

1* 1 2 3 3 3

¹Max Planck Institute for Evolutionary Anthropology, Department of Evolutionary Genetics, Deutscher Platz 6, D-04103, Leipzig, Germany

²Center of Integrative Genomics University of Lausanne CH-1015 Lausanne Dorigny, Switzerland

³Vavilov Institute of General Genetics, Gubkin str.3, 119991, Moscow, Russia

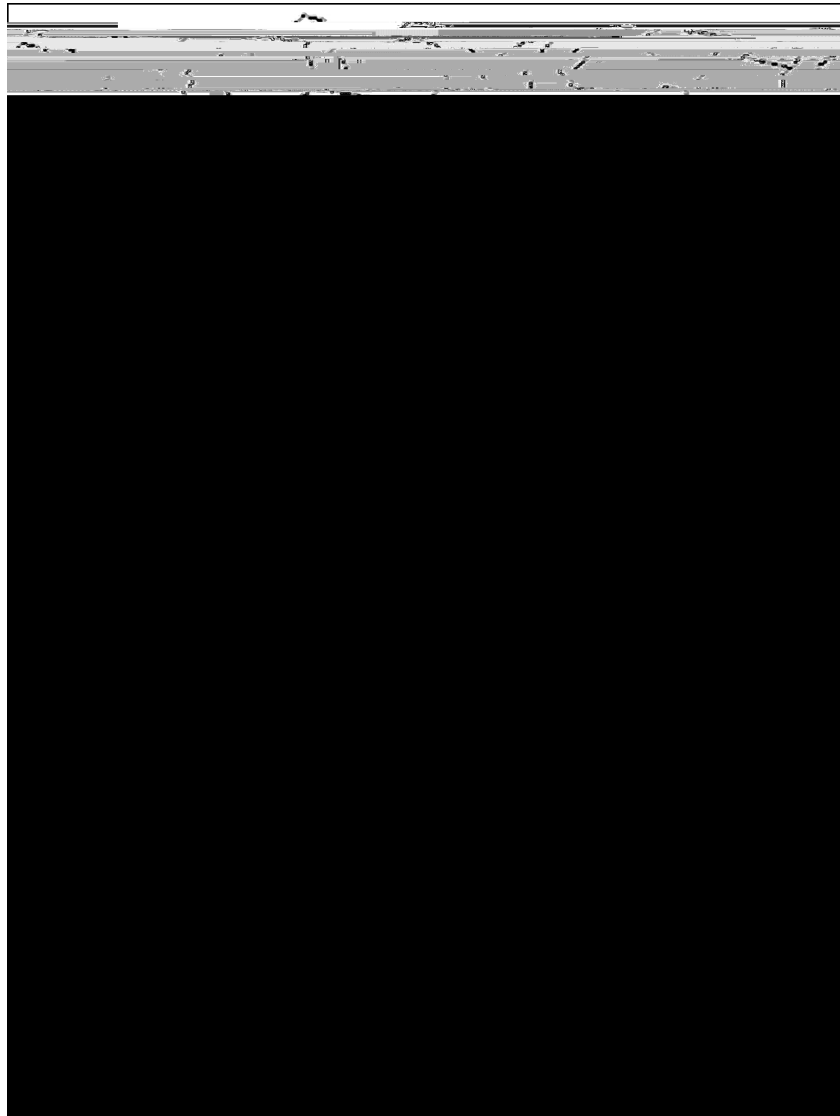
Ossetians are a unique group in the Caucasus, in that they are the only ethnic group found on both the north and south slopes of the Caucasus, and moreover they speak an Indo-European language in contrast to their Caucasian-speaking neighbours. We analyzed mtDNA HV1 sequences, Y chromosome binary genetic markers, and Y chromosome short tandem repeat (Y-STR) variability in three North Ossetian groups and compared these data to published data

Caucasian groups. Rychkov *et al.* (1996) also studied classical genetic markers in a number of North Ossetian groups. Genetic differentiation between Ossetian groups was the highest among all comparisons of groups from the Caucasus. Surprisingly, North Ossetians showed greater similarity with Russians than with neighbouring populations from the Caucasus. This observation was interpreted as genetic evidence of an ancient northern Alanian origin for Ossetians.

Eleven polymorphic restriction sites in the first hypervariable segment (HVI) of the mtDNA control region were studied in 40 North Ossetians and in additional groups from Siberia (Rychkov *et al.* 1995). A comparison of these groups with each other and with data from Western and Eastern Europe, as well as Central and Eastern Asia, revealed close similarity of all the Eurasian groups with respect to the frequency of each polymorphic restriction site.

Eleven bi-allelic loci and 9 short tandem repeat (STR) loci on the Y-chromosome, as well as mtDNA HV1 sequence variability, were previously studied in two populations from North Ossetia (Nasidze *et al.* 2004). One South Ossetian population, from Georgia, has also been studied for a number of Y chromosome SNPs (Wells *et al.* 2001) and for mtDNA HV1 sequence variability (Kivisild *et al.* 1999). However, the focus of these studies was on a much larger regional scale and did not address specific questions concerning the origin of Ossetians.

To investigate the origins and relationships of North and South Ossetian groups, here we present data on mtDNA and Y chromosome variation in 70 individuals from three additional North Ossetian groups. These new data were combined with the previously-published data from two other North and one South Ossetian group (Kivisild *et al.* 1999; Wells *et al.* 2001; Nasidze *et al.* 2003; Nasidze *et al.* 2004), to address the following questions: (1) are North Ossetians genetically more closely related to South Ossetians or to other North



A map of the geographic location of Ossetian populations used in this study. 1 - Alagir, 2 - Digora, 3 - Ardon, 4 - Zamankul, 5 - Zil'ga, 6 - South Ossetians.

(Comas *et al.* 2000), 55 Kazakhs and 94 Kyrgyz (Comas *et al.* 1998).

Ten Y chromosomal SNP markers were genotyped: RPS4Y (M130), M9, M89, M124, M45, M173, M17, M201, M170, and M172 (Underhill *et al.* 2000 and references therein); the YAP *Alu* insertion polymorphism (Hammer & Horai, 1995) was also typed. M9 and RPS4Y were typed by PCR-RFLP assays as described elsewhere (Kayser *et al.* 2000); M89 was typed as described in Ke *et al.* (2001); and the remaining Y-SNP markers were typed using PIRA-PCR assays (Yoshimoto *et al.* 1993) as described in Cordaux *et al.* (2004) and Nasidze *et al.* (2004). The YAP *Alu* inser-

tion was typed as described in Hammer & Horai (1995). The samples were genotyped according to the hierarchical order of the markers (Underhill *et al.* 2000). The Y-SNP haplogroup nomenclature used here is according to the recommendations of the Y Chromosome Consortium (2002).

Published Y-SNP data (Semino *et al.* 2000, Wells *et al.* 2001) for European, West Asian, and Central Asian populations were also included in some analyses.

R

Samples belonging to the SNP haplogroup G* (M201) were genotyped for nine Y chromosome short tandem repeat (Y-STR) markers: DYS19 (DYS394), DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391,

DYS392, and DYS393. These loci were amplified in pentaplex and quadruplex PCRs, or alternatively in a single nanoplex PCR, and detected on an ABI PRISM 377 DNA sequencer (Applied Biosystems), as described elsewhere (Kayser *et al.* 1997, 2001). In order to distinguish between the DYS385a and DYS385b loci, an additional PCR was carried out as described in Kittler *et al.* (2003) and detected on an ABI PRISM 377 DNA sequencer (Applied Biosystems).

Basic parameters of molecular diversity and population genetic structure, including analyses of molecular variance (AMOVA), were calculated using the software package Arlequin 2.000 (Schneider *et al.* 2000). The statistical significance of F_{st} values was estimated by permutation analysis, using 10,000 permutations. The statistical significance of the correlation between genetic distance matrices based on mtDNA and Y chromosome SNP data was evaluated by the Mantel test with 10,000 permutations. The STATISTICA package (StatSoft Inc.) was used for multi dimensional scaling (MDS) analysis (Kruskal, 1964). Network analysis for Y-STR and mtDNA HVI sequence data was carried out using the software package NETWORK version 3.1 (Bandelt *et al.* 1999).

In order to define geographic regions of large genetic changes (or genetic barriers), spatial analysis of molecular variance (SAMOVA) was used (Dupanloup *et al.* 2002). SAMOVA determines groups of populations that are geographically and genetically homogeneous and maximally differentiated from each other, based solely on genetic data. The method is based on a simulated annealing procedure that aims at maximizing

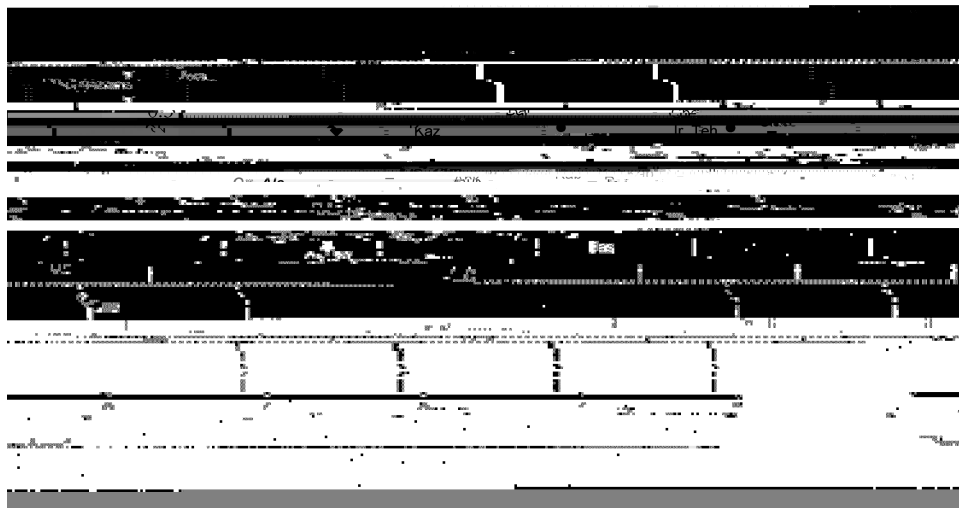
P Mean pairwise F_{st} values between North and South Ossetian groups, and caucasian, Iranian, West and East European, and West and Central Asian groups. Below diagonal - pairwise F_{st} values based on Y-SNP haplogroups; above diagonal - pairwise F_{st} values based on mtDNA HVI sequences

	N.Oss.	S.Oss.	N.Cau.	S.Cau.	Iranians	West Eu.	East Eu.	Centr. Asia	West Asia
N. Oss.		0.037	0.020	0.030	0.024	0.043	0.035	0.035	0.029
S. Oss.	0.330		0.037	0.033	0.015	0.036	0.030	0.073	0.018
N. Cau.	0.231	0.101		0.019	0.021	0.022	0.026	0.023	0.018
S. Cau.	0.298	0.126	0.190		0.016	0.025	0.027	0.021	0.020
Iranians	0.199	0.112	0.105	0.124		0.018	0.016	0.055	0.011
West Eu.	0.387	0.338	0.316	0.209	0.264		0.023	0.078	0.026
East Eu.	0.317	0.292	0.296	0.191	0.116	0.275		0.087	0.027
Centr. Asia	0.303	0.166	0.245	0.225	0.106	0.306	0.149		0.056
West Asia	0.211	0.087	0.135	0.253	0.093	0.243	0.183	0.121	

within the range of haplotype diversity values for other Caucasian groups. The mean number of pairwise nucleotide differences (MPD) varied from 4.65 (Ardon) to 6.70 (Alagir); these values are at, or exceed, the upper limit of the range of MPD values for European populations (3.15-5.03; Comas *et al.* 1997) as well as for the rest of the Caucasus (4.40-5.87; Nasidze *et al.* 2004). Tajima's D values were negative in all Ossetian groups (Table 1), significantly so in all groups except the Digora and the South Ossetian group. The negative D values suggest past population expansions in the Ossetian groups.

For overall pairwise comparisons of North and South Ossetians with neighbouring populations from the Caucasus, as well as groups from West and East Europe and West and Central Asia, we calculated average pairwise F_{st} values, pooling the five North Ossetian groups, as F_{st} values for almost all comparisons between North Ossetian groups are not significantly different from zero (not shown). Pairwise F_{st} comparisons (Table 2) indicate that South and North Ossetians are no closer to each other (average $F_{st} = 0.037$) than to North or South Caucasus groups (average F_{st} values range from 0.020-0.037). These average F_{st} values are not significantly different from one another ($t = 0.976$, $p = 0.332$). Indeed, South and North Ossetians are more distant from each other than are other South Caucasus groups from other North Caucasus groups (average $F_{st} = 0.019$), although the difference between these values is also not significant ($t = 0.817$, $p = 0.460$). North Ossetian populations are more similar to their geographic neighbours from the North Caucasus (average $F_{st} = 0.020$) than South Ossetians are to their

Nor=0/F1NorO52rai30(o Tm(=)Tj/F1 1 Tf1.200(o60 T0(erage)-336.8(F)]TJ6.9738 0 839rage)-336J6.9i.904601 Tm(st)Tj/F8 1 T



MDS plots based on pairwise F_{st} values, showing relationships among the North and South Ossetians, Caucasian, European, Central and West Asian populations. Ossetians are represented by stars; other Caucasus groups are represented by circles; squares correspond to populations from Europe; Central Asian groups are represented by diamonds; and West Asian populations by triangles. A. Based on mtDNA HVI sequence data. The stress value for the MDS plot is 0.104. B. Based on Y chromosome SNP data. The stress value for the MDS plot is 0.126. The populations are given the following abbreviations: Os_Dig – Ossetians from Digora, Os_Ala – Ossetians from Alagir, Os_Zil – Ossetians from Zil'ga, Os_Zam – Ossetians from Zamankul, Os_Ard – Ossetians from Ardon, S.Os – South Ossetians, Sv – Svans, Rut – Rutulians, Kazb – Kazbegi, Lez_Dag – Lezginians (Arletians),-2721.864 0 e.a14BT8.966g[(foa14BT8.321s/

P Fixation indices corresponding to the groups of populations as inferred by SAMOVA

A. Only Ossetian groups were considered in SAMOVA analysis. B. Along with Ossetian groups, other populations from the Caucasus and East Europe were included.

Locus	Group composition	F _{SC}	F _{ST}	F _{CT}			
mtDNA	<i>A. Between Ossetian groups</i>	-0.011	0.035***	0.045			
	1. North Ossetians (Alagir)						
	2. North Ossetians (Zamankul)						
	3. Other samples						
Y chromosome	<i>B. Between Ossetians and other neighboring groups</i>	0.019***	0.061***	0.043***			
	1. Karachaians						
	2. North Ossetians (Alagir)						
	3. North Ossetians (Zamankul)						
Y chromosome	<i>A. Between Ossetian groups</i>	-0.011	0.035***	0.045			
	1. South Ossetians						
	2. North Ossetians (Ardon)						
	3. North Ossetians (Zamankul)						
	4. Other samples						
	<i>B. Between Ossetians and other neighboring groups</i>				0.100***	0.270***	0.189**
	1. Russians						
	2. Georgians (Kazbegi)						
	3. Other samples						

***P < 0.001, **P < 0.005.

populations in the North and South Caucasus, respectively)? The results are somewhat different for mtDNA vs. the Y-chromosome. North and South Ossetians do cluster somewhat in the MDS plot based on mtDNA (Fig. 2A), which may indicate a common origin. However, for the Y-chromosome, North Ossetians are more similar to other North Caucasian populations, and South Ossetians to other South Caucasian populations, than to each other. The SAMOVA analysis also iden-

- human Y chromosome explained by population history. *Eur J Hum Genet* , 304–314.
- Kivisild, T., Bamshad, M. J., Kaldma, K., Metspalu, M., Metspalu, E., Reidla, M., Laos, S., Parik, J., Watkins, W.S., Dixon, M. E., Papiha, S. S., Mastana, S. S., Mir, M. R., Ferak, V. & Villems, R. (1999) Deep common ancestry of Indian and western-Eurasian mitochondrial DNA lineages. *Curr. Biol* , 1331–1334.
- Kruskal, J. B. (1964) Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* , 1–27.
- Macaulay, V., Richards, M., Hickey, E., Vega, E., Cruciani, F., Guida, V., Scozzari, R., Bonne-Tamir, B. & Sykes, B., Torroni, A. (1999) The emerging tree of West Eurasian mt DNAs: a synthesis of control-region sequences and RFLPs. *Am. J. Hum. Genet.* , 232–249.
- Maliarchuk, B. A., Derenko, M. V. & Solovenchuk, L. L. (1995): Types of regulatory regions in mitochondrial DNA in eastern Slavs. *Genetika* , 846–851.
- Miller, V. F. (1992) Ossetian essays. Vladikavkaz, SOIGI.
- Maniatis, T., Fritsh, E. F., Sambrook, J. (1982) Molecular cloning. A laboratory manual. New York: Cold Spring Laboratory.
- Nasidze, I., & Stoneking, M. (2001) Mitochondrial DNA variation and language replacements in the Caucasus. *Proc R Soc Lond B* , 1197–1206.
- Nasidze, I., Schadlich, H. & Stoneking, M. (2003) Haplotypes from the Caucasus, Turkey and Iran for nine Y-STR loci. *Forensic Sci Int* , 85–93.
- Nasidze, I., Ling, E Y S & Quinque, D., Dupanloup, I., Cordaux, R., Rychkov, S., Naumova, O., Zhukova, O., Sarraf-Zadegan, N., Naderi, G. A., Asgary, S., Sardas, S., Farhud, D. D., Sarkisian, T., Asadov, C., Kerimov, A. & Stoneking, M.. (2004) Mitochondrial DNA and Y-chromosome variation in the Caucasus. *Ann. Hum. Genet.* (in press)
- Orekhov, V., Poltoraus, A., Zhivotovsky, L.A., Spitsyn, V., Ivanov, P. & Yankovsky, N. (1999) Mitochondrial DNA sequence diversity in Russians. *FEBS letters* , 197–201.
- Piercy, R., Sullivan, K. M., Benson, N. & Gill, P. (1993) The application of mitochondrial DNA typing to the study of white Caucasian genetic identification.