Received: 14 December 2016 | Revised: 19 July 2017 | Accepted: 10 September 2017

DOI: 10.1002/ajpa.23324

RESEARCH ARTICLE

WILEY American Journal of PHYSICAL ANTHROPOLOGY

Genetic diversity in Svaneti and its implications for the human settlement of the Highland Caucasus

Aram Yardumian[1,2*] | Ramaz Shengelia[3*] | David Chitanava[4] | Shorena Laliashvili[4] | Lia Bitadze[4] | Irma Laliashvili[4] | Fernando Villanea[5] | Akiva Sanders[2] | Andrew Azzam[2] | Victoria Groner[2] | Kristi Edleson[2] | Miguel G. Vilar[2,6] | Theodore G. Schurr[*2]

[1]Department of History and Social Sciences, Bryn Athyn College, Pennsylvania 19009

[2]Department of Anthropology, University of Pennsylvania, Philadelphia, Pennsylvania 19104

[3]Department of the History of Medicine and Bioethics, Tbilisi State Medical University, Tbilisi 01747, Georgia

[4]Laboratory for Anthropologic Studies, Ivane Javakhishvili Institute of History and Ethnology, Tbilisi 0102, Georgia

[5]Grant Programs, Science and Exploration, National Geographic Society, Washington, DC 20036

[6]Grant Programs, Science and Exploration, National Geographic Society, Washington, DC 20036

Correspondence
Aram Yardumian, Ph.D., Department of History and Social Sciences, Bryn Athyn College, Bryn Athyn, PA 19009, USA.
Email: aram2@sas.upenn.edu

Funding information
National Science Foundation, Grant/Award Number: BCS-1249281; University of Pennsylvania Museum of Archeology and Anthropology; University of Pennsylvania Department of Anthropology; American Philosophical Society; Penn Faculty Research Funds Bryn Athyn College Carpenter Scholars Program, Bryn Athyn College

Abstract

Objectives: In this study, we characterized genetic diversity in the Svans from northwestern Georgia to better understand the phylogeography of their genetic lineages, determine whether genetic diversity in the highland South Caucasus has been shaped by language or geography, and assess whether Svan genetic diversity was structured by regional residence patterns.

Materials and Methods: We analyzed mtDNA and Y-chromosome variation in 184 individuals from 13 village districts and townlets located throughout the region. For all individuals, we analyzed mtDNA diversity through control region sequencing, and, for males, we analyzed Y-chromosome diversity through SNP and STR genotyping. The resulting data were compared with those for populations from the Caucasus and Middle East.

Results: We observed significant mtDNA heterogeneity in Svans, with haplogroups U1-U7, H, K, and W6 being common there. By contrast, ~78% of Svan males belonged to haplogroup G2a, with the remainder falling into four other haplogroups (J2a1, I2, N, and R1a). While showing a distinct genetic profile, Svans also clustered with Caucasus populations speaking languages from different families, suggesting a deep common ancestry for all of them. The mtDNA data were not structured by geography or linguistic affiliation, whereas the NRY data were influenced only by geography.

Discussion: These patterns of genetic variation confirm a complex set of geographic sources and settlement phases for the Caucasus highlands. Such patterns may also reflect social and cultural practices in the region. The high frequency and antiquity of Y-chromosome haplogroup G2a in this region further points to its emergence there.

KEYWORDS
Georgia, Svan, Kartvelian, haplogroup, haplotype, mtDNA, Y-chromosome

1 | INTRODUCTION

The Svan people dwell in a geographically remote, though not culturally isolated, highland region of northwest Georgia (Figure 1). Their habitus, Svaneti, which sits in the upper valleys of the Enguri and Tskhenistskali Rivers, has only recently been reachable via paved roads (Supporting Information Fig. S1). The Svan language is an outlier in the Kartvelian Family, and one not mutually intelligible with the other living members (Klimov, 1969). Its status as the oldest and most differentiated of the Kartvelian languages suggests that it first split off from the proto-

*These authors contributed equally to the study.

FIGURE 1 An administrative map of the country of Georgia (http://www.mappery.com/map-of Georgia-Administrative-Map). The key in the upper right corner provides details about the provinces within the country. Fieldwork was conducted in the northern portion of the khaki green colored province, Samgrelo-Zemo Svaneti

Kartvelian speech-group as early as the third millennium BCE (Chirikba, 2008; Supporting Information Fig. S2).

Very little is known about the population biology or history of Svans. Pre-Neolithic archaeological assemblages have not yet been found in Svaneti, and only sparse scattered scrapers and blades attest to Neolithic human presence in the region (Chartolani, 1967). Cultural continuity is thought to begin with the presence of copper spearheads and axes (some supposedly with Kura-Arax affinities), linking Iron Age Georgia to its later periods (Chartolani, 1967). The first mention of Svans in Classical literature occurs in Strabo, in his descriptions of the seaside emporium, Dioscurias (modern day Sukhumi), and the various peoples who traded there and "hold possession of the heights of the Caucasus above Dioscurias" (Strabo XI: II: 19). That the Svans are mentioned as a distinct people around the time of the Advent, established in the same general area as they are today, is significant.

Of the few genetics studies that have featured Svan samples (Alfonso-Sánchez et al., 2008; Sánchez-Velasco and Leyva-Cobián, 2001; Tarkhnishvili, Gavashelishvili, Murtskhvaladze, Gabelaia, & Tevzadze, 2014; Wells et al., 2001), none directly investigated their population history, or examined potential relationships between Svans and their neighbors in the Caucasus. These studies did, however, reveal that Svans have broad genetic affinities with European, Caucasus, and Near East populations, although they were unable to resolve questions stemming from Svaneti's relative isolation from lowland Georgia and highland proximity to neighbors speaking languages in different families.

Although geographically and linguistically remote, Svans are nonetheless the physical neighbors to a great number of cultural traditions, nations, and languages (Supporting Information Fig. S3). To the north, beyond the high Caucasus peaks, are the Russian oblasts of Karachay-Circassia and Kabardino-Balkaria, whose national languages are divided between the Altaic (Karachay, Balkar) and Northwest Caucasian (Circassian, Kabardine) language families. To the west is Abkhazia, a Black Sea coast breakaway region whose language clusters with Abaza, Adyghe, Circassian, and the extinct Ubykh to form the Northwest Caucasian family (Chirikba, 2008). To the east and southeast is the Kartvelian-speaking Georgian province of Racha-Lechkhumi, which Svans have sporadically inhabited, and to the south Samegrelo (Mingrelia). Beyond Racha-Lechkhumi to the east lies South Ossetia, a breakaway region from Georgia's Shida Kartli district whose inhabitants speak an Indo-European language.

The linguistic distinctiveness of Svans within an otherwise very biologically diverse region raises some interesting questions. What are the phylogeographic roots of the genetic lineages present in Svans and what do they reflect about the prehistoric settlement of the South Caucasus? Is genetic diversity in Svaneti and the South Caucasus more strongly shaped by language or geography? At a local level, was Svan genetic variation structured by regional residence within Svaneti or by clan ancestry? Our research brings new genetic insights to this little

known and remote region of Georgia, and to the history of the culturally diverse Caucasus region in which it is situated.

2 | MATERIALS AND METHODS

2.1 | Sample and data collection

In August 2012, we conducted ethnographic fieldwork in village districts and townlets throughout Upper Svaneti. During this month, we enrolled in the study a total of 184 Svan adults from Mestia, Laghami, Mulakhi, Kala, Ipari, Ushguli, Tskhumari, Becho, Etseri, Latali, Chuberi, Khaishi, and Idliani (Supporting Information Fig. S1). All but two participants were unrelated through three generations, with these individuals sharing a maternal ancestor; one of the identical haplotypes was excluded from the statistical analysis. In addition, sixteen unrelated Georgians from Tbilisi provided samples for the study, bringing the total number of participants to 200. Of this total, 103 were males, with 96 being Svan and seven Georgian in their ancestry.

Written informed consent was obtained with Georgian-language consent forms prior to the collection of buccal swabs and genealogical information from each participant. During interviews, participants were asked to identify their age and birthplace, their parents' names, ethnicity, and birthplaces, as well as their four grandparents' names, ethnicity, and birthplace, when known. Such information, along with extended genealogical interviews conducted with a subset of participants, yielded important details about the demography and history of Svan villages and the region of Svaneti as a whole. All unrelated males and females from each Svan settlement were encouraged to participate. However, when working with men, emphasis was placed on obtaining samples from individuals having different Svan surnames, as they might represent different patrilineal clan-affiliated lineages.

The sampling and consent procedures were carried out with permission of the University of Pennsylvania IRB #8 and the Georgian National Council on Bioethics.

2.1.1 | Comparative populations

To explore the genetic affinities of Svans, we compared their mtDNA and Y-chromosome data to those of populations from the Caucasus and Middle East.

The mtDNA data were drawn from Abkhazians, Armenians, Circassians, Georgians, Kabardians, North and South Ossetians (Nasidze et al., 2004); Anatolians (Gökçümen et al., 2011); and Iranians (Terreros et al., 2011). For this analysis, we evaluated the genetic affinities of the study populations based on the frequencies of haplogroups C, D, H, I1a, I1c, J, K, M, R0a1a, T*, T1, T2, U*, U1, U2, U3, U4, U5, U5a, U6, U7, W*, W6, X2*, X2b-e, and X4, which were derived from extended HVS1 sequences (np 16,024–16,400; Supporting Information Table S1). For Georgians, we used mtDNA sequence data reported in summary form by Nasidze et al. (2004) that were kindly provided by Dr. Mark Stoneking from the Institute of Evolutionary Anthropology in Leipzig, Germany.

Comparative NRY data were drawn from Abkhazians, Circassians, and North Ossetians (Balanovsky et al., 2011), Anatolians (Gökçümen et al., 2011), and Armenians (Herrera et al., 2012), and Georgians (Georgia DNA Project, 2017), with the latter representing all areas of Georgia except Samtshke-Javakheti. Statistical analysis was conducted on the basis of the relative proportions of NRY haplogroups (G2a, I2a, J2a, J1, N, R1a, and R1b), as ascertained from Y-SNP and Y-STR data, and with the Y-STR haplotype data themselves (Supporting Information Table S2). Data for the Kabardian and South Ossetian populations were excluded from the analysis due to the low number of STR loci used to characterize their Y-chromosomes (Nasidze et al., 2004). Unfortunately, we were unable to find appropriately annotated data for Iranian populations (e.g., Grugni et al., 2012) for inclusion in the analysis. In addition, Herrera et al. (2012) did not indicate sample sizes for Y-STR haplotypes in Armenians, thus forcing us to assume, however improbable, that all such haplotypes appeared only once in this population.

2.2 | Genetic analysis

2.2.1 | DNA preparation

Genomic DNAs were extracted from buccal swab samples using Puregene Blood Core B kits (Qiagen), using a slightly modified version of the manufacturer's protocol.

2.2.2 | mtDNA analysis

For all 200 samples, the entire mtDNA control region (CR), which encompasses hypervariable regions 1 and 2 (HVS1 and HVS2), was PCR amplified and sequenced using published methods (Gaieski et al., 2011; Schurr et al., 2012; Zhadanov et al., 2010). All polymorphic nucleotides were reckoned relative to the revised Cambridge reference sequence (Anderson et al., 1981; Andrews et al., 1999) and the Reconstructed Sapiens Reference Sequence (Behar et al., 2012). The CR sequence data defined maternal haplotypes in these individuals, and all haplogroups were ascertained relative to existing mtDNA databases (e.g., Phylotree version 17, Kloss-Brandstätter et al., 2011; van Oven and Kayser, 2009).

2.2.3 | Y-chromosome analysis

The Y-chromosome of each male participant was characterized using several methods. To identify paternal haplotypes, we surveyed phylogenetically informative SNPs in the nonrecombining region of the Y-chromosome (NRY; Cruciani et al., 2011; Francalacci et al., 2013; Karafet et al., 2008). Most of the SNPs and fragment length polymorphisms were characterized using custom ABI TaqMan® assays (Gaieski et al., 2011; Schurr et al., 2012; Zhadanov et al., 2010). The SNPs included (MI68, M89, M45, M9, M201, P15, M304, M69, MI70, M207, M96, and LLY22), with these markers being characterized using custom TaqMan assays read on an ABI Prism® 7900 HT Real-Time PCR System. Paternal haplotypes were further defined through the analysis of male samples with 17 short tandem repeat (STR) loci in the ABI AmpFℓSTR Y-filer® PCR Amplification Kit (DYS19, 385I, 385II, 389I, 389II, 390, 391, 392, 393, 437, 438, 439, 448, 456, 458, and 635 and GATA H4), as previously described (Gaieski et al., 2011; Schurr et al., 2012; Zhadanov et al., 2010). A separate multiplex reaction was used to

characterize six additional fragment length polymorphisms (M17, M60, M91, M139, M175, M186) and two additional Y-STRs (DYS388, DYS426; Gaieski et al., 2011; Schurr et al., 2012; Zhadanov et al., 2010). These markers were run on a 3130xl Genetic Analyzer and read with GeneMapper ID v3.2 software.

The assignment of each male sample to a Y-chromosome SNP haplogroup followed the conventions outlined by the Y Chromosome Consortium (2002) and Karafet et al. (2008) and detailed in PhylotreeY (van Oven et al., 2014). All Y-STR haplotypes were also checked for their haplogroup status using Athey's Haplogroup Predictor (http://www.hprg.com/hapest5/) and Nevgen Y-DNA Haplogroup Predictor (http://www.nevgen.org/). The combination of SNPs and STR alleles defined haplogroups and haplotypes, respectively, for each male individual.

2.3 | Data analyses

2.3.1 | Statistical analysis

We calculated basic descriptive statistics for the mtDNA and Y-chromosome data using Arlequin v3.11 (Excoffier and Lischer, 2010). These statistics included the nucleotide diversity, gene diversity, haplotype sharing, mean pairwise differences for mtDNA sequences, and Y-STR haplotypes. In addition, Tajima's D (Tajima and Nei, 1984) and Fu's F_S (Fu, 1997) were calculated from mtDNA sequences to assess the possible influence of selection on genetic variation. To assess the influence of language and geography on genetic differentiation in the Caucasus and the greater Near East, we performed analyses of molecular variation (AMOVA; Excoffier and Lischer, 2010) based on frequencies of both mtDNA and NRY haplotypes. Similarly, the mtDNA and Y-STR data were used to conduct Mantel tests in Arlequin (Smouse et al., 1986) assessing the relationship between genetic and geographic distances for Svans and the comparative populations.

To better understand Svan genetic ancestry, we compared our data with those from other Eurasian populations (see below). For the mtDNA data, F_{ST} values were estimated for the Svans and comparative populations using extended HVS1 sequences and relative haplogroup frequencies with Arlequin v.3.5.1.2 (Excoffier et al., 2010). The F_{ST} estimates from these analyses were visualized through multidimensional scaling (MDS) using R version 3.1.2. For the Y-chromosome data, pairwise R_{ST} values between Svans and comparative populations, and between the Svan villages, were estimated using 17 Y-STR haplotypes (DYS388 and DYS426 were excluded to make the haplotypes commensurate with published data sets) and NRY haplogroup frequencies. The R_{ST} estimates from these analyses were also visualized through MDS using R version 3.1.2.

The Svan mtDNA and NRY haplotypes were organized by the home village of the participants to look for local patterns of genetic differentiation. For the mtDNA HVS1 sequence data, participants belonged to thirteen separate village districts. Home villages reported less than three times were regarded as too small for statistical treatment and left out of the analysis, as were those located outside Svaneti or Abkhazia. In one case, a reported home village (Jamushi) could not be geographically verified as being in either Svaneti or Abkhazia, and therefore was left out. These subtractions reduced the total number of HVS1 haplotypes analyzed by 10. The data organized in this fashion were then subjected to pairwise F_{ST} estimation and population differentiation analysis.

For the Y-STR data, male participants reported coming from 10 different village districts. As noted above, home villages reported less than three times were regarded as too small for statistical treatment and left out of the analysis, as were those located outside Svaneti or Abkhazia. These subtractions reduced the total number of Y-STR haplotypes analyzed by eight. The data organized in this fashion were then subjected to pairwise R_{ST} value estimation and population differentiation analysis.

2.3.2 | Phylogenetic analysis

To explore the phylogenetic history of genetic lineages present in Svaneti and between Svaneti and the greater Near East, we analyzed both mtDNA HVS1 sequence and Y-STR data sets with NETWORK 4.6.1.3 (Bandelt et al., 1999). For the mtDNA HVS1 sequences, the mutation-weighting scheme was based on that described in Bandelt et al. (2002), in which fast-evolving sites were given lower weights relative to other less mutable sites. All variants known to result from homopolymeric C expansions (e.g., A16182C, A16183C) or to occur at mutational hotspots in the mtDNA CR (e.g., T16519C) were excluded from the haplotypes used in this analysis. For the time to most recent common ancestor (T_{MRCA}) estimates for haplogroups, a mutation rate of one mutation per 16,667 years was used for the HVS1 region (np 16,024–16,383; Soares et al., 2009).

For the Y-STR haplotypes, we calculated T_{MRCA} estimates using ρ-statistics, with the founder haplotype being inferred as the one most central to the network (Dulik et al., 2011, 2012; Sengupta et al., 2006). We used both a pedigree-based mutation rate (one mutation per 453 years and a generation time of 25 years; Chandler, 2006; Vilar et al., 2014), and an evolutionary mutation rate (one mutation per every 2,778 years; Rootsi et al., 2012) for these estimates. All networks were visualized using Network Publisher v1.2.0.0 (Fluxus Technology).

2.3.3 | Bayesian coalescent analysis

To understand the demographic dynamics of the Svan population, we used BEAST v1.8.4 to conduct Bayesian coalescent analyses (Drummond and Rambaut, 2007). All analyses were performed with the HKY substitution model, a Γ distribution of site-specific rates, a proportion of invariant sites, and a strict molecular clock. An extended Bayesian skyline plot (eBSP; Heled and Drummond, 2008) was used as the demographic tree prior, using a 1/X population size prior. Previously published estimates for HVR-I substitution rate (16.4% replacement per million years; Soares et al., 2009) were used to translate past population dynamics into a yearly time scale. All Markov chains were run for 1×10^8 generations and sampled every 1×10^4 generations to produce posterior distributions of 10,000 samples. The companion software Tracer v1.5 (http://tree.bio.ed.ac.uk/software/tracer) was used to calculate posterior distributions and effective sample sizes (ESS) for all parameters. All analyses were duplicated and compared to ensure Markov chain convergence, with posterior estimates made from the

TABLE 1 mtDNA haplogroup frequencies in Svans and Georgians

Haplogroup	Svans		Georgians		Likely Geographic Source
	n	%	*n*	%	
C	7	3.80	2	2.90	East Eurasia
D	3	1.63	1	1.45	East Eurasia
H	33	17.90	24	34.80	West Eurasia
HV	2	1.09	0	0.00	West Eurasia
I1	3	1.63	3	4.35	North-Central Europe
J	1	0.54	3	4.35	Middle East
K	29	15.80	4	5.78	Europe/Middle East
M	1	0.54	0	0.00	East Eurasia
M1	1	0.54	1	1.45	South-West Asia
N1b	2	1.09	1	1.45	Africa/West Asia
R01a	1	0.54	0	0.00	Africa/West Asia
T	17	9.24	11	15.94	Middle East
U	3	1.63	1	1.45	West Eurasia
U1	14	7.61	0	0.00	West Eurasia
U2	11	5.98	2	2.90	West Eurasia
U3	4	2.17	2	2.90	West Eurasia
U4	4	2.17	3	4.35	West Eurasia
U5	5	2.72	4	5.78	West Eurasia
U6	3	1.63	0	0.00	West Eurasia
U7	3	1.63	1	1.45	West Eurasia
V	0	0.00	2	2.90	Europe
W	24	13.00	2	2.90	Middle East/West Asia
X2	12	6.52	2	2.90	Middle East
X4	1	0.54	0	0.00	Middle East
Total	184		69		

Note: The Svan data come from this study; the Georgian data are taken from this study and Nasidze et al. (2004).

combined samples of both Markov chains. To estimate the time the most recent common ancestor (T_{MRCA}) of the mitochondrial haplogroups present in the Svan population, we used a Yule tree prior, a strict molecular clock prior, and a molecular rate of 16.4% replacement per million years.

3 | RESULTS

3.1 | mtDNA diversity in Svaneti

3.1.1 | Haplogroup diversity

Svans exhibited a wide array of mtDNA haplogroups, with the majority being of putative West Eurasian or North African origin (Table 1). Of the 11 distinct maternal lineages present in the Svans, haplogroups H (~18%), K (~16%), W (13%), T (9.2%), and X2 (6.5%) together accounted for well over half of them. In particular, the frequency of haplogroup W6 (11%) was quite high relative to those observed in regional populations. The diversity of haplogroup U in the Svans was also remarkable inasmuch as seven of its nine major subhaplogroups (U1–U7) were present (excluding K) and comprised 25.5% of the Svan mtDNAs. Several other haplogroups with broadly Near Eastern origins (M1, N1b1, R0a1; Brandstätter et al., 2008; Fernandes et al., 2012; González et al., 2007) were also present at low percentages, while East Eurasian haplogroups C and D (Comas et al., 1998; Li et al., 2010; Torroni et al., 1993) accounted for 6% of the total. With respect to the latter two haplogroups, East Eurasian mtDNAs have also been observed at low frequency in North Caucasus populations, including both Caucasian- (e.g., Chechens, Circassians) and Turkic-speaking (e.g., Kabardians) groups (Yunusbayev et al., 2011).

Our characterization of mtDNA sequences from all 200 participants revealed a considerable diversity of haplotypes amongst them. We observed 145 distinct CR sequences (inclusive of HVS1 and HVS2), with 130 being unique to Svans, 11 unique to Georgians and 4 shared (Supporting Information Table S3). When considering only the extended HVS1 sequence (np 16,024–16,400), we observed 99 different haplotypes (Supporting Information Table S4), with 74 being exclusive to Svans and 7 unique to Georgians. Overall, Svans and Georgians shared eight HVS1 haplotypes (occurring in haplogroups H, K, T1, U2, U3, U5, and W6), with additional haplogroups appearing in both Svans and Georgians including C, J, U1, and U7. This general pattern was affirmed when the Georgian mtDNA data from Nasidze et al. (2004) were added to this comparison, with haplogroups H and T occurring at the highest frequency in Georgians (Table 1). These data pointed to the moderate divergence between Svans and Georgians in terms of their maternal lineage composition.

3.1.2 | Phylogenetic analysis of Svan mtDNA haplotypes

As noted above, Svans exhibited a high frequency of haplogroup W and, in particular, W6 mtDNAs. For this reason, we constructed a network using Svan W6 haplotypes and a second adding those from published studies (Metspalu et al., 2004; Olivieri et al., 2013; Quintana-Murci et al., 2004; Terreros et al., 2011) to explore its phylogeographic features. The resulting network of Svan haplotypes showed that W6 did not have any well-defined subclades (Figure 2). When haplotypes from neighboring populations (i.e., Iranian, Turk, Georgian) were included, Svan haplotypes usually appeared only 1–2 steps away from them (data not shown). These results suggested the broad easterly expansion of W6-bearing individuals across West Asia, perhaps emanating in or near the South Caucasus.

Using the HVS1 mutation rate, we calculated a T_{MRCA} for W6 at 20,614 YBP (Table 2). This date fell within the range of post-glacial dispersals from the Near East (Olivieri et al., 2013; Posth et al., 2016). Interestingly, the T_{MRCA} for W6 was somewhat older than the 18,500 YBP estimate from Richards et al. (1998), which was based largely on Finnish and other European data available at the time, and also the 12,600 YBP estimate by Oliveri et al. (2013) based on mitogenome sequence data.

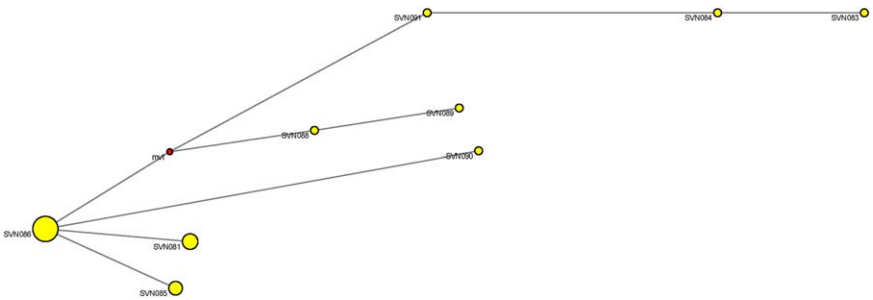

FIGURE 2 A median-joining network of mtDNA haplogroup W6 HVS1 sequences from Svans

Svans also exhibited an unusually high frequency of haplogroup X2. We therefore constructed a median-joining network for this mtDNA haplogroup (Figure 3) and calculated a T_{MCRA} for it. The resulting T_{MCRA} of 26,938 YBP fell within the range of Upper Paleolithic in Europe and the Near East, although its standard deviation of 13,117 years was quite large (Table 2). Given the wide geographic range of haplogroup X but more limited distribution of X2 (the Levant, Caucasus, and southern Europe; Reidla et al., 2003), this lineage may have arisen and diversified during Neolithic human dispersals (Hofmanová et al., 2015; Mathieson et al., 2015), although also possibly through an earlier post-LGM expansion around 17,900–21,600 YBP (Reidla et al., 2003).

We also used Bayesian MCMC analysis to estimate the T_{MCRAS} of the major mtDNA haplogroups in Svans (Table 2). While the T_{MRCAS} estimated in this way were older than the rho estimates, the former were still within the credibility intervals of the latter, which were very broad. Overall, the resulting eBSPs for haplogroups H, K, and U exhibited similar demographic profiles, with each showing a constant population size through 0–10 kya of time (Figure 4). This result, along with the Tajima's D and Fu's Fs neutrality indices for Svan mtDNA HVS1 sequences (Table 3), indicated no clear evidence for population bottlenecks or reductions in the recent past.

3.1.3 | Statistical analysis of mtDNA data

We analyzed mtDNA sequence data in Svans and comparative populations using various descriptive statistics. The Svans were relatively similar in their HVS1 sequence diversity to other populations in the Near East and Caucasus (Table 3). Gene diversity in all populations was also fairly high, while nucleotide diversity values were relatively similar across all groups. In addition, the mean pairwise differences in Svans were greater than those of Circassians or Kabardians but less than those for Abkhaz and South Ossetians. Thus, Svans fell between these two groups of populations in terms of their haplotypic diversity.

To better understand the relationship between Svans and their neighbors in the Caucasus and greater Near East, we generated F_{ST} estimates from HVS1 sequences for Svans and comparative populations (Table 4), and visualized them in a MDS plot (Figure 5). In this plot, North Ossetians were genetically similar to Georgians and Karbardians, and Armenians and Anatolians were positioned nearby in a tight cluster with each other. Otherwise, all populations, including the Svans,

TABLE 2 Coalescence time estimates for mtDNA and Y-chromosome haplogroups

(a) T_{MRCA} estimates for mtDNA haplogroups in Svans (YBP)					
		T_{MCRA} (rho)		T_{MRCA} (Bayesian)	
Haplogroup	n	Date	C.I.	Mean	95% HDP interval
H	33	41,668	30,938–52,398	52,035.5	24,925.1–84,152.0
K	29	20,115	10,447–29,783	30,881.0	10,334.3–54,075.3
T	17	22,549	13,404–31,704	65,968.4	25,771.5–112,860.0
U1	14	55,556	39,240–71,872	73,732.4	27,615.3–124,820.0
W6	24	20,139	13,736–26,542	50,399.8	17,810.0–86,407.6
X2	12	20,833	11,726–29,940	38,798.5	12,518.2–73,620.8
(b) T_{MRCA} Estimates for NRY Haplogroups in Svans (YBP)					
Haplogroup		***n***			**TMCRA (rho)**
G2a		79			12,952 (± 1,781)
I2a		4			11,807 (± 3,331)
J2a		6			33,336 (± 5,708)
N		1			–
R1a		10			8,334 (± 2,204)

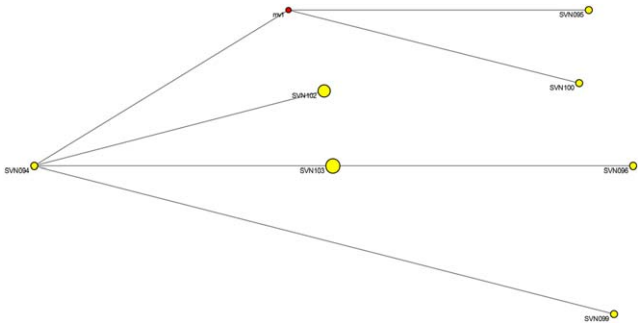

FIGURE 3 A median-joining network of mtDNA haplogroup X2 HVS1 sequences from Svans. The reticulations appearing in this network were likely due to recurrent mutations, such as T16093C, G16129A, and C16192T, which could not be resolved by adjusting the weighting scheme for HVS1 mutations

differed significantly from each other except for Georgians, Circassians, and Kabardians, each insignificantly different from the others, and also Iranians and Armenians, who were not significantly different from each other. Svans and South Ossetians appeared to be the most genetically divergent from each other, compared with the rest. This pattern of population relationships was largely the same when visualizing the F_{ST} data in a PCA plot, with the first two components explaining 78% of the variation (Supporting Information Figure S4). It is unclear whether this pattern reflects longstanding genetic differentiation of these geographically proximate ethnic groups or the effects of admixture with populations from outside the region.

To assess the genetic structure of Svans and regional populations, we conducted an AMOVA based on F_{ST} estimates for mtDNA haplotypes. Overall, the percentage of mtDNA variation among groups was 2.27% and that within groups 97.73%, respectively (Table 5). The high degree of genetic similarity of these populations was not entirely surprising, given the number of haplogroups that they shared (Supporting Information Table S1). When Svans and comparative populations were separated by their geographic location in the North or South Caucasus, a similar result was obtained, with only a small percentage of variance (0.48%) being attributable to geography. Likewise, when these populations were analyzed by linguistic affiliation, they showed no interpopulational differences, although the extent of variation within each language grouping increased to ~16% (Table 5).

We also investigated possible substructure at the village-level in Svaneti through AMOVA using F_{ST} values estimated from HVS1 haplotypes (Supporting Information Table S5). Most villages clustered relatively close to each other, with Kala and Ushguli being outliers, probably due to their small sample sizes (Figure 6). Given the nonsignificance of most F_{ST} values (Supporting Information Table S5), this analysis did not yield evidence of long-term geographic differentiation of local Svan populations at the maternal lineage level.

3.2 | Y-chromosome diversity in Svaneti

3.2.1 | Haplogroup diversity

Haplogroup G2a represented ~78% of Svan male lineages (Table 6; Supporting Information Table S6). The next most common haplogroup was R1a (10.6%), followed by J2alb (6.4%) and I2a (4.3%) and N (1.1%). The fact that these were the predominant Y-chromosome haplogroups in the Caucasus was unsurprising. The putative homelands of the two most prevalent of these lineages (G2a, J2a) lay in adjacent regions to the south and southwest (i.e., Anatolia and the greater Near East; Chiaroni et al., 2010; Rootsi et al., 2012) and that of the third (I2a) to the west (eastern Europe and Balkans; Underhill et al., 2007). In fact, G2a itself may, indeed, be an autochthonous lineage within the Caucasus (Cinnioğlu et al., 2004; Rootsi et al., 2012). However, although difficult to make inferences about population history from a single sample,

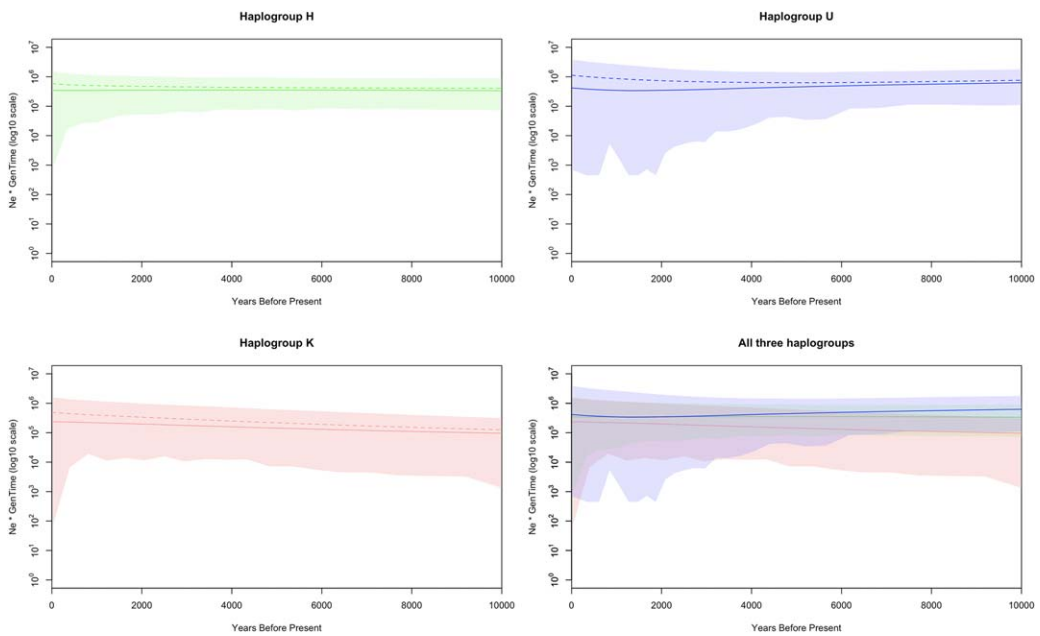

FIGURE 4 Extended Bayesian Skyline Plots for mtDNA haplogroups

TABLE 3 Descriptive statistics for mtDNA HVS1 haplotypes in Svaneti and comparative populations

Population	Source	*n*	# of haplotypes	Gene diversity	Nucleotide diversity	Mean pairwise differences	Tajima's D	Fu's *F*
Svans	This Study	184	103	0.982 ± 0.003	0.029 ± 0.015	5.645 ± 2.720	−1.772	−25.029
Abkhaz	Nasidze et al., 2004	27	19	0.969 ± 0.018	0.029 ± 0.016	5.801 ± 2.863	−1.196	−8.093
Anatolians	Gökçümen et al., 2011	163	104	0.990 ± 0.003	0.026 ± 0.014	5.218 ± 2.537	−2.115	−25.213
Armenians	Nasidze & Stoneking, 2001	42	36	0.981 ± 0.014	0.027 ± 0.015	5.271 ± 2.598	−2.144	−25.459
Circassians	Nasidze & Stoneking, 2001	43	39	0.985 ± 0.011	0.024 ± 0.013	4.761 ± 2.374	−2.041	−25.609
Georgians	This study & Nasidze et al., 2004	72	55	0.981 ± 0.009	0.029 ± 0.016	5.718 ± 2.771	−1.965	−25.322
Iranians	Terreros et al., 2011	109	94	0.996 ± 0.002	0.035 ± 0.019	6.989 ± 3.309	−2.117	−24.909
Kabardians	Nasidze & Stoneking, 2001	51	36	0.975 ± 0.011	0.024 ± 0.013	4.762 ± 2.367	−2.212	−25.615
N. Ossetians	Nasidze et al., 2004	126	67	0.967 ± 0.009	0.026 ± 0.014	5.071 ± 2.478	−1.891	−25.376
S. Ossetians	Kivisild et al., 1999	201	76	0.972 ± 0.005	0.031 ± 0.016	6.062 ± 2.899	−1.445	−24.871

it is likely that the N-bearing individual descends from a population living somewhere in eastern Eurasia, where it originated (Shi et al., 2013). The relative lack of East Eurasian haplogroups in Svans was also consistent with data from North Caucasus populations (Balanovsky et al., 2011; Yunusbayev et al., 2012).

We detected 75 distinct Y-STR haplotypes among 94 male Svan participants (Supporting Information Table S7). Of these, haplogroup G2a encompassed some 55 different haplotypes in 73 men. Two samples yielded incomplete STR profiles, although the partial Y-STR and Y-SNP data indicated that both belonged to haplogroup G2a.

In addition, we observed seven different Y-STR haplotypes in the seven male Georgian participants from our study (Supporting Information Table S7). These haplotypes belonged to haplogroups E1b1b, G2a, J2a1b, and R1b, with most falling into G2a. A generally similar pattern was seen in the Georgia DNA Project (2017) data. Haplogroups E1b1b and R1b did not appear in the Svans. In addition, although being somewhat similar, none of the Georgian G2a Y-STR haplotypes, including ones identified in this study and those from the Georgian DNA Project (2017), exactly matched those appearing in Svan men. Thus, Svans were genetically distinctive from Georgians on the basis of haplogroup composition and Y-STR haplotype diversity.

3.2.2 | Phylogenetic analysis of Y-STR haplotypes

We also subjected the Y-STR haplotypes from haplogroup G2a to network analysis. The resulting Svan G2a network lacked a central node or founder haplotype and exhibited a complex set of branches (Figure 7). Within these branches, a number of haplotypes occurred multiple times in Svans, with no one type predominating. In addition, the network showed no clear geographic structure, that is, no branches were found exclusively in one village of Svaneti. Overall, the topology of this network likely reflects the diversification of an old lineage within this region.

TABLE 4 F_{ST} estimates based on mtDNA HVS1 sequences for Svans and comparative populations

Pop.	Svan	Abk	Anat	Arm	Circ	Geo	Iran	Kab	NOss	SOss
Svan	0	+	+	+	+	+	+	+	+	+
Abk	0.024	0	+	+	+	+	+	+	+	+
Anat	0.016	0.025	0	+	+	+	+	+	+	+
Arm	0.029	0.041	0.017	0	+	+	–	+	+	+
Circ	0.021	0.034	0.021	0.012	0	–	+	–	+	+
Geo	0.020	0.023	0.01	0.011	0.004	0	+	–	+	+
Iran	0.025	0.021	0.008	0.006	0.018	0.011	0	+	+	+
Kab	0.013	0.016	0.012	0.011	0.001	0.006	0.014	0	+	+
NOss	0.025	0.024	0.014	0.029	0.018	0.018	0.019	0.007	0	+
SOss	0.047	0.044	0.018	0.022	0.041	0.029	0.014	0.034	0.037	0

Note: The F_{st} estimates are presented in the lower triangle of the table, while the significance values are presented in the upper triangle. The significance values are shown as pluses (+) and minuses (–), with (+) indicating a significant difference at $p = .050$.
Abbreviations: Svan = Svan; Abk = Abkhaz; NOss = North Ossetian; SOss = South Ossetian; Kab = Kabardian; Iran = Iranian; Anat = Anatolian; Arm = Armenian; Circ = Circassian; Geo = Georgian.

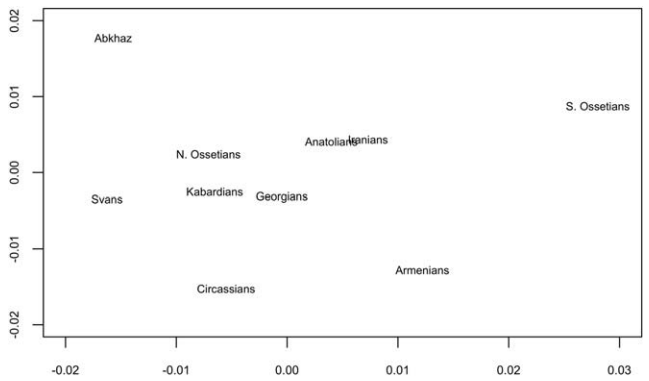

FIGURE 5 An MDS plot of F_{ST} values estimated using mtDNA HVS1 sequences from Svans and neighboring Caucasus populations (normalized raw stress = 0.03190)

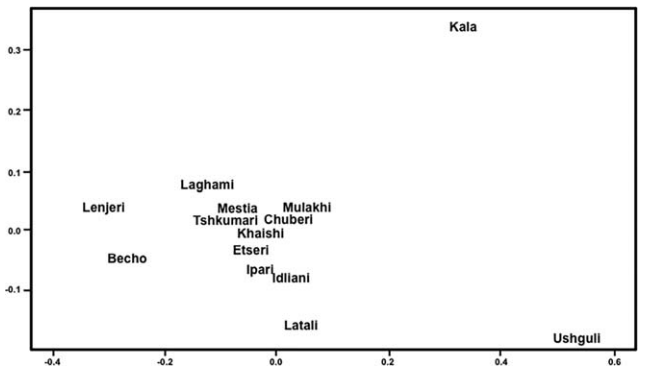

FIGURE 6 An MDS plot based on F_{ST} estimates generated from mtDNA HVS1 sequences from 14 Svan village districts

We used both the pedigree-based and evolutionary mutation rates to estimate T_{MCRA} for the NRY haplogroups present in Svans (Table 2). Given the time depth for both G2a in the Caucasus (Rootsi et al., 2012) and Svan as a language [emerging in the second millennium BCE (Deeters, 1930; Klimov, 1994) or possibly the third millennium BCE (Nichols, 1998)], the pedigree-based mutation rate was viewed as being too fast for these calculations. For this reason, we made coalescence estimates using the evolutionary rate. Since the nodes of this lineage most central to the network were represented by more than one sample (SV015 and SV028; Figure 7), we ran the coalescence analysis for haplogroup G2a using each one as the founder haplotype. The average T_{MCRA} for the G2a networks using these two haplotypes was ~12,600 YBP (Table 2). These results suggested that G2a emerged during the late Epipaleolithic or Neolithic period, and possibly in the South Caucasus, given its distribution in West Asia.

We also produced networks and estimated T_{MRCA}s for haplogroups I2a and R1a in Svans, despite their having a limited number of samples. The T_{MRCA} for I2a was 11,807 YBP. Although published T_{MRCA} estimates for I2a are difficult to find, Underhill et al. (2007) calculated the divergence of subclades I1 and I2 at 28,400 YBP, respectively. Lazaridis et al. (2014) also determined that I2a1b was present among Mesolithic northern Europeans, and therefore was not exclusively, a consequence of Neolithic demic diffusion. However, no study to date has focused exclusively on the phylogeographic history of haplogroup I and its subclades.

Our estimated T_{MRCA} for R1a in Svans was 8,334 YBP. This date was somewhat earlier than that produced by Underhill et al. (2014), whose coalescence time estimate for R1a-M417 was ~5800 YBP. Because all Svan R1a haplotypes were different from each other, it was impossible with so few individuals to make an accurate inference about the geographic sources for these haplotypes without additional Y-SNP data.

We also used Bayesian MCMC analysis to estimate the T_{MCRA} of haplogroup G2a in Svans, and to generate an eBSP (Supporting Information Figure S5). Despite the large number of Y-STR haplotypes in the Svans, these analyses failed to meet the minimum requirement for statistical power (ESS < 200), even after running for twice as long as the mtDNA counterparts (2×10^8 generations). This result likely reflects insufficient information in the Y-STR data to parameterize complex models such as an eBSP, unlike for mtDNA sequence data.

While generally informative, such T_{MRCA} estimates should be viewed as approximations of the ages of the paternal lineages being analyzed. Recent studies have indicated that STR-based age calculations tend to overestimate coalescence dates for Y-chromosome haplogroups compared to those estimated from complete NRY sequences (Batini et al., 2015; Hallast et al., 2015; Karmin et al., 2015). Thus, our results are suggestive of the dates at which these paternal lineages

TABLE 5 AMOVA of pairwise F_{ST} values based on mtDNA HVS1 haplotype data

Grouping	# Groups	Among Groups	Within Groups	Within Populations
All	1	–	2.27**	97.73**
Geography	2	0.48*	2.06**	97.46**
Language	4	−0.07*	15.91**	84.16**

Note: The populations analyzed in this AMOVA included Svans, Abkhaz, Anatolians, Armenians, Circassians, Georgians, Iranians, Kabardians, North Ossetians, and South Ossetians.
For the Geography analysis, populations assigned to "North Caucasus" included Abkhaz, Circassians, Kabardians, and North Ossetians, while those assigned to "South Caucasus" included Svans, Anatolians, Armenians, Georgians, Iranians, and South Ossetians.
For the Language analysis, "North Caucasus" = Abkhaz, Circassians, and Kabardians; "Indo-European" = Armenian, Iranian, North Ossetian and South Ossetian; "Kartvelian" = Georgians and Svans; "Turkic" = Anatolian.
*Nonsignificant at $p = .05$.
**Significant at $p = .05$.

TABLE 6 Y-chromosome haplogroup frequencies in Svans

Haplogroup	N	Frequency	Likely geographic source
G2a	79	79.0%	Caucasus
J2a1	6	6.0%	Middle East
I2	4	4.0%	Balkans/Europe
N	1	1.0%	East Eurasia
R1a	10	10.0%	West Eurasia

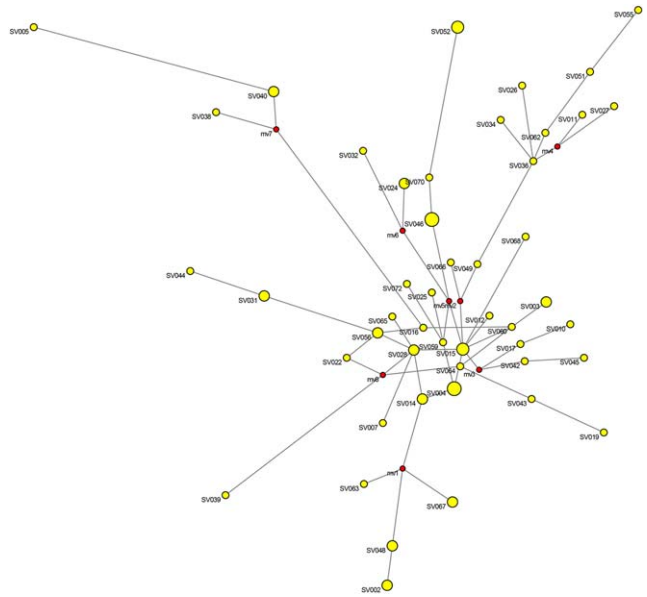

FIGURE 7 A Median-Joining network of haplogroup G2a Y-STR haplotypes in Svans

emerged, and not necessarily the times that they first appeared in the South Caucasus.

3.2.3 | Statistical analysis of Y-chromosome data

We carried out statistical analyses using Y-STR haplotype data (Supporting Information Table S7). Paternal haplotypic diversity in the Caucasus was fairly high overall (Table 7). While the Armenian and Georgian diversity values reflected the fact that their sample sizes matched the number of haplotypes present in them (see Methods), those of the Svans were somewhat lower compared to neighboring populations. In addition, Anatolians had a higher gene diversity value than other populations, a finding pointing to this region as a potentially important source of paternal lineages in the Caucasus.

To investigate the genetic relationship between Svans and comparative populations from the Caucasus and Anatolia, we generated an MDS plot based on R_{ST} estimates generated from their Y-STR haplotypes (Table 8). In this plot, Svans and the Ossetians stood out from the other populations, while Abkhaz, Georgians, and Circassians clustered closely with one another, and Armenians and Anatolians appeared at some distance from the rest (Figure 8). In terms of *p*-values, the genetic differences between all populations were statistically significant. This pattern of population relationships was largely the same when visualizing the R_{ST} data in a PCA plot, with the first two components explaining 80% of the variation (Supporting Information Figure S6). Given the geographic proximity and shared history of these groups (more certain in the case of Armenians and Anatolians; Payaslian, 2007), the statistical differences between Anatolian and Caucasus populations were intriguing.

We also analyzed paternal genetic diversity at the village level in Svaneti using R_{ST} estimates based on Y-STR haplotypes (Table 9). Due to the lack of significance for most pairwise comparisons, as reflected in the proximity of villages in the MDS plot (Figure 9), there was no clear differentiation of the different villages. This result suggested that, despite the diversity seen in them, Svan villages have not been isolated long enough from one another to accumulate distinct patterns of genetic diversity within them.

To further assess, the paternal genetic structure of Svans and regional populations, we conducted an AMOVA based on R_{ST} estimates for Y-STR haplotypes from Svan and comparative populations. Overall, the results revealed significant population-level differentiation, with 25% of the variance in genetic diversity being attributable to interpopulational differences (Table 10). When study populations were partitioned into their North and South Caucasus locations, the AMOVA revealed nearly significant differentiation between the groups, with 6% of the variance being attributable to geography. This finding was generally consistent with the clustering of Caucasus, Anatolian, Iranian, and Levantine populations in the MDS plot shown in Figure 8, and the Mantel test results, which indicated a nonsignificant correlation between genetic and geographic distances (data not shown). By contrast, when these populations were sorted by language, the AMOVA revealed no significant differentiation of the groups.

TABLE 7 Descriptive statistics for Y-STR haplotypes in Svaneti and neighboring regions

Population	Source	*n*	# of haplotypes	Gene diversity	Average gene diversity over loci
Svan	This study	88	72	0.995 ± 0.003	0.511 ± 0.264
Abkhaz	Balanovsky et al., 2011	52	46	0.988 ± 0.010	0.672 ± 0.344
Anatolians	Gökçümen et al., 2011	171	126	0.994 ± 0.002	0.664 ± 0.335
Armenians	Herrera et al., 2012	260	260[a]	1.000 ± 0.000	0.627 ± 0.317
Circassians	Balanovsky et al., 2011	132	105	0.996 ± 0.002	0.650 ± 0.329
Georgians	This study & Georgia History Project	166	166	1.000 ± 0.096	0.674 ± 0.341
N. Ossetians Digor	Balanovsky et al., 2011	126	90	0.988 ± 0.004	0.638 ± 0.324
N. Ossetians Iron	Balanovsky et al., 2011	229	151	0.988 ± 0.003	0.502 ± 0.258

[a]Herrera et al. (2012) did not indicate sample sizes for the observed Y-STR haplotypes in Armenians, thus forcing us to assume that all of them appeared only once in this population.

TABLE 8 R_{ST} values estimated from Y-STR haplotype data for Svans and comparative populations

	Svans	Georgians	Abkhazians	Circassians	Armenians	N. Ossetians Digor	N. Ossetians Iron	Anatolians
Svans	0.0000	+	+	+	+	+	+	+
Georgians	0.1823	0.0000	+	+	+	+	+	+
Abkhazians	0.1144	0.0543	0.0000	+	+	+	+	+
Circassians	0.1948	0.0504	0.0209	0.0000	+	+	+	+
Armenians	0.3335	0.1144	0.1444	0.1220	0.0000	+	+	+
N. Ossetians Digor	0.0580	0.1405	0.0775	0.1512	0.2149	0.0000	+	+
N. Ossetians Iron	0.1022	0.2870	0.2448	0.3039	0.3958	0.0776	0.0000	+
Anatolians	0.2241	0.0770	0.0691	0.0758	0.0635	0.1405	0.2959	0.0000

Note: The R_{ST} estimates are presented in the lower triangle of the table, while the significance values are presented in the upper triangle. The significance values are shown as pluses (+) and minuses (−), with (+) indicating significant differences at $p = .050$.

4 | DISCUSSION

Our study of genetic variation in Svans revealed a complex population structure in the highland South Caucasus. Broadly speaking, both the Svan paternal and maternal gene pools consist mainly of West Eurasian lineages, along with some East Eurasian types. While this pattern is generally similar to genetic diversity seen in other populations of the Caucasus, Iran, and eastern Anatolia (Nasidze et al., 2004; Khusnutdinova et al., 2012), Svans are distinctive in having unusually high frequencies of mtDNA haplogroups W6 and X2 and the presence of U1-U7 haplotypes, as well as an extremely high frequency of Y-chromosome haplogroup G2a. This distinct pattern is consistent across all Svan communities, which show no genetic structure within the region.

While precise geographic origins of several mtDNA haplogroups (e.g., H, K) found in Svans remain unclear, the majority are of West Eurasian or North African origin. Many of these lineages expanded into Europe especially during the Neolithic, although some show signs of much earlier, pre-LGM diversification in the Caucasus (Metspalu et al., 1999). Since these haplogroups are all thought to have originated somewhere in the Near East or West Asia, and all have coalescence dates well before the LGM (Achilli et al., 2004; Metspalu et al., 1999; Richards et al., 1999; Roostalu et al., 2007; Soares et al., 2009), their time depth in the Caucasus may be significant. Bayesian analysis of the major mtDNA haplogroups present among Svans also yielded coalescence ages extending back into the Upper Paleolithic, thus suggesting they might have been present in the Caucasus at that time. This finding also supports the notion that Svans and other populations of the highland Caucasus descend from early human settlement events in Eurasia during the Upper or Epipaleolithic eras of this region. However, these coalescence dates have wide confidence intervals and may be influenced by sample sizes for the haplogroups being analyzed.

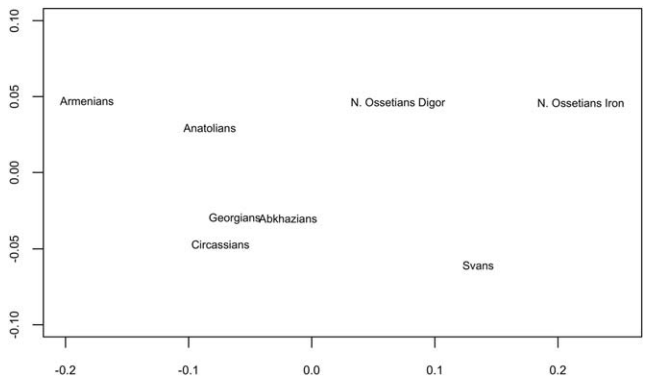

FIGURE 8 An MDS plot based on R_{ST} values generated from Y-STR haplotypes in Svans and their Caucasus neighbors

Furthermore, the AMOVA using HVS1 sequences shows no clear distinction in mtDNA diversity between Svan and neighboring populations. The similar frequencies of mtDNA haplogroups (H, HV, J, K, T, W) across ethno-linguistic boundaries would seem to belie an ancestral Caucasus population having a maternal gene pool consisting of these haplogroups, as well as different lineages of haplogroup U. Geographically speaking, this haplogroup has a continuous presence across Eurasia, from Iberia to Baikal, and its various branches (U1-U7) arose ~20–35,000 years ago in this broad region (e.g., Metspalu et al., 2004; Richard et al., 2000). For this reason, the presence of all of the U haplogroups in Svaneti is worth noting. Our data also suggest that W6 might have originated in the South Caucasus, given its significant frequency there. Together, these findings further suggest that Iran, Anatolia, and North and South Caucasus share a common set of maternal lineages, and that geography and language have not profoundly influenced mtDNA diversity in these populations.

While the phylogeographic pattern underlying Svan mtDNA diversity largely reflects that of the greater Middle East (west of the Indus Valley), including Anatolia and the Caucasus, the phylogeographic origin of male diversity seems to be more localized (Rootsi et al., 2012). Y-chromosome haplogroup diversity among Svan males is relatively low compared to that seen in other West Asian populations (e.g., Badro et al., 2013; Grugni et al., 2012), although haplotypic diversity within them is quite high. Four of the Y-chromosome haplogroups (G2a, R1a, J2a1b, I2a) found in Svans are also associated with male-mediated migrations related to Neolithic farming (Balaresque et al., 2010; Chiaroni et al., 2008; Chikhi et al., 1998, 2002; Haak et al., 2010; Lacan

TABLE 9 Pairwise R_{ST} values based on Y-STR haplotype data from Svan villages

	Becho	Chuberi	Etseri	Idliani	Ipari	Latali	Mestia	Mulakhi	Tshkumari	Ushguli
Becho	0	–	+	–	–	+	–	–	–	–
Chuberi	0.077	0	–	–	–	–	–	+	+	+
Etseri	0.014	0.025	0	–	–	–	–	–	–	–
Idliani	0.053	−0.005	0.003	0	–	–	–	+	–	+
Ipari	0.142	−0.042	0.073	0.26	0	–	–	+	–	+
Latali	−0.056	0.037	−0.019	0.036	0.09	0	–	–	–	–
Mestia	0.047	−0.017	−0.014	−0.011	0.048	0.003	0	–	–	–
Mulakhi	0.116	0.098	0.033	0.274	0.237	0.06	0.033	0	–	–
Tshkumari	−0.004	0.081	−0.027	0.121	0.16	−0.056	−0.006	0.027	0	–
Ushguli	0.059	0.238	0.061	0.304	0.384	0.023	0.148	0.105	0.009	0

Note: The R_{st} estimates are presented in the lower triangle of the table, while the significance values are presented in the upper triangle. The significance values are shown as pluses (+) and minuses (–) in the upper section of the table, with (+) indicating significant differences at $p = .050$.

et al., 2011). However, it must be noted that a recent Late Upper Paleolithic individual from Kotias (Georgia) belongs to J2a (Jones et al., 2015), suggesting that not all of these lineages were fully associated with the spread of Neolithic farming.

High haplotypic diversity in a region with otherwise low haplogroup diversity may point to one of several scenarios leading to this pattern. It could indicate the presence of a few old, well-rooted male lineages whose diverse lineages have evolved *in situ*. On the other, it may point to a region geographically prone to accumulating settlers from a diverse source area (or areas) with similar Y-STR haplotypes. Given its history, the South Caucasus would partly fit both scenarios. With the epicenter of haplogroup G2a being the eastern Black Sea area (Rootsi et al., 2012), this region may have served as a long-standing source population for all three highland populations—Svans, Abkhaz, and Ossetians—despite the observed genetic distances between them (Figure 8). Such distances may reflect their haplogroup composition relative to other study populations and most likely the range of haplotypes present in them. The lower haplogroup diversity among Svans could further be attributable to their geographic remoteness, and possibly to reduced male gene flow, rather than to a recent founder effect, which the data do not reveal.

These results were also broadly consistent with the analysis of linguistic affiliation, which showed no differences between study populations on the basis on language family (Table 5). This result suggests that geography is a stronger influence on genetic variation in this culturally complex region. Our findings contrast with those of Balanovsky et al. (2011), who observed geographic "zones" of language-gene coevolution in the North Caucasus. In the South Caucasus, there do not appear to be any geographic or haplogroup-dominated zones, most likely due to the fact that regional populations are largely unconstrained by the terrain that isolated speakers of Chechen, Circassian, Avar, and other languages in the North Caucasus.

The T_{MRCA} estimate (~12,600 YBP) for NRY haplogroup G2a in Svans suggests an origin for this lineage in the Epipaleolithic. The overall distribution of G2a in the Caucasus, and among Svans, may be

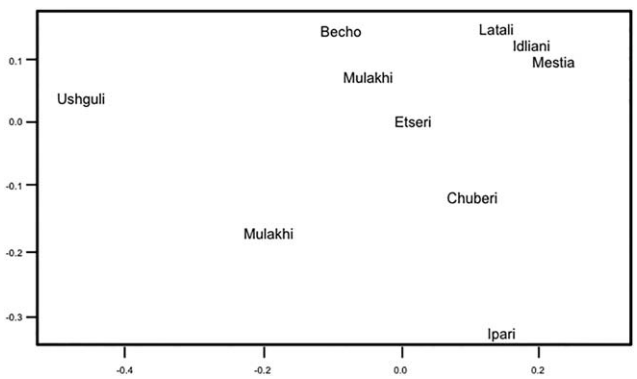

FIGURE 9 An MDS plot based on R_{ST} estimates generated from Y-STR haplotypes in 14 Svan village districts

TABLE 10 AMOVA of R_{ST} values estimated with Y-STR haplotype data

Grouping	# Groups	Among Groups	Within Groups	Within Populations
All	1	–	18.95**	81.05**
Geography	2	6.06*	14.96**	78.89**
Language	4	−7.71*	25.27**	82.45**

Note: The populations analyzed in this AMOVA included Svans, Abkhaz, Anatolians, Armenians, Circassians, Georgians, North Ossetians (Digor), and North Ossetians (Iron).
For the Geography analysis, populations assigned to "North Caucasus" included Abkhaz, Circassians, and North Ossetians, while those assigned to "South Caucasus" included Svans, Anatolians, Armenians, Georgians, and South Ossetians.
For the Language analysis, "North Caucasus" = Abkhaz and Circassians; "Indo-European" = Armenian, North Ossetian and South Ossetian; "Kartvelian" = Georgians and Svans; and "Turkic" = Anatolian.
*Nonsignificant at $p = .05$.
**Significant at $p = .05$.

interpreted as a genetic signal of the expansion of early Neolithic settlers from the Black Sea coast and nearby lowlands following the retreat of the glaciers in the Caucasus (Rootsi et al., 2012). Although G2a haplotyps may have originated in lowland Georgia, eastern Anatolia, or beyond, their close genetic relationship to their descendants' current highland neighbors in the Caucasus is suggestive of a nearby geographical locus. Given the high Y-STR haplotype diversity of G2a, it is reasonable to posit the long-term presence and differentiation of this haplogroup among the ancestors of the Svans, perhaps beginning in the late Epipaleolithic or early Neolithic of the South Caucasus. Alternatively, this diversity could possibly reflect recurrent gene flow, that is, the regular introgression of males bearing G2a Y-chromosomes from Georgia or the South Caucasus.

We further explored the genetic relationships between Svans and their western and eastern neighbors, the Abkhaz and South Ossetians, respectively. Our analysis of Y-chromosome data indicate a common ancestry for, or an exchange of, male lineages between Svans, Ossetians, and Abkhaz, in spite of the fact that these populations speak languages from three distinct families. Although premature to suggest that a putative common ancestral population emerged in the highlands, these haplotype distributions were likely established well before the formation of the current ethno-linguistic groups in the Caucasus. Relatively close paternal lineage affinities between Svans and neighboring populations (i.e., Ossetians) further point to a common source population for, and/or subsequent unrestricted gene flow between them despite the observed differences at the language family level.

Our examination of Svan mtDNA and Y-chromosome diversity in relation to village residence in Svaneti revealed no phylogeographic patterning within Svaneti. These results imply that the observed mtDNA diversity accumulated in the highland Caucasus region before the formation of the current ethnolinguistic communities there. This pattern may also have been influenced by the post-abolition (1871) demographic reorganization of Svaneti in recent years (Avaliani, 1913; Gasviani, 1980). Neither haplotype nor haplogroup frequencies at the village level were significant enough to suggest any geographic or clan-based patterning in maternal lineages. We also surmise that the general Svan mtDNA pool was likely formed over many millennia, an observation consistent with data from the greater West Asian region.

The high maternal genetic diversity in Svans may further reflect social and cultural practices in the region. For one thing, it may point to the possibility of greater exchange of mtDNA lineages between highland Caucasus regions due to patrilocality, bride theft, and other cultural processes (Grant, 2005; Marchani et al., 2008; Nasidze et al., 2004). The reluctance of Svans to marry someone sharing a surname going back 10 generations (Tuite, 1994) may also have led to the necessity of finding marital partners outside the region, although these social practices cannot solely account for the incredible mtDNA diversity in this small population.

To conclude, while being the first anthropological genetic study undertaken in Svaneti, this analysis also represents the beginning of a comprehensive examination of genetic variation in Georgia that will situate its history more firmly within the broader context of the Caucasus and Near East. Future studies focusing on genetic diversity in other parts of Georgia and eastern Anatolia will also help to clarify the geographic movement of some of the mtDNA and Y-chromosome haplogroups seen in Svans. This research will further result in a much better comparative understanding of the historical ethnology of the Caucasus, as well as its geographic importance to early human expansions.

ACKNOWLEDGMENTS

We wish to express our gratitude to the people of Svaneti for their hospitality and participation in this study. We would also like to acknowledge the organizations that have provided administrative and logistical support for this project. These include the University of Pennsylvania Museum of Archeology and Anthropology, the University of Pennsylvania Department of Anthropology, Tbilisi State Medical University, Ivane Javakhishvili Institute of History and Ethnology, and the Georgian Academy of Sciences. In addition, we would also like to thank Ilya Japaridze, mayor of Mestia, for his facilitation of their work in remote Svan communities; Adilar and Zaur Chartolani for assistance with logistics in Svaneti; Dr. Giorgi Kavtaradze for valuable suggestions for and guidance with this project; Rezo and Izolda Chitanava and Dr. and Mrs. Gerald Hurst for their material and moral support of the project; Dr. Eugene Potapov of Bryn Athyn College and Irakli Akhvlediani for assistance with data analysis; and Drs. Laurent Ristvet, Brian Spooner, and Clark Erickson of the University of Pennsylvania and Dr. Kevin Tuite from the Université de Montréal for their constructive reading of the manuscript. We received financial support for this project from the National Science Foundation (BCS-1249281) the University of Pennsylvania Museum of Archeology and Anthropology, the University of Pennsylvania Department of Anthropology, the American Philosophical Society, and Penn Faculty Research Funds. Finally, we acknowledge the infrastructural support provided to the Laboratory of Molecular Anthropology at Penn by the National Geographic Society.

ORCID

Theodore G. Schurr http://orcid.org/0000-0001-9323-9237

REFERENCES

Achilli, A., Rengo, C., Magri, C., Battaglia, V., Olivieri, A., Scozzari, R., … Torroni, A. (2004). The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. *American Journal of Human Genetics*, *75*, 910–918.

Alfonso-Sánchez, M. A., Cardoso, S., Martínez-Bouzas, C., Peña, J. A., Herrera, R. J., Castro, A., …, de Pancorbo, M. M. (2008). Mitochondrial DNA haplogroup diversity in Basques: A reassessment based on HVI and HVII polymorphisms. *American Journal of Human Biology*, *20*, 154–164.

Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H., Coulson, A. R., Drouin, J., … Young, I. G. (1981). Sequence and organization of the human mitochondrial genome. *Nature*, *290*, 457–465.

Andrews, R. M., Kubacka, I., Chinnery, P. F., Lightowlers, R. N., Turnbull, D. M., & Howell, N. (1999). Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nature Genetics*, *23*, 147.

Avaliani, S. L. (1913). *Krestyanskiy vopros v zakavkazye*. Vol. *II*. Odessa: Tipografiya "Tekhnik" Yekaterininskaya (In Russian).

Badro, D. A., Douaihy, B., Haber, M., Youhanna, S. C., Salloum, A., Ghassibe-Sabbagh, M., … Zalloua, P. A. The Genographic Consortium. (2013). Y-chromosome and mtDNA genetics reveal significant contrasts in affinities of modern Middle Eastern populations with European and African populations. *PLoS One*, *8*, e54616.

Balanovsky, O., Dibirova, K., Dybo, A., Mudrak, O., Frolova, S., Pocheshkhova, E., … Balanovska, E. (2011). Parallel evolution of genes and languages in the Caucasus region. *Molecular Biology and Evolution*, *28*, 2905–2920.

Balaresque, P., Bowden, G. R., Adams, S. M., Leung, H.-Y., King, T. E., Rosser, Z. H., … Jobling, M. A. (2010). A predominantly Neolithic origin for European paternal lineages. *PLoS Biology*, *8*, e1000285.

Bandelt, H.-J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, *16*, 37–48.

Bandelt, H.-J., Quintana-Murci, L., Salas, A., & Macaulay, V. (2002). The fingerprint of phantom mutations in mitochondrial DNA data. *American Journal of Human Genetics*, *71*, 1150–1160.

Batini, C., Hallast, P., Zadik, D., Delser, P. M., Benazzo, A., Ghirotto, S., … Jobling, M. A. (2015). Large-scale recent expansion of European patrilineages shown by population resequencing. *Nature Communications*, *19*, 7152.

Battaglia, V., Fornarion, S., Al-Zahery, N., Oliveri, A., Pala, M., Myres, N. M., … Semino, O. (2009). Y-chromosomal evidence of the cultural diffusion of agriculture in southeast Europe. *European Journal of Human Genetics*, *17*, 820–830.

Behar, D. M., van Oven, M., Rosset, S., Metspalu, M., Loogväli, E.-L., Silva, N. M., … Villems, R. (2012). A "Copernican" Reassessment of the Human Mitochondrial DNA Tree from its Root. *American Journal of Human Genetics*, *90*, 675–684.

Brandstätter, A., Zimmermann, B., Wagner, J., Göbel, T., Röck, A. W., Salas, A., … Parson, W. (2008). Timing and deciphering mitochondrial DNA macro-haplogroup R0. *BMC Evolutionary Biology*, *8*, 191.

Chandler, J. F. (2006). Estimating per-locus mutation rates. *Journal of Genetic Genealogy*, *2*, 27–33.

Chartolani, A. (1967). *Svaneti*. Tbilisi: Metsniereba (in Georgian).

Chiaroni, J., King, R. J., Myres, N. M., Henn, B. M., Mitchell, M. J., Boetsch, G., … Underhill, P. A. (2010). The emergence of Y-chromosome haplogroup J1e among Arabic-speaking populations. *European Journal of Human Genetics*, *18*, 348–353.

Chiaroni, J., King, R. J., & Underhill, P. A. (2008). Correlation of annual precipitation with human Y-chromosome diversity and the emergence of Neolithic agricultural and pastoral economies in the Fertile Crescent. *Antiquity*, *82*, 281–289.

Chikhi, L., Destro-Bisol, G., Bertorelle, G., Pascali, V., & Barbujani, G. (1998). Clines of nuclear DNA markers suggest a largely Neolithic ancestry of the European gene pool. *Proceedings of the National Academy of Sciences of the United States of America*, *95*, 9053–9058.

Chikhi, L., Nichols, R. A., Barbujani, G., & Beaumont, M. A. (2002). Y genetic data support the Neolithic demic diffusion model. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 11008–11013.

Chirikba, V. A. (2008). The problem of the Caucasian *Sprachbund*. In P. Muysken (Ed.), *From linguistic areas to areal linguistics* (pp. 25–93). Amsterdam: John Benjamins Publishing Co..

Chirikba, V. A. (2015). Between Christianity and Islam: Heathen Heritage in the Caucasus. In: U. Bläsing, V. Arakelova, M. Weinreich M (Eds.), *Studies on Iran and the Caucasus: In honour of Garnik Asatrian* (pp. 145–191). Leiden: Brill.

Cinnioğlu, C., King, R., Kivisild, T., Kalfoğlu, E., Atasoy, S., Cavalleri, G. L., … Underhill, P. A. (2004). Excavating Y-chromosome haplotype strata in Anatolia. *Human Genetics*, *14*, 127–148.

Comas, D., Calafell, F., Mateu, E., Pérez-Lezaun, A., Bosch, E., Martínez-Arias, R., … Bertranpetit, J. (1998). Trading genes along the Silk Road: mtDNA sequences and the origin of central Asian populations. *American Journal of Human Genetics*, *63*, 1824–1838.

Cruciani, F., Trombetta, B., Massaia, A., Destro-Bisol, G., Sellitto, D., & Scozzari, R. (2011). A revised root for the human Y Chromosomal phylogenetic tree: The origin of patrilineal diversity in Africa. *American Journal of Human Genetics*, *88*, 814–818.

Deeters, G. (1930). Das Karthwelische Verbum. Vergleichende Darstellung des Verbalbaus der südkaukasischen Sprachen. Sächsische Forschungsinstitute Leipzig. Forschungsinstitut für Indogermanistik, Sprachwissenschaftliche Abteilung. Band 1. Leipzig: Merkert & Petters.

Dulik, M. C., Osipova, L. P., & Schurr, T. G. (2011). Y-chromosome variation in Altaian Kazakhs reveals a common paternal gene pool for Kazakhs and the influence of Mongolian expansions. *PLoS One*, *6*, e17548.

Dulik, M. C., Owings, A. C., Gaieski, J. B., Vilar, M. G., Andre, A., Lennie, C., … Schurr, T. G. &The Genographic Consortium. (2012). Y-chromosome analysis reveals genetic divergence and new founding native lineages in Athapaskan- and Eskimoan-speaking populations. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 8471–8476.

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin Suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*, 564–567.

Fernandes, V., Alshamali, F., Alves, M., Costa, M. D., Pereira, J. B., Silva, N. M., … Pereira, L. (2012). The Arabian cradle: mitochondrial relicts of the first steps along the southern route out of Africa. *American Journal of Human Genetics*, *90*, 347–355.

Francalacci, P., Morelli, L., Angius, A., Berutti, R., Reinier, F., Atzeni, R., … Cucca, F. (2013). Low-pass DNA sequencing of 1200 Sardinians reconstructs European Y-chromosome phylogeny. *Science*, *341*, 565–569.

Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, *147*, 915–925.

Gaieski, J. B., Owings, A. C., Vilar, M. G., Dulik, M. C., Gaieski, D. F., Gittelman, R. M., … Schurr, T. G. (2011). Genetic ancestry and indigenous heritage in a Native American descendant community in Bermuda. *American Journal of Physical Anthropology*, *146*, 392–405.

Gasviani, G. A. (1980). *Sotsialno-ekonomicheskaya struktura Svaneti v XI-XVIII vv*. Tbilisi: Metsniereba (in Georgian).

Georgia DNA Project. (2017). FamilyTreeDNA.com Available at: https://www.familytreedna.com/public/georgia/default.aspx?section=yresults.

Gökçümen, Ö., Gultekin, T., Alakoc, Y. D., Tug, A., Gulec, E., & Schurr, T. G. (2011). Biological Ancestries, Kinship Connections and Projected Identities in Four Central Anatolian Settlements. *American Anthropology*, *113*, 116–131.

González, A. M., Larruga, J. M., Abu-Amero, K. K., Shi, Y., Pestano, J., & Cabrera, V. M. (2007). Mitochondrial lineage M1 traces an early human backflow to Africa. *BMC Genomics*, *8*, 223.

Grant, B. (2005). The good Russian prisoner: Naturalizing violence in the Caucasus Mountains. *Cultural Anthropology*, *20*, 39–67.

Grugni, V., Battaglia, V., Hooshiar Kashani, B., Parolo, S., Al-Zahery, N., Achilli, A., … Semino, O. (2012). Ancient migratory events in the Middle East: new clues from the Y-chromosome variation of modern Iranians. *PLoS One*, *7*, e41252.

Haak, W., Balanovsky, O., Sanchez, J. J., Koshel, S., Zaporozhchenko, V., Adler, C. J., … Cooper, A. & The Genographic Consortium. (2010). Ancient DNA from European Early Neolithic farmers reveals their Near Eastern affinities. *PloS Biology*, *8*, e1000536.

Hallast, P., Batini, C., Zadik, D., Maisano Delser, P., Wetton, J. H., Arroyo-Pardo, E., … Jobling, M. A. (2015). The Y-chromosome tree bursts into leaf: 13,000 high-confidence SNPs covering the majority of known clades. *Molecular Biology and Evolution*, *32*, 661–673.

Herrera, K. J., Lowery, R. K., Hadden, L., Calderon, S., Chiou, C., Yepiskoposyan, L., … Herrera, R. J. (2012). Neolithic patrilineal signals indicate that the Armenian plateau was repopulated by agriculturalists. *European Journal of Human Genetics*, *20*, 313–320.

Hofmanová, Z., Kreutzer, S., Hellenthal, G., Sell, C., Diekmann, Y., Díez del, M. D., … Burger, J. (2015). Early farmers from across Europe directly descended from Neolithic Aegeans. Available at: http://biorxiv.org/content/early/2015/11/25/032763. Accessed June 1st, 2017.

Jones, E. R., Gonzalez-Fortes, G., Connell, S., Siska, V., Eriksson, A., Martiniano, R., … Bradley, D. G. (2015). Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nature Communications*, *6*, 8912.

Karafet, T. M., Mendez, F. L., Meilerman, M. B., Underhill, P. A., Zegura, S. L., & Hammer, M. F. (2008). New binary polymorphisms reshape and increase resolution of the human Y-chromosomal haplogroup tree. *Genome Research*, *18*, 830.

Karmin, M., Saag, L., Vicente, M., Wilson Sayres, M. A., Järve, M., Talas, U. G., … Kivisild, T. (2015). A recent bottleneck of Y chromosome diversity coincides with a global change in culture. *Genome Research*, *25*, 459–466.

Kavtaradze, G. L. (2000). Some problems of the interrelation of Caucasian and Anatolian Bronze Age cultures. *Quaderni Di Archeologia Università Di Messina*, *1*, 107–123.

Khusnutdinova, E. K., Litvinov, S. S., Kutuev, I. A., Yunusbayev, B. B., Khusainova, R. I., Ahmetova, V. L., … Villems, R. (2012). Gene pool of ethnic groups of the Caucasus: Results of integrated study of the Y chromosome and mitochondrial DNA and genome-wide data. *Russian Journal of Genetics*, *48*, 640–650.

Klimov, G. A. (1969). [W. Boeder, Trans.]. *Die kaukasische Sprachen*. Hamburg: Helmut Buske.

Klimov, G. A. (1994). [J. Gippert, Trans.]. *Einffuhrung in die kaukasische sprachwissenschaft*. Berlin: Helmut Buske.

Kloss-Brandstätter, A., Pacher, D., Schönherr, S., Weissensteiner, H., Binna, R., Specht, G., & Kronenberg, F. (2011). HaploGrep: A fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. *Human Mutation*, *32*, 25–32.

Lacan, M., Keyser, C., Ricaut, F.-X., Brucato, N., Tarrús, J., Bosch, A., … Ludes, B. (2011). Ancient DNA suggests the leading role played by men in the Neolithic dissemination. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 18255–18259.

Li, C., Li, H., Cui, Y., Xie, C., Cai, D., Li, W., … Zhou, H. (2010). Evidence that a West-East admixed population lived in the Tarim Basin as early as the early Bronze Age. *BMC Biology*, *8*, 15.

Marchani, E. E., Watkins, W. S., Bulayeva, K., & Jorde, L. B. (2008). Culture creates genetic structure in the Caucasus: Autosomal, mitochondrial, and Y-chromosomal variation in Daghestan. *BMC Genetics*, *9*.

Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S. A., … Reich, D. (2015). Genome-wide patterns of selection in 230 ancient Eurasians. *Nature*, *528*, 499–503.

Metspalu, E., Kivisild, T., Kaldma, K., Parik, J., Reidla, M., Tambets, K., & Villems, R. (1999). The Trans-Caucasus and the expansion of the Caucasoid-specific human mitochondrial DNA. In S. S.Papiha, R. Deka, & R. Chakraborty (Eds.), *Genomic diversity: Applications in human population genetics* (pp. 121–133). New York: Kluwer Academic/Plenum Publishers.

Metspalu, M., Kivisild, T., Metspalu, E., Parik, J., Hudjashov, G., Kaldma, K., … Villems, R. (2004). Most of the extant mtDNA boundaries in South and Southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans. *BMC Genetics*, *5*, 26.

Nasidze, I., Schädlich, H., & Stoneking, M. (2003). Haplotypes from the Caucasus, Turkey and Iran for nine Y-STR loci. *Forensic Science International*, *137*, 85–93.

Nasidze, I., Ling, E. Y., Quinque, D., Dupanloup, I., Cordaux, R., Rychkov, S., … Stoneking, M. (2004). Mitochondrial DNA and Y chromosome variation in the Caucasus. *Annals of Human Genetics*, *68*, 205–221.

Nichols, J. (1998). The origin and dispersal of languages: Linguistic evidence. In: N. G. Jablonski & L. C. Aiello (Eds.), *The origin and diversification of language*. Memoirs of the California Academy of Sciences No. 24. San Francisco: California Academy of Sciences.

Olivieri, A., Pala, M., Gandini, F., Hooshiar Kashani, B., Perego, U. A., Woodward, S. R., … Torroni, A. (2013). Mitogenomes from two uncommon haplogroups mark late glacial/postglacial expansions from the Near East and Neolithic dispersals within Europe. *PLoS One*, *8*, e70492.

Payaslian, S. (2007). *The History of Armenia: From The Origins To The Present*. New York: Palgrave Macmillan.

Posth, C., Renaud, G., Mittnik, A., Drucker, D. G., Rougier, H., Cupillard, C., … Krause, J. (2016). Pleistocene mitochondrial genomes suggest a single major dispersal of non-Africans and a late glacial population turnover in Europe. *Current Biology*, *26*, 827–833.

Quintana-Murci, L., Chaix, R., Wells, R. S., Behar, D. M., Sayar, H., Scozzari, R., … McElreavey, K. (2004). Where west meets east: The complex mtDNA landscape of the southeast and Central Asian corridor. *American Journal of Human Genetics*, *74*, 827–845.

Reidla, M., Kivisild, T., Metspalu, E., Kaldma, K., Tambets, K., Tolk, H. V., … Villems, R. (2003). Origin and diffusion of mtDNA haplogroup X. *American Journal of Human Genetics*, *73*, 1178–1190.

Richards, M. B., Macaulay, V. A., Bandelt, H. J., & Sykes, B. C. (1998). Phylogeography of mitochondrial DNA in western Europe. *Annals of Human Genetics*, *62*, 241–260.

Richards, M., Macaulay, V., Hickey, E., Vega, E., Sykes, B., Guida, V., … Bandelt, H. J. (2000). Tracing European founder lineages in the Near Eastern mtDNA pool. *American Journal of Human Genetics*, *67*, 1251–1276.

Richards, M., Macaulay, V., Torroni, A., & Bandelt, H. J. (2002). In search of geographical patterns in European mitochondrial DNA. *American Journal of Human Genetics*, *71*, 1168–1174.

Roostalu, U., Kutuev, I., Loogväli, E. L., Metspalu, E., Tambets, K., Reidla, M., … Villems, R. (2007). Origin and expansion of haplogroup H, the dominant human mitochondrial DNA lineage in West Eurasia: the Near Eastern and Caucasian perspective. *Molecular Biology And Evolution*, *24*, 436–448.

Rootsi, S., Myres, N. M., Lin, A. A., Järve, M., King, R. J., Kutuev, I., … Underhill, P. A. (2012). Distinguishing the co-ancestries of haplogroup G Y-chromosomes in the populations of Europe and the Caucasus. *European Journal of Human Genetics*, *20*, 1275–1282.

Sánchez-Velasco, P., & Leyva-Cobián, F. (2001). The HLA class I and class II allele frequencies studied at the DNA level in the Svanetian population (Upper Caucasus). *Tissue Antigens*, *58*, 223–233.

Schurr, T. G., Dulik, M. C., Owings, A. C., Zhadanov, S. I., Gaieski, J. B., Vilar, M. G., … Natkong, F. & The Genographic Consortium. (2012). Clan, language, and migration history has shaped genetic diversity in Haida and Tlingit populations from Southeast Alaska. *American Journal of Physical Anthropology*, *148*, 422–435.

Sengupta, S., Zhivotovsky, L. A., King, R., Mehdi, S. Q., Edmonds, C. A., Chow, C. E., … Underhill, P. A. (2006). Polarity and temporality of high-resolution Y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralists. *American Journal of Human Genetics*, *78*, 202–221.

Shi, H., Qi, X., Zhong, H., Peng, Y., Zhang, X., Ma, R. Z., & Su, B. (2013). Genetic evidence of an East Asian origin and Paleolithic northward migration of Y-chromosome haplogroup N. *PLoS One*, *8*, e66102.

Smouse, P. E., Long, J. C., & Sokal, R. R. (1986). Multiple regression and correlation extensions of the Mantel Test of matrix correspondence. *Systematic Zoology*, *35*, 627–632.

Soares, P., Ermini, L., Thomson, N., Mormina, M., Rito, T., Röhl, A., … Richards, M. B. (2009). Supplemental data correcting for purifying selection: An improved human mitochondrial molecular clock. *American Journal of Human Genetics*, *84*, 82–93.

Strabo. (1924). *Geography*. H. L. Jones (Ed.), Cambridge: Harvard University Press.

Tajima, F., & Nei, M. (1984). Estimation of evolutionary distance between nucleotide sequences. *Molecular Biology and Evolution*, *1*, 269–285.

Tarkhnishvili, D., Gavashelishvili, A., Murtskhvaladze, M., Gabelaia, M., & Tevzadze, G. (2014). Human paternal lineages, languages and environment in the Caucasus. *Human Biology*, *86*(2): 113–30.

Terreros, M. C., Rowold, D. J., Mirabal, S., & Herrera, R. J. (2011). Mitochondrial DNA and Y-chromosomal stratification in Iran: relationship between Iran and the Arabian Peninsula. *Journal of Human Genetics*, *56*, 235–246.

Torroni, A., Sukernik, R. I., Schurr, T. G., Starikorskaya, Y. B., Cabell, M. F., Crawford, M. H., … Wallace, D. C. (1993). mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. *American Journal of Human Genetics*, *53*, 591–608.

Tuite, K. (1994). The Svans. In: P. Friedrich, N. Diamond, & D. Levinson (Eds.). *The Encyclopedia of World Cultures* (Vol. *6*; pp. 343–347). Boston: GK Hall & Sons.

Tuite, K. (2000). Anti-marriage in ancient Georgian society. *Anthropological Linguistics*, *42*, 37–60.

Underhill, P. A., Myres, N. M., Rootsi, S., Chow, C. T., Lin, A. A., Otillar, R. P., … Woodward, S. R. (2007). New phylogenetic relationships for Y-chromosome haplogroup I: Reappraising its phylogeography and prehistory. In: P. Mellars, K. Boyle, O. Bar-Yosef, & C. Stringer C (Eds.), *Rethinking human evolution: New behavioural and biological perspectives on the origin and dispersal of modern humans* (pp. 33–42). Cambridge: McDonald Institute for Archaeological Research.

Underhill, P. A., Myres, N. M., Rootsi, S., Metspalu, M., Zhivotovsky, L. A., King, R. J., … Kivisild, T. (2010). Separating the post-glacial coancestry of European and Asian Y chromosomes within haplogroup R1a. *European Journal of Human Genetics*, *18*, 479–484.

van Oven, M., & Kayser, M. (2009). Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Human Mutation*, *30*, E386–E394.

van Oven, M., Van Geystelen, A., Kayser, M., Decorte, R., & Larmuseau, M. H. (2014). Seeing the wood for the trees: A minimal reference phylogeny for the human Y chromosome. *Human Mutation*, *35*, 187–191.

Vilar, M. G., Melendez, C., Sanders, A. B., Walia, A., Gaieski, J. B., Owings, A. C., & Schurr, T. G. (2014). Genetic diversity in Puerto Rico and its implications for the peopling of the Island and the West Indies. *American Journal of Physical Anthropology*, *155*, 352–368.

Wells, R. S., Yuldasheva, N., Ruzibakiev, R., Underhill, P. A., Evseeva, I., Blue-Smith, J., … Bodmera, W. F. (2001). The Eurasian heartland: A continental perspective on Y-chromosome diversity. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 10244–10249.

Y Chromosome Consortium. (2002). A nomenclature system for the tree of human Y-Chromosomal binary haplogroups. *Genome Research*, *12*, 339–348.

Yunusbayev, B., Metspalu, M., Järve, M., Kutuev, I., Rootsi, S., Metspalu, E., … Villems, R. (2011). The Caucasus as an asymmetric semipermeable barrier to ancient human migrations. *Molecular Biology and Evolution*, *29*, 359–365.

Yunusbayev, B., Metspalu, M., Metspalu, E., Valeev, A., Litvinov, S., Valiev, R., … Villems, R. (2015). The genetic legacy of the expansion of Turkic-Speaking nomads across Eurasia. *PLoS Genetics*, *11*, e1005068.

Zhadanov, S. I., Dulik, M. C., Markley, M., Jennings, G. W., Gaieski, J. B., Elias, G., & Schurr, T. G. (2010). Genetic heritage and native identity of the Seaconke Wampanoag Tribe of Massachusetts. *American Journal of Physical Anthropology*, *142*, 579–589.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Yardumian A, Shengelia R, Chitanava D, et al. Genetic diversity in Svaneti and its implications for the human settlement of the Highland Caucasus. *Am J Phys Anthropol*. 2017;00:000–000. https://doi.org/10.1002/ajpa.23324