” can best be understood by their  
35 joint analysis of populations across the geographical divide, so can the SA best be understood by its joint analysis with the complementary “Northern Arc” to the north of the Black Sea, with the still woefully undersampled regions of Arabia and North Africa to its south, and the world of mainland Europe and central Asia to its west and east. Future studies, transcending traditional categories of geography and population and revealing hidden and distant connections, will be  
40 necessary to obtain an even richer understanding of population history.

## References and Notes

1. F. Broushaki *et al.*, Early Neolithic genomes from the eastern Fertile Crescent. *Science* **353**, 499 (2016).
2. L. D. Jack, R. S. Sharon, The Lord of the Gold Rings: The Griffin Warrior of Pylos. *Hesperia: The Journal of the American School of Classical Studies at Athens* **85**, 627-655 (2016).
3. I. Mathieson *et al.*, The genomic history of southeastern Europe. *Nature* **555**, 197-203 (2018).
4. M. E. Allentoft *et al.*, Population genomics of Bronze Age Eurasia. *Nature* **522**, 167-172 (2015).
5. Detailed information is provided in the supplementary materials.
6. S. Eisenmann *et al.*, Reconciling material cultures in archaeology with genetic data: The nomenclature of clusters emerging from archaeogenomic analysis. *Scientific Reports* **8**, 13003 (2018).
7. E. R. Jones *et al.*, Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nature Communications* **6**, 8912 (2015).
8. W. Haak *et al.*, Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207-211 (2015).
9. I. Mathieson *et al.*, Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499-503 (2015).
10. I. Lazaridis *et al.*, Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419-424 (2016).
11. Z. Hofmanová *et al.*, Early farmers from across Europe directly descended from Neolithic Aegeans. *Proceedings of the National Academy of Sciences* **113**, 6886 (2016).
12. G. M. Kilinc *et al.*, The Demographic Development of the First Farmers in Anatolia. *Curr Biol* **26**, 2659-2666 (2016).
13. M. Feldman *et al.*, Late Pleistocene human genome suggests a local origin for the first farmers of central Anatolia. *Nature Communications* **10**, 1218 (2019).
14. R. Yaka *et al.*, Variable kinship patterns in Neolithic Anatolia revealed by ancient genomes. *Current Biology*, (2021).
15. E. Skourtanioti *et al.*, Genomic History of Neolithic to Bronze Age Anatolia, Northern Levant, and Southern Caucasus. *Cell* **181**, 1158-1175.e1128 (2020).
16. O. Barge *et al.*, Diffusion of Anatolian and Caucasian obsidian in the Zagros Mountains and the highlands of Iran: Elements of explanation in 'least cost path' models. *Quaternary International* **467**, 297-322 (2018).
17. C.-C. Wang *et al.*, Ancient human genome-wide data from a 3000-year interval in the Caucasus corresponds with eco-geographic regions. *Nature Communications* **10**, 590 (2019).
18. É. Harney *et al.*, Ancient DNA from Chalcolithic Israel reveals the role of population mixture in cultural transformation. *Nature Communications* **9**, 3336 (2018).
19. A. F.-H. Egjford *et al.*, Genomic Steppe ancestry in skeletons from the Neolithic Single Grave Culture in Denmark. *PLOS ONE* **16**, e0244872 (2021).
20. V. M. Narasimhan *et al.*, The formation of human populations in South and Central Asia. *Science* **365**, eaat7487 (2019).
21. L. Saag *et al.*, Genetic ancestry changes in Stone to Bronze Age transition in the East European plain. *Science Advances* **7**, eabd6535 (2021).
22. J. Diakonoff, S. A. e. Starostin, Hurro-Urartian as an Eastern Caucasian Language. *Muenchener Studien zur sprachwissenschaft*, (1986).

23. I. Lazaridis *et al.*, Genetic origins of the Minoans and Mycenaeans. *Nature* **548**, 214-218 (2017).
24. D. Adamov, V. M. Gurianov, S. Karzhavin, V. Tagankin, V. Urasin, Defining a New Rate Constant for Y-Chromosome SNPs based on Full Sequencing Data. *Russian Journal of Genetic Genealogy* **7**, 1920-2997 (2015).
- 5 25. H. Ringbauer, J. Novembre, M. Steinrücken, Parental relatedness through time revealed by runs of homozygosity in ancient DNA. *Nature Communications* **12**, 5425 (2021).
26. B. E. Preda-Bălănică, A. Frînculeasa, V. Heyd, The Yamnaya Impact North of the Lower Danube. *A Tale of Foreigners and Locals*. **117**, 85-101 (2020).
- 10 27. I. Olalde *et al.*, The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature* **555**, 190-196 (2018).
28. I. Olalde *et al.*, The genomic history of the Iberian Peninsula over the past 8000 years. *Science* **363**, 1230 (2019).
29. M. Feldman *et al.*, Ancient DNA sheds light on the genetic origins of early Iron Age Philistines. *Science Advances* **5**, eaax0061 (2019).
- 15 30. F. Clemente *et al.*, The genomic history of the Aegean palatial civilizations. *Cell*, (2021).
31. M. L. Antonio *et al.*, Ancient Rome: A genetic crossroads of Europe and the Mediterranean. *Science* **366**, 708 (2019).
32. M. Unterländer *et al.*, Ancestry and demography and descendants of Iron Age nomads of the Eurasian Steppe. *Nature Communications* **8**, 14615 (2017).
- 20 33. P. d. B. Damgaard *et al.*, 137 ancient human genomes from across the Eurasian steppes. *Nature* **557**, 369-374 (2018).
34. M. Krzewińska *et al.*, Ancient genomes suggest the eastern Pontic-Caspian steppe as the source of western Iron Age nomads. *Science Advances* **4**, eaat4457 (2018).
- 25 35. J. P. Mallory, *In search of the Indo-Europeans : language, archaeology, and myth / J.P. Mallory*. (Thames and Hudson, New York, N.Y., 1989).
36. E. H. Sturtevant, On the Position of Hittite among the Indo-European Languages. *Language* **2**, 25-34 (1926).
37. D. Ringe, T. Warnow, A. Taylor, Indo-European and Computational Cladistics. *Transactions of the Philological Society* **100**, 59-129 (2002).
- 30 38. M. J. Kümmel, paper presented at the European Association of Archaeologists (EAA), Bern, 2019.
39. D. W. Anthony, *The Horse, the Wheel, and Language: How Bronze-Age Riders from the Eurasian Steppes Shaped the Modern World*. (Princeton University Press, 2010).
- 35 40. D. W. Anthony, D. Ringe, The Indo-European Homeland from Linguistic and Archaeological Perspectives. *Annual Review of Linguistics* **1**, 199-219 (2015).
41. J. Nichols, The Origin and Dispersal of Uralic: Distributional Typological View. *Annual Review of Linguistics* **7**, 351-369 (2021).
42. 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*. (1995).
- 40 43. S. A. Grigoriev, *Ancient Indo-Europeans*. (RIFEI, 2002).
44. 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).
- 45 45. L. Poliakov, *Le mythe aryen. essai sur les sources du racisme et du nationalisme*. (Calmann-Levy, Paris, 1971).
46. J. V. Day, *Indo-European origins : the anthropological evidence*. (Institute for the Study of Man, Washington, D.C., 2001).

**Acknowledgments:** We are grateful to the authorities and sample stewards including museums, museum curators, and archaeologists, for providing official permits to sample human remains and carry out DNA analysis in diverse contexts in Turkey, Armenia, Croatia, Iraq, Romania, Bulgaria, Greece, Albania, North Macedonia, Montenegro, Moldova, Serbia, Cyprus, Jordan, Bosnia and Herzegovina. We acknowledge the many ancient individuals whose permission we could not directly ask, but who we studied after consulting with the institutions and individuals who could best represent them and the modern communities that may feel connected to them based on geography or shared culture. We thank the many students whose names we cannot list here, the workers on the excavations, and the numerous local institutions and private individuals who supported this challenging work, without which this study would not have been possible. We thank M. Wiener, J. Bennett, J. Sedig, A. Shaus, and L. Vokotopoulos for critical comments; D. Mitrevski, M. Pantelidou-Gofa for archaeological work; H. Ringbauer for discussions on the hapROH analysis; and N. Adamski for lab work. Most of the samples from Albania (sites of Podgorie; Tren Cave 2; Dukat; Çinamak; Kënetë; Bardhoc; 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 Gregory Areshian, Milos Bilbija, Edgar Peltenburg, who would have been our co-authors if they had not passed away during the preparation of this study.

### **Funding:**

The ancient DNA work for this study was supported by the National Institutes of Health (NIGMS GM100233), the John Templeton Foundation (grant 61220), a private gift from Jean-Francois Clin, and by the Allen Discovery Center program, a Paul G. Allen Frontiers Group advised program of the Paul G. Allen Family Foundation and the Howard Hughes Medical Institute (DR)

European Research Council (ERC) Starting Grant Project HIDDEN FOODS (Grant No. 639286). (ECr)

Grant of the Romanian Ministry of Research and Innovation, CCCDI – UEFISCDI, project number 351PED (PN-III-P2-2.1-PED-2019-4171), within PNCDI III. (CL)

Financial support of the FF-MAC project (Face to Face: Meet an Ancient Cypriot; Grant No. INTEGRATED/0609/29) and the BioMERA project (Platform for Biosciences and Human Health in Cyprus: MicroCT and Synchrotron Radiation Enabled Analyses; Grant No. INFRASTRUCTURES/1216/09) co-financed by the European Regional Development Fund and the Republic of Cyprus through the Research and Innovation Foundation. (KOL)

Grant of the Hungarian Research, Development and Innovation Office (FK128013). (TH, TSz, KK)

Croatian Science Foundation grant HRZZ IP-2016-06-1450. (MN, IJ, JB)

Grant NCN 2015/17/B/HS3/01327 (PW)

## Author Contributions:

Conceived the study: ILa, SA, RP, DR

Supervised the study: SA, DKen, NPat, NR, RP, DR

Assembled archaeological material and prepared the site descriptions: SA, AAca, AAçı, AAg,  
5 LA, UA, DAnd, GA, DAnt, IA, AAt, PA, AIA, KBa, RBa, JB, LB, ABe, HB, ABi,  
MBod, MBon, CB, DB, NB, MCa, SCho, M-EC, SChr, IC, NC, MCo, ECr, JD, TID,  
VD, ZD, SDes, SDev, VDj, NEI, AE, NEr, SE-P, AF, MLG, BGas, BGay, EG, TG, SG,  
TH, VH, SH, NH, Ill, SI, İİv, IJ, LJ, PKa, BKK, EK, SDK, AK, KK, SKı, PKI,  
10 SKBNV, SKo, MK-N, MKŠ, RK, PKu, CL, KLe, TEL, ILi, KLo, SŁ, KM-O, RM,  
WM, KMc, VM, LM, DMi, JMM, GN, RN, AGN, VN, MN, AO, CÖ, NÖ, NPap,  
NPap, APa, LPa, EPa, IP, LPe, VP, APh-T, APi, NPocK, HP, BP-B, ZP, DP, SRad,  
KRA, PRŠ, KRR, SRaz, AR, JR, RR, VR, MŞa, AŞar, ES, AS LS, TSe, AS-E, MSh-P,  
HMS, ASid, ASim, HS, VS, GS, MŠl, ASof, BS, ASoł, ÇS, MSta, MSt, SS, FSA, AS-  
15 N, TSz, NTe, STe, NTo, UT, GT, STr, AT, MU, FV, ZV, CV, SW, PW, AYar, EY,  
AYYav, HY, RÇ, RP

Performed laboratory work: SA, GBM, KBu, KC, FC, BJC, ECu, KSDC, LRE, DMF, MF,  
SF, BGam, LI, DKea, AML, KMa, MMi, JO, KTO, LQ, CS, KSi, KSt, AW, JNW, FZ,  
AZ, NR

Performed population genetic analyses: ILa, DR

20 Analyzed data: ILa, SA, RBe, OC, MMA, SM, AMic, AMit, IO, ZZ, NR, DR

Wrote the manuscript and compiled the supplementary sections with the input of all other co-  
authors: ILa, SA, DR

**Competing interests:** The authors declare that they have no competing interests.

**Data and Materials availability:** Genotype data for individuals included in this study can be  
25 obtained from the Reich Lab datasets page (<https://reich.hms.harvard.edu/datasets>). BAM files of  
aligned reads can be obtained from the European Nucleotide Archive (Accession number xxx).

## 30 Supplementary Materials:

Materials and Methods

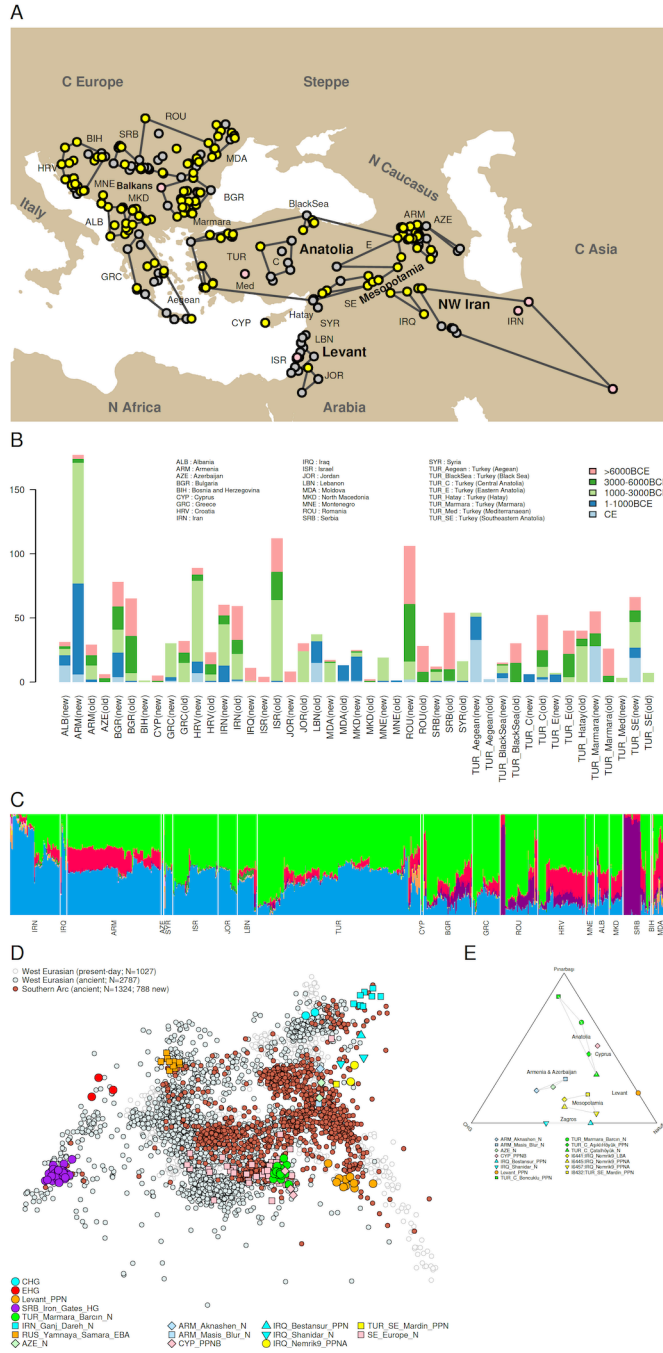
Supplementary Text, S1-9

Figs. S1 to S85

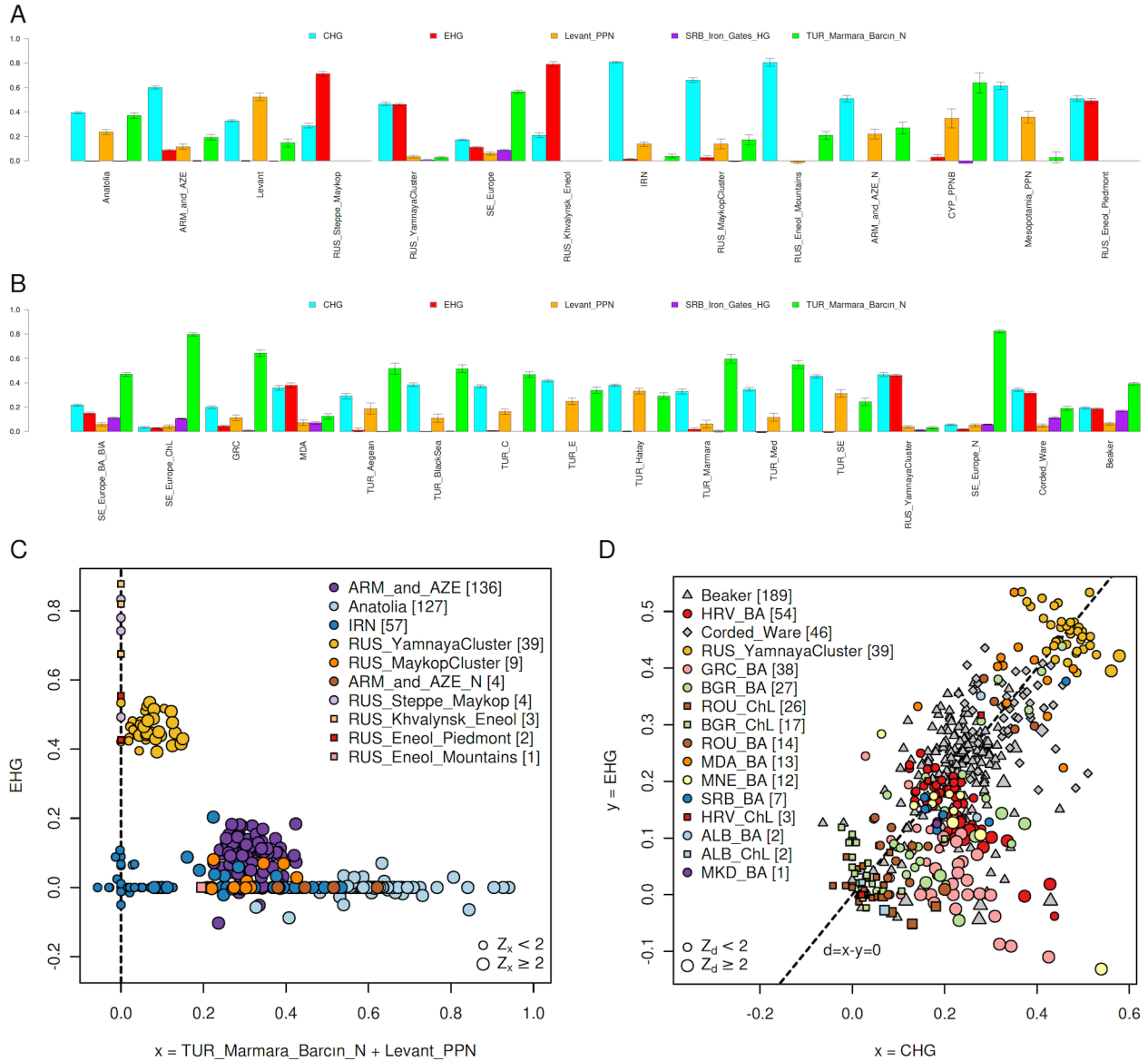
Tables S1 to S44

35 References (47–484)

Data S1 to S4

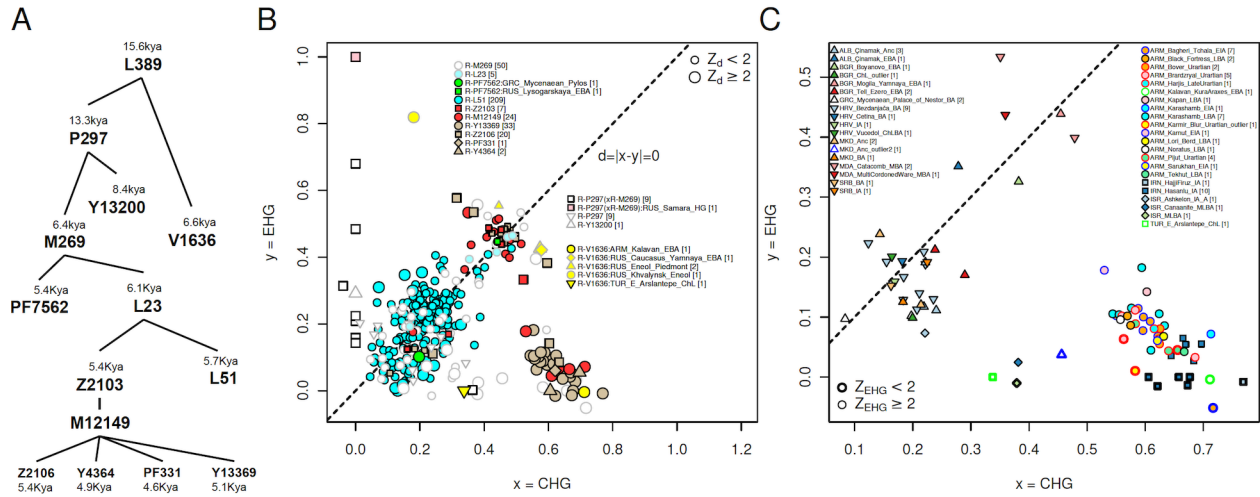


**Fig. 1. Studied individuals, PCA, and ADMIXTURE analysis.** (A) The geography of the Southern Arc; Sampling locations of previously published individuals (grey), new data on published individuals (pink), new individuals (yellow); convex hulls of individuals from each present-day country are also shown. (B) Correspondence of ISO codes with names and proportion of newly reported individuals and time histogram. (C) ADMIXTURE analysis: blue (Caucasus), green (Anatolia), red (Eastern European hunter-gatherers), purple (Balkan hunter-gatherers). (D) Principal components analysis. (E) Ancestry in Neolithic West Asia from pre-Neolithic Anatolian (Pınarbaşı), Caucasus (CHG), and Levant (Natufian) sources.



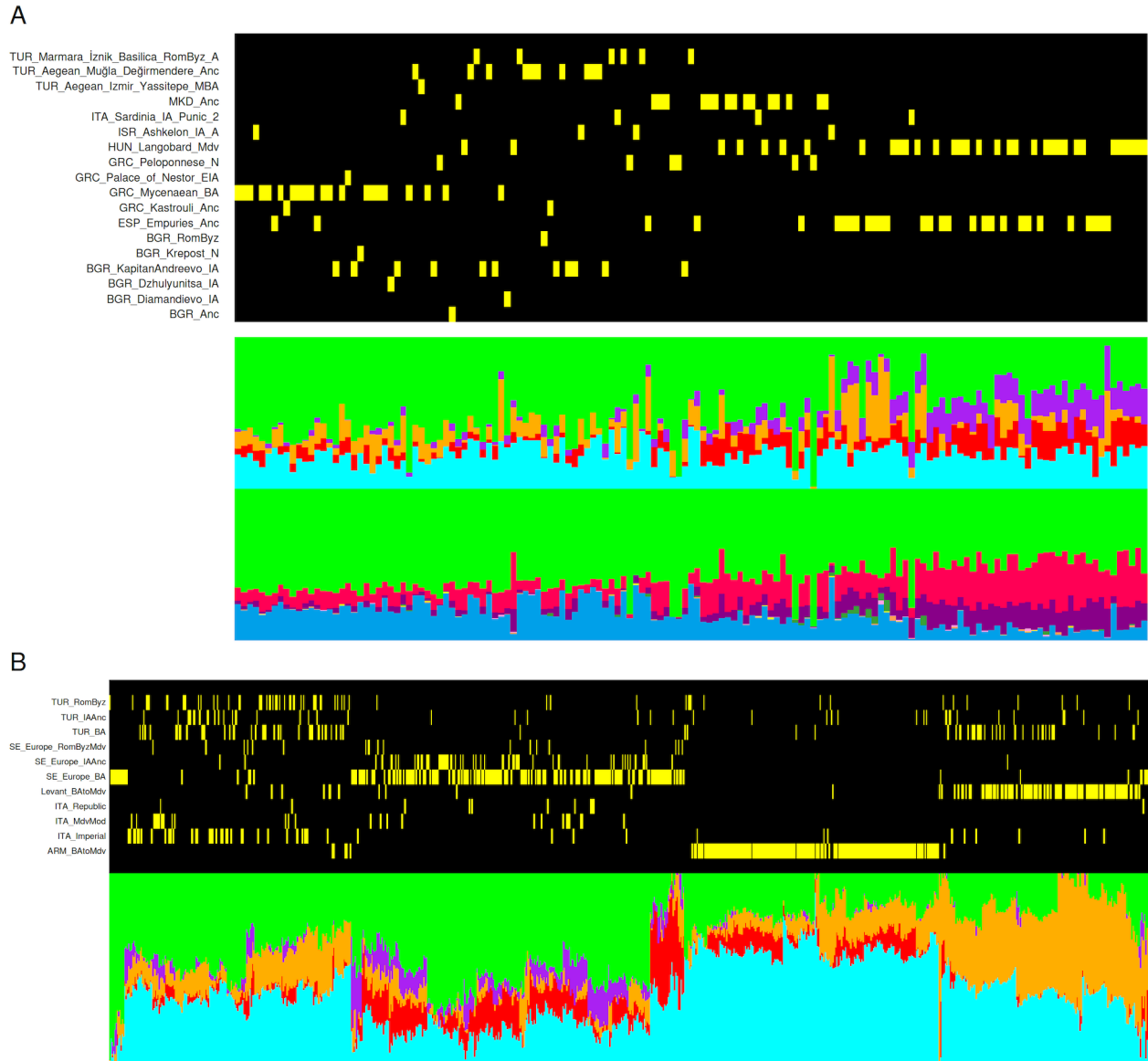
**Fig. 2: Spread of CHG and EHG ancestry during the Chalcolithic, Bronze, and Early Iron Age.** (A) Along the *Caucasus bridge* CHG ancestry flowed northwards to the steppe and EHG ancestry spread southward into SE Europe and Armenia but not into Anatolia or the Levant. (B) Along the *Anatolian bridge* CHG ancestry exceeds EHG ancestry in SE Europe and Anatolia, and EHG ancestry spread into SE Europe with a maximum in Moldova, a minimum in the Aegean, and hardly any in Anatolia. (C) Variation along the *Caucasus bridge* shows that Armenia/Iran individuals experienced an influx of EHG ancestry and Yamnaya cluster individuals an influx of Anatolian/Levantine ancestry. (D) Individuals from SE Europe at the western end of the *Anatolian bridge* have an excess of CHG over EHG ancestry.





**Fig. 3: Y-chromosome links between the Southern Arc and the Eurasian steppe. (A)**

Phylogeny of haplogroup R-L389 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 panel b. Individuals more than 2,000 years old are shown.



**Fig. 4: Expansions from the Southern Arc after the Bronze Age.** (A) In the Iron Age, individuals similar to the Bronze Age Mycenaeans were found in Greece, North Macedonia, and Bulgaria and can be detected from Spain to Lebanon to Hungary as outliers. Individuals are ordered left-right by their rank of similarity to Mycenaeans with respect to their proportions of their five inferred ancestry components. Individuals (yellow marks on black background) are shown in conjunction with F4admix and ADMIXTURE proportions at the bottom. (B) Roman/Byzantine period individuals from Anatolia and Imperial Italy(31) were diverse and genetically indistinguishable from each other, and drawn primarily from the preceding population of Anatolia.

Topic	Background	Genetic inferences
What was the ancestry of the world's very first farmers from Mesopotamia?	Three early groups of farmers (from Anatolia, the Zagros, and the South Levant) were highly genetically differentiated. Did early farmers from Northern Mesopotamia resemble any of these groups?	We report the first PPN individuals from the SE Anatolia/N Mesopotamia region, including from the previously unsampled PPNA period. These were unlike any of the three previously known groups and genetically intermediate between populations of the Levant and Caucasus/Iran. Together with other new data from early Neolithic individuals of Armenia, the north Zagros, and Cyprus, these suggest the existence of a Neolithic continuum between the highly differentiated populations at the edges of ancient West Asia. This continuum was formed as Natufian ancestry spread via Mesopotamian intermediaries into the rest of West Asia. Anatolian Epipaleolithic and Mesopotamian ancestry alone formed the pre-pottery Neolithic of Anatolia, but pottery Neolithic populations had extra Levantine ancestry.
Whence the Anatolian languages?	Did Anatolian languages like Hittite and Lydian spread from the steppe as is likely for other Indo-European (IE) languages of Europe and South Asia?	We sampled across Anatolia and the immediately neighboring regions of Armenia and Aegean/South Balkans during the Chalcolithic and Bronze Age (where putative steppe ancestors may have first arrived) and later periods (where any such ancestry would have diffused into the general population). There was a distinction between Anatolia and its immediate neighbors marked by absence/presence of EHG ancestry typical of the steppe. By contrast, steppe pastoralists had experienced at least two gene flows from the south prior to the Yamnaya expansion. Anatolian languages likely did not originate on the steppe but in a West Asian population without EHG ancestry, and steppe expansions were responsible only for the dispersal of non-Anatolian IE languages.
Reverberations of the Great Steppe Takeover across the Caucasus	Yamnaya-cluster steppe populations ~3000BCE were dominated by Y-chromosome haplogroup R-M12149, but these were close to wholly replaced by R-Z93 bearers by ~2000BCE.	Bronze and Iron Age individuals from Armenia and NW Iran have (i) autosomal steppe ancestry, (ii) R-M12149 as their most frequent Y-chromosome haplogroup, and (iii) a ~2500BCE admixture date estimate. Patrilineal descendants of the Yamnaya—whether independently or pressured by the expansion R-Z93 bearers—must have left the steppe and settled south of the Caucasus. Remarkably, while this expansion did not have the genetic impact of those into Europe and South Asia it was the only one that preserved the patriline of the Yamnaya population.
The origin of Paleo-Balkan speakers	Southeastern Europe was populated by groups speaking numerous Indo-European languages (e.g., Illyrians and Thracians) that were replaced, except Albanian, by exogenous Latin and Slavic languages. What accounts for this linguistic diversity?	We show that people genetically indistinguishable from Yamnaya pastoralists penetrated deep into the Balkans in Bulgaria, Serbia, and Albania during the early Bronze Age. Admixture dating of individuals from the Balkans is consistent with a major pulse of admixture ~3000BCE coinciding with the Yamnaya expansion. Steppe ancestry drastically diminishes north-to-south and is blended with varying amounts of indigenous hunter-gatherer and Anatolian Neolithic ancestry, but also with eastern ancestry from West Asia. This remarkable genetic heterogeneity contrasts with the geographically expansive Corded Ware and Beaker phenomena of mainland Europe, and may explain the diverse linguistic milieu of ancient Southeastern Europe.
The Kingdom of Urartu	An extinct non-IE Urartian language is recorded in the Iron Age Kingdom of Urartu. The IE Armenian language was in early contact with this language. Who inhabited the Urartian Kingdom?	The Kingdom of Urartu in the Iron Age had a diverse population marked by absence (Çavuştepe, in Lake Van in Turkey) or presence (in Armenia) of steppe ancestry and greater/lesser Levantine ancestry respectively. The population from the central Lake Van part of the Kingdom was in continuity with the pre-Urartian population (Muradiye) and had a West Asian genetic composition suggesting it spoke the local Urartian language of the Hurro-Urartian family, while the peripheral Urartian period population of Armenia may trace the steppe-admixed early Armenian speakers from the north.
Elites and commoners in the Late Bronze Age Aegean	Did elites of the Mycenaean world represent exogenous influences or were they locally drawn?	We show that the genetic pattern of the Mycenaeans was not limited to individuals from the Peloponnese but included central Greece and as far north as Thrace. In the Palace of Nestor in Pylos there is some evidence of consolidation of power (people who were genetically related or the offspring of cousin marriage were buried there), and exogenous links (a pair of 1 <sup>st</sup> degree relatives having a patriline of steppe origin), but the autosomal ancestry was indistinguishable on average from the rest of the Mycenaean world.
The Roman Empire in West and East	Was there any relationship between the Imperial Romans of Italy prior to the Fall of Rome in the 5 <sup>th</sup> c. CE and the people who called themselves Roman in the eastern Roman Empire but are usually called “Byzantine” by modern historians?	Roman and Byzantine era Anatolians were diverse but genetically indistinguishable from Imperial Romans. Both were drawn primarily from people that were similar to the inhabitants of Anatolia prior to the Roman conquest. The Empire was Roman in name and institutions but substantially Anatolian in population in west and east and from the time of its founding to its demise.
The origin of Anatolian Turks	Anatolian Turkish people are linguistically related to Central Asian Turkic speakers, but genetically cluster with European and West Asian populations. What are their origins?	Ancient individuals likely speaking Turkic languages from Anatolia and previously published present-day Anatolians show that they have admixture of both pre-Turkic Anatolians and Central Asians dating to the early 1000s CE. The data point to Anatolian Turks carrying the genetic legacy of both ancient people who lived in Anatolia for thousands of years covered by our study and also people coming from Central Asia and bearing Turkic languages who contributed ~9-22% of their ancestry.

**Table 1: Archaeological, genetic, and linguistic questions newly informed by our study.**