Int J Legal Med (2009) 123:315–325
DOI 10.1007/s00414-009-0348-5

ORIGINAL ARTICLE

Pigment phenotype and biogeographical ancestry from ancient skeletal remains: inferences from multiplexed autosomal SNP analysis

Caroline Bouakaze · Christine Keyser · Eric Crubézy · Daniel Montagnon · Bertrand Ludes

Received: 16 December 2008 / Accepted: 7 April 2009 / Published online: 5 May 2009
© Springer-Verlag 2009

Abstract In the present study, a multiplexed genotyping assay for ten single nucleotide polymorphisms (SNPs) located within six pigmentation candidate genes was developed on modern biological samples and applied to DNA retrieved from 25 archeological human remains from southern central Siberia dating from the Bronze and Iron Ages. SNP genotyping was successful for the majority of ancient samples and revealed that most probably had typical European pigment features, i.e., blue or green eye color, light hair color and skin type, and were likely of European individual ancestry. To our knowledge, this study reports for the first time the multiplexed typing of autosomal SNPs on aged and degraded DNA. By providing valuable information on pigment traits of an individual and allowing individual biogeographical ancestry estimation, autosomal SNP typing can improve ancient DNA studies and aid human identification in some forensic casework situations when used to complement conventional molecular markers.

C. Bouakaze (✉) · C. Keyser · D. Montagnon · B. Ludes
Laboratoire d'Anthropologie Moléculaire,
Institut de Médecine Légale,
11 rue Humann,
67 085 Strasbourg Cedex, France
e-mail: caroline.bouakaze@ulp.u-strasbg.fr

C. Keyser · E. Crubézy · B. Ludes
Laboratoire d'Anthropobiologie, FRE 2960 CNRS,
Université Paul Sabatier,
39 Allée Jules Guesdes,
31 000 Toulouse, France

Keywords Autosomal SNPs · Pigmentation · Ancestry · Ancient DNA · SNaPshot

Introduction

It is well known that ancient DNA (aDNA) studies are plagued by a high risk of contamination with contemporary DNA due to the limited amount and degraded nature of endogenous DNA that is retrievable from ancient specimens [1]. Most of these studies have so far relied on the analysis of variation in maternally inherited mitochondrial DNA (mtDNA). This is due to the fact that hundreds to thousands of copies of mtDNA are available per cell making mtDNA more likely to survive in degraded specimens than single copy nuclear DNA (nuDNA) [2]. Nevertheless, recently, there have been a growing number of aDNA studies that report the successful recovery of nuDNA [3–7]. Almost all of these studies performed on human remains have focused on the amelogenin gene for sex identification and on short tandem repeat loci (STRs) using commercially available typing kits developed for forensic applications [8–12]. Autosomal STR typing generally enables individual identification and deciphering close kinship relationship, while Y-chromosomal STRs are useful to follow-up paternal lineages analogous to the conventional sequencing of mtDNA hypervariable region I (HV-I) that allows the following of maternal lineages. A lot of additional information could be obtained by typing nuclear single nucleotide polymorphisms (SNPs) since these polymorphisms, which are the most abundant in the human genome, have been associated to several human phenotypes. Very few aDNA studies have attempted to type

nuclear SNPs for investigating phenotypic characteristics (e.g., disease susceptibility [13] and lactase persistence [14]). Evidence that pigment features, one of the most observable phenotypic traits, can be inferred by analyzing variation within ancient nuDNA genes was recently provided by Lalueza-Fox and colleagues on Neanderthal specimens [15]. By retrieving a sequence of the melanocortin 1 receptor (MC1R) gene, a key pigmentation gene, which showed a mutation altering the MC1R activity, their study revealed that at least some Neanderthals probably had red hair and pale skin.

Pigmentation, including skin eye and hair color, is also one of the most variable human phenotypic traits. Variation between individuals is mostly dependent on the shape and distribution of melanosomes (which are specialized organelles of melanocytes in which melanin pigments are synthesized and stored) and the type of melanin synthesized (black/brown eumelanin or yellow/red pheomelanin) [16]. The genetic architecture underlying normal pigmentation variation in humans has not yet been fully elucidated. Although more than 100 color loci have been identified in the mouse genome [17], only around 12 have been characterized in humans. A simplified representation of melanogenesis with the major proteins involved is shown in Fig. 1. Briefly, the eumelanin biosynthesis pathway is triggered by stimulation of the key protein MC1R via binding of the melanocyte stimulating hormone (α-MSH). This results in increased levels of cAMP in the melanocyte, which activates the microphtalmia transcription factor (MITF). MITF stimulates tyrosinase (TYR), tyrosine-related protein 1 (TYRP1), and dopachrome tautomerase (DCT), enzymes required for the transformation of tyrosine to eumelanin. Other proteins are required in the eumelanin biosynthesis and eumelanosome maturation such as P-protein, SLC24A5, and MATP. Pheomelanogenesis seems to be the default pathway when MC1R signaling is altered, which can be the result of antagonism of MC1R by agouti-signaling protein (ASIP) or high concentration of cysteine.

Recently, various SNPs in pigmentation candidate genes have been identified and associated to specific pigment phenotypes. Some of these polymorphisms also show large allele frequency differences among populations, which is not surprising since differences in pigmentation are very high between individuals of various geographical origins. Such SNPs belong to a class of genetic marker known as ancestry-informative markers (AIMs) that are widely used to infer biogeographical ancestry at the population level in order to study population genetic structure but can also be used at the individual level. Both types of SNPs, i.e., phenotype-informative and ancestry-informative SNPs, are greatly interesting for the fields of biomedical research as genetic differences between geographical groups may account for health disparities between populations [18] and forensic genetics because providing information on an individual's physical appearance and geographical origin based on DNA analysis would aid the tracing of unknown suspects [19].

The aim of the present study was (1) to evaluate the possibility of typing several autosomal SNPs simultaneously on DNA retrieved from ancient skeletal remains and (2) to investigate the possibility of inferring pigment traits and individual biogeographical ancestry of individuals from the autosomal genotype data obtained. For this purpose, a panel

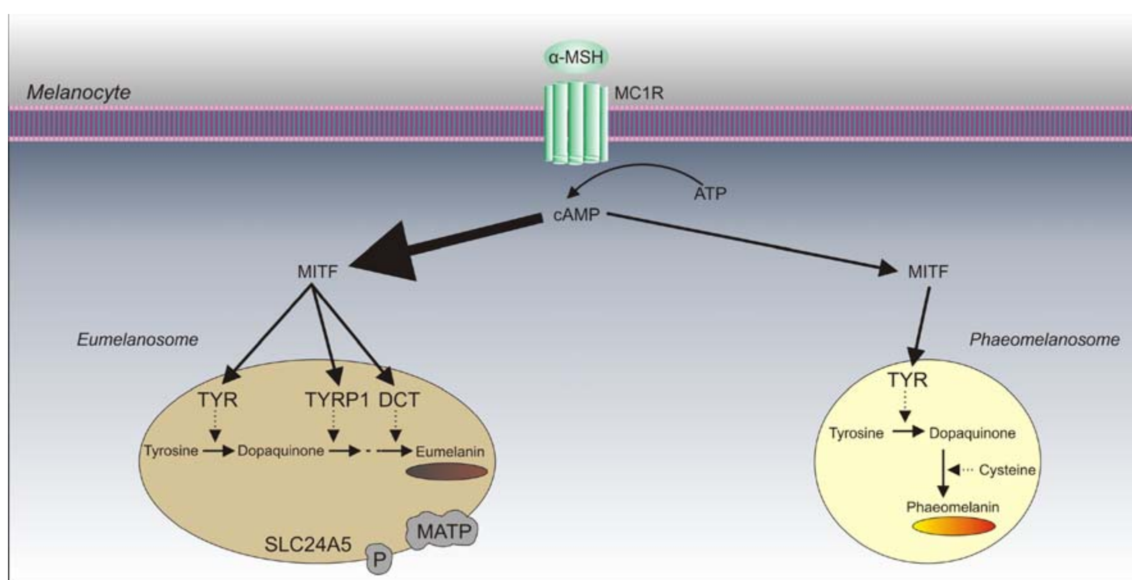

Fig. 1 Schematic representation of key proteins involved in melanogenesis

of ten phenotype and ancestry-informative SNPs were selected. A multiplex assay was developed using the SNaPshot multiplex kit for the genotyping of these SNPs on modern DNA samples as it was previously showed that this methodology is convenient, robust, and effective for SNP typing on low amounts of aged and degraded DNA [20]. Finally, it was applied to 25 aDNA samples from southern Siberia that have previously been successfully analyzed for variation in mtDNA and the Y-chromosome [20, 21].

Materials and methods

Modern and archeological samples

Modern biological samples consisted of buccal swabs or blood dropped onto blotting paper from 27 unrelated Europeans (France), three Africans (Togo), and six Asians (Yakutia and Mongolia). All donors signed an informed consent for use of their DNA, and all except Asians were asked to determine their own pigment characteristics, including eye (blue, green, gray, or brown), hair (red/auburn, blond, light or dark brown, or black), and skin (fair, medium, or black) color.

Archeological samples consisted of 25 long bone fragments from the Krasnoyarsk region of southern central Siberia, which covered four successive cultures from the Middle Bronze Age to the Iron Age (third millennium B.C. to the fourth century A.D.). These samples had previously yielded reproducible autosomal STR profiles and mtDNA and Y-chromosomal haplotypes and haplogroups.

DNA extraction

Modern samples were subjected to genomic DNA extraction using standard phenol–chloroform procedure, purification, and quantitation as described elsewhere [20].

Archeological samples were pre-treated to remove surface contamination using a rotary drill, and genomic DNA extraction and purification were done carefully according to a published protocol [22].

SNP selection and genotyping

Ten potentially informative SNPs were selected from the literature according to their association to normal human pigmentation variation or differences in published allele frequencies among Europeans, East Asians, and Africans. Phenotype-informative SNPs include two non-synonymous MC1R mutations, rs1805007 and rs1805008, which have been associated to red hair phenotype [23] and four intron SNPs associated to eye color variation. The latter ones included three SNPs, rs7495174, rs6497268, and rs11855019, located on the OCA2 gene, which when grouped together into a single haplotype block have also been associated

Table 1 General information about the ten SNPs selected in this study

Gene	Location	Protein	Reference SNP ID (rs#)[a]	Alleles	Variation type
MC1R	16q24.3	MC1R: melanocortin 1 receptor	rs1805007	C/T	ns coding, c.451C>T, p.R151C
			rs1805008	C/T	ns coding, c.478C>T, p.R160W
HERC2	15q13	Unknown	rs12913832	A/G	Non-coding, intron 86
OCA2	15q11.2-15q12	P-protein: NA+/H+ antiporter or glutamate transporter	rs7495174	T/C	Non-coding, intron 1
			rs6497268 or rs4778241	G/T	
			rs11855019 or rs4778138	T/C	
			rs1545397	G/A	Non-coding intronic
SLC45A2	5p13.3	MATP: membrane-associated transporter protein	rs16891982	C/G	ns coding, c.1122C>G, p.F374L
SLC24A5	15q21.1	SLC24A5 (or NCKX5): solute carrier family 24, member 5; potassium-dependent sodium–| calcium ion exchanger	rs1426654	G/A	ns coding, p.A111T
DCT	13q32	DCT or TYRP2/TRP-2: dopachrome tautomerase or tyrosinase-related protein-2	rs2031526	G/A	Non-coding, intronic

ns non-synonymous

[b] Reference SNP ID refer to the reference sequence identifier given to the SNP in the dbSNP database

to hair and skin color variation [24], and one SNP, rs12913832, located on the HERC2 gene [25, 26]. Ancestry-informative SNPs include two non-synonymous SNPs, rs16891982 and rs1426654, that have been associated to light-skinned European ancestry and are located on the SLC45A2 and SLC24A5 genes, respectively [27–29], and two SNPs, rs2031526 and rs1545397, that show large allele frequency differences between Asians and other populations and are located on the DCT and OCA2 genes, respectively [30, 31]. Detailed information regarding the SNPs selected is reported in Table 1.

Genotyping consisted of a minisequencing-based approach as previously described [20]. The multiplexed reactions were developed according to the recommendations of Römpler and colleagues [32] and Sanchez and colleagues [33]. Briefly, the ten DNA fragments encompassing the SNP positions of interest were amplified within a multiplex PCR. Sequences and concentrations of PCR primers used in the 10-plex PCR mix are specified in Table 2. PCR products were purified using the NucleoSpin® Extract II isolation Kit (Macherey-Nagel, Düren, Germany) following the manufacturer's recommendations and subjected to a multiplex single base extension (SBE) reaction using the ABI PRISM® SNaPshot® Multiplex Kit (Applied Biosystems, AB, Foster City, CA, USA). Sequences and concentrations of SBE primers in the 10-plex SBE mix are specified in Table 3. After enzymatic treatment, extension products were subjected to capillary electrophoresis on an ABI PRISM 3100® Genetic Analyzer (AB), and data were analyzed using the GeneMapper® ID v.3.2.1 software (AB).

The same procedure as for contemporary DNA was applied to aDNA samples except that the number of PCR and SBE reaction cycles was increased to 37 and 30, respectively. Moreover, SNPs that failed to amplify in the multiplex reactions were further genotyped in singleplex reactions. For samples where the amount of available aDNA was limited, genotyping of the four ancestry-informative SNPs and rs12913832 was privileged over other markers.

Precautions against contamination

Precautions to avoid contamination with modern DNA have been previously published [10]. In addition, new extracts were subjected to STR typing using the AmpFlSTR® Identifiler® PCR Amplification Kit (AB) according to the manufacturer's instructions. STR profiles obtained in this study were compared to those previously obtained on the same ancient samples using AmpFlSTR® Profiler® Plus PCR Amplification Kit (AB) and to those from all laboratory personnel.

Ancestry inference

The individual biogeographical ancestry of modern and ancient specimens was estimated using the STRUCTURE 2.2 software [34]. The Bayesian method implemented in this program is widely used to detect population genetic structure but can also be used to assign individuals to predefined populations. The three HapMap population groups, i.e., Africans (Yoruba people of Ibadan, Nigeria), Europeans (U.S. residents with ancestry from northern and western Europe), and East Asians (Japanese from Tokyo and Han Chinese from Beijing), consisting of 251 individuals were used as predefined populations [35]. The complete genotype data of these individuals for the four ancestry-informative SNPs selected in this study was obtained on the dbSNP database of the National Center for Biotechnology Information. To test whether the markers selected were discriminant enough, the program was first run without population information.

Table 2 PCR primers for the multiplex amplification of the ten SNPs used in this study

Reference SNP ID (rs#)	PCR primer sequence (5′→3′)		Amplicon size (bp)	Concentration (μM)
	Forward	Reverse		
rs12913832	TTGTTCTTCATGGCTCTCTGTGTCTG	AGAGAAGCCTCGGCCCCTGA	108	0.4
rs1805007	CCGCTACATCTCCATCTTCTACGC	GCGTGCTGAAGACGACACTGG	110	0.55
rs1805008	ATCTTCTACGCACTGCGCTACCAC	GAAGAGCGTGCTGAAGACGACAC	102	0.55
rs7495174	AGGCCCAGGCGGACTCAG	AGGCAGGGAGGGTTTACACAGC	128	0.5
rs6497268	TACAGCCACTCTGGAAAGCAGTTTG	TTGCGTGTAGGGTTTTGTGTGAATG	133	0.4
rs11855019	CCAGGATTCAAAAAGAAAGTCTCAAGG	TTTCCTCCCATCACTGATTTAGCTG	115	0.55
rs16891982	CCCTATAGTGCACACAACTCCACAGAG	TGAGGAAAACACGGAGTTGATGC	96	0.3
rs1426654	CCCTTGGATTGTCTCAGGATG	TGAGTAAGCAAGAAGTATAAGGAGCAA	118	0.1
rs2031526	CCTTGAATTGCTCTTGAAAAACTAA	CAGCCCAATGATACACTTTCATTTAAC	149	0.55
rs1545397	TGGAATTGGATACTGACAATGGTTG	CATGGGGGAGAGAGAATGACTCAG	144	0.55

Results

Electropherograms obtained from three modern samples of different descent and one ancient sample are presented in Fig. 2. The successful amplification and typing of target SNPs from both modern and ancient samples demonstrate the robustness of this genotyping method.

Genotyping of modern samples

We were able to reliably genotype all SNPs for all modern samples. Self-reported phenotypic characteristics (eye, hair, and skin color) and complete genotypes obtained for the modern samples are given in Table 4. Eye, hair, and skin color as well as individual ancestry estimated from the genotypes were in complete agreement with those reported by the donors themselves.

Among the six red/auburn-headed Europeans tested, only one was found to be T/T homozygous for one of the two MC1R SNPs. Interestingly, all the remaining subjects were heterozygous for only one of the two variants, while Branicki and co-workers [27] reported an association between a compound heterozygosity or a T/T homozygosity for one of two MC1R variants and red hair phenotype.

The majority of modern European samples carried the TGT/TGT OCA2 intron 1 diplotype and of those 83% had blue or green eye color and 17% had brown eye color. All except one had fair hair color (blond, red/auburn, or light brown) and a fair skin type. These results are consistent with frequencies reported by Duffy et al. [24] who found this genotype at the highest frequency in subjects with blue or green eye color, light hair, and fair or medium skin type. The remaining Europeans had either the TGT/CTC, TGT/TGC, TGT/TTT, or TTT/CTC diplotype. The TGT haplotype was only found in modern European samples, which is also consistent with the strong positive selection for this haplotype in Europeans [24]. The three African subjects tested had brown eye color, black hair, and dark skin, which could be predicted from their genotype data. All had the TGC/TTC diplotype, which was found at a high frequency in subjects with brown eye color, black hair, and dark/olive skin type by Duffy et al. [24]. Unfortunately, no phenotypic information was available for the six Asian samples tested but they probably had brown eye color, dark brown or black hair, and medium or olive/dark skin color since they had the CTC/CTC, TTT/CTC, or TTT/CGC diplotype. The TTC/CGC diplotype was found for one Asian individual but this diplotype was not found in the populations analyzed in the study of Duffy et al.

All 16 blue or green-eyed individuals tested had the G/G genotype for rs12913832. This is in accordance with previous studies, which suggested that homozygosity of the rs12913832*G allele causes blue eye color phenotype in humans, defined as a complete lack of brown pigmentation [25, 26]. Other modern samples tested, including all Africans and Asians, which were either blue/hazel, green/hazel, or brown eyed subjects, had either the A/A or A/G genotype for rs12913832, which is predictive of brown eye color phenotype.

Self-reported geographical origin of all modern samples agreed well with individual ancestry inferred by the program STRUCTURE, which shows that the four SNPs selected are informative and complementary enough to estimate the biogeographical ancestry of our ancient specimens.

Genotyping of ancient samples

The genotype data for the 25 ancient samples are given in Table 5. Complete genotype data for the ten SNPs were obtained from eight aDNA samples among the 25 tested. All were distinctive at least at one loci demonstrating that systemic contamination had not occurred.

The genotype for the two non-synonymous SNPs located on MC1R gene was obtained for 11 aDNA samples.

Table 3 Minisequencing primers for the detection of the ten SNPs used in this study

Reference SNP ID (rs#)	Poly(dC)	Neutral sequence (5′→3′)	Target specific sequence (5′→3′)	(*)	Primer size (nt)	Concentration (μM)
rs12913832	–	AGTCTGACAA	TAGCGTGCAGAACTTGACA	R	29	0.3
rs1805007	–	AAAGTCTGACAA	CATCTCCATCTTCTACGCACTG	F	34	0.58
rs1805008	–	GCCACGTCGTGAAAGTCTGACAA	CATCGTGACCCTGCCG	F	39	0.58
rs7495174	–	GCCACGTCGTGAAAGTCTGACAA	AAGGCAAGTTCCCCTAAAGGT	R	44	0.58
rs6497268	–	ACTAGGTGCCACGTCGTGAAAGTCTGACAA	TTGGCTGGTAGTTGCAATT	R	49	0.5
rs11855019	–	ACTAGGTGCCACGTCGTGAAAGTCTGACAA	CATCACTGATTTAGCTGTGTTCTG	R	54	0.58
rs16891982	1	AACTGACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAA	GGTTGGATGTTGGGGCTT	F	59	0.4
rs1426654	5	AACTGACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAA	GTCTCAGGATGTTGCAGGC	F	64	0.1
rs2031526	4	AACTGACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAA	CACTTTCATTTAACTTTTGCTTTTG	R	69	0.5
rs1545397	10	AACTGACTAAACTAGGTGCCACGTCGTGAAAGTCTGACAA	ATCTTGCAAAATTATATCATTCAG	R	74	0.58

(*) Orientation of primer

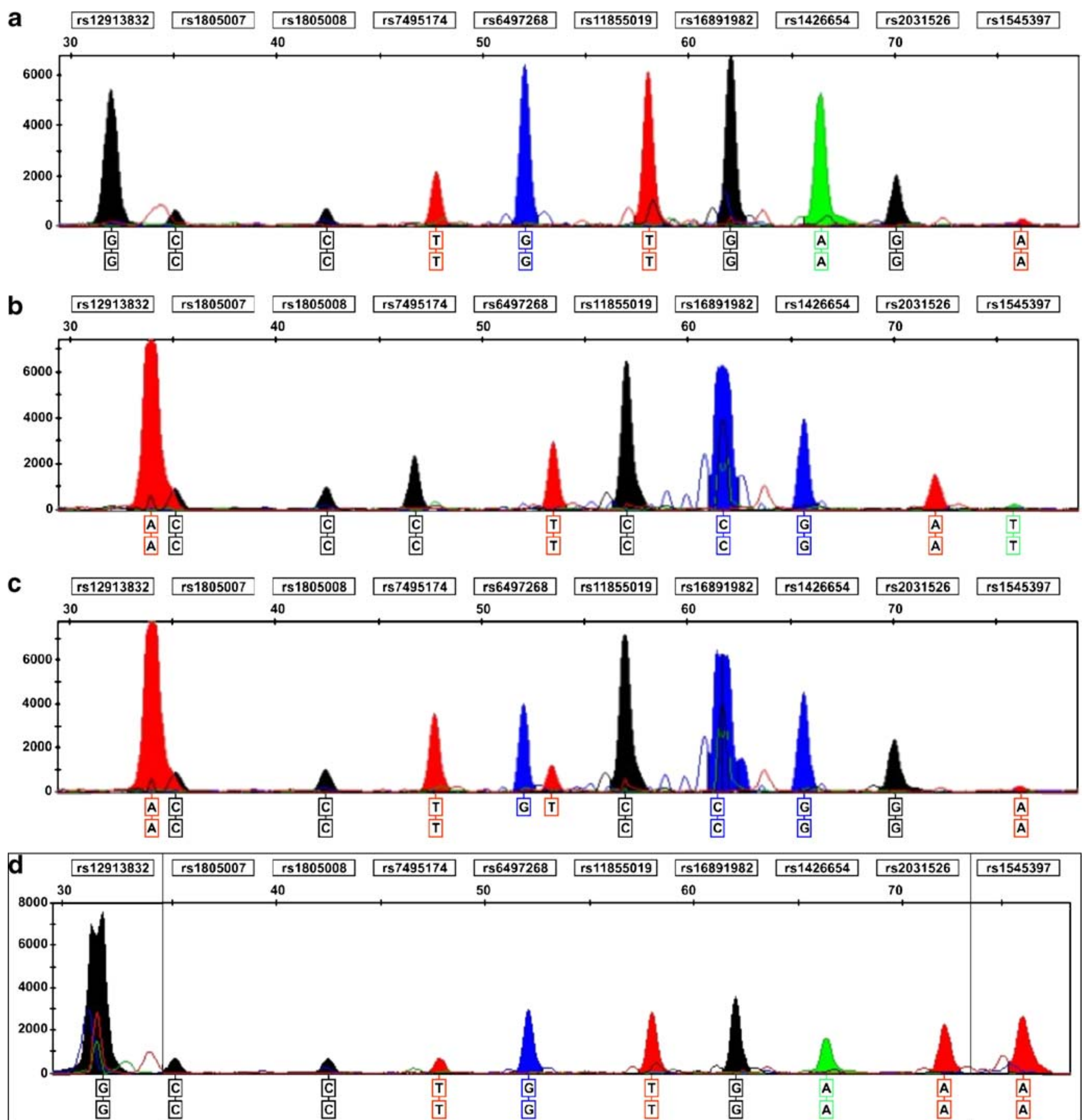

Fig. 2 Electropherograms obtained from the genotyping of three modern DNA samples of European (**a**), Asian (**b**), and African (**c**) origin and one aDNA sample, S33 (**d**). The last electropherogram (**d**) was constructed from three electropherograms resulting from the singleplex amplification of rs12913832 and rs1545397 and the multiplex amplification of the eight remaining SNPs

Neither T/T homozygosity nor compound heterozygosity for the two MC1R variants was observed.

The complete OCA2 diplotype was obtained for 12 samples, most of which carried the TGT/TGT diplotype (n=8), and the others had TGT/TGC (n=2), TGT/TTT (n=1), or TTT/CTC (n=1) diplotypes. These results strongly suggest that most of the ancient specimens had a typical light European pigment phenotype (i.e., blue or green eyes, light hair color, and skin type). This high frequency of light eye color was further confirmed by the analysis of the intron HERC2 SNP.

The genotype for rs12913832 was obtained for 23 out of the 25 samples, and most had the G/G genotype (n=15), which indicates that at least 60% of ancient specimens were

Table 4 Genotypes obtained for the ten SNPs typed on 30 modern DNA samples and inferred ancestry

Sample	Self-reported pigmentary traits			rs12913832 HERC2	rs1805007 MC1R	rs1805008 MC1R	OCA2 diplotype[a]	rs16891982 SLC24A2	rs1426654 SLC24A5	rs2031526 DCT	rs1545397 OCA2	Inferred ancestry of individuals[b]		
	Eye color	Hair color	Skin color									European	Asian	African
E1	Blue	Red	Fair	G/G	C/C	C/T	TGT/TGT	G/G	A/A	G/G	A/A	**0.963**	0.012	0.024
E2	Green	Light brown	Fair	G/G	C/C	C/C	TGT/TGT	G/G	A/A	A/G	A/A	**0.954**	0.021	0.025
E3	Blue	Blond	Fair	G/G	C/C	C/C	TGT/TGT	G/G	A/A	A/G	A/A	**0.954**	0.024	0.022
E4	Blue	Blond	Fair	G/G	C/C	C/C	TGT/TGT	G/G	A/A	A/G	A/A	**0.960**	0.020	0.020
E5	Blue/gray	Auburn	Fair	G/G	C/T	C/C	TGT/TGT	G/G	A/A	G/G	A/A	**0.961**	0.013	0.026
E6	Green/gray	Light brown	Fair	G/G	C/C	C/C	TGT/TGT	C/G	A/A	G/G	A/A	**0.787**	0.038	0.175
E7	Green/hazel	Light brown	Fair	A/G	C/C	C/C	TGT/TGT	G/G	A/A	A/G	A/A	**0.955**	0.022	0.024
E8	Green/hazel	Dark brown	Fair	A/A	C/C	C/C	TGT/CTC	G/G	A/A	G/G	A/A	**0.961**	0.013	0.027
E9	Green/hazel	Dark brown	Fair	A/A	C/C	C/C	TTT/CTC	G/G	A/A	G/G	A/A	**0.963**	0.013	0.024
E10	Blue	Light brown	Fair	G/G	C/C	C/C	TGT/TGT	C/G	A/A	G/G	A/A	**0.789**	0.049	0.163
E11	Green	Auburn	Fair	G/G	C/T	C/C	TGT/TGC	G/G	A/A	G/G	A/A	**0.958**	0.014	0.028
E12	Blue/hazel	Light brown	Fair	A/G	C/C	C/C	TGT/TTT	G/G	A/A	G/G	A/A	**0.962**	0.012	0.026
E13	Blue/hazel	Light brown	Fair	A/G	C/C	C/C	TGT/TTT	G/G	A/A	G/G	A/A	**0.965**	0.013	0.022
E14	Green	Light brown	Fair	G/G	C/C	C/C	TGT/TGT	C/G	A/A	G/G	A/T	**0.763**	0.165	0.073
E15	Brown	Dark brown	Fair	A/G	C/C	C/C	TGT/TGT	G/G	A/A	A/G	A/A	**0.957**	0.022	0.021
E16	Brown	Dark brown	Fair	A/A	C/C	C/C	TGT/CTC	C/G	A/A	A/G	A/T	**0.669**	0.283	0.048
E17	Green/hazel	Dark brown	Medium	A/G	C/C	C/C	TGT/TTT	C/G	A/A	G/G	A/T	**0.755**	0.170	0.076
E18	Blue	Light brown	Fair	G/G	C/C	C/C	TGT/TGT	G/G	A/A	G/G	A/T	**0.935**	0.045	0.021
E19	Brown	Red	Fair	A/G	C/T	C/C	TGT/TGT	G/G	A/A	G/G	A/A	**0.964**	0.013	0.022
E20	Green	Light brown	Fair	G/G	C/C	C/C	TGT/TGT	C/G	A/A	G/G	A/A	**0.792**	0.047	0.161
E21	Green/gray	Blond	Fair	G/G	C/C	C/C	TGT/TGT	G/G	A/A	A/G	A/A	**0.957**	0.022	0.021
E22	Blue	Light brown	Fair	G/G	C/C	C/C	TGT/TGT	G/G	A/A	G/G	A/A	**0.959**	0.014	0.026
E23	Green/hazel	Light brown	Fair	A/G	C/C	C/C	TGT/TTT	G/G	A/A	A/G	A/A	**0.957**	0.020	0.022
E24	Green	Light brown	Fair	G/G	C/C	C/C	TGT/TGT	C/G	A/A	G/G	A/A	**0.786**	0.049	0.166
E25	Brown	Red	Fair	A/G	C/C	T/T	TGT/TGC	G/G	A/A	G/G	A/A	**0.963**	0.014	0.023
E26	Blue	Light brown	Fair	G/G	C/C	C/C	TGT/TGT	G/G	A/A	A/G	A/A	**0.954**	0.021	0.025
E27	Blue	Red	Fair	G/G	C/C	C/T	TGT/TGT	G/G	A/A	G/G	A/A	**0.958**	0.014	0.028
Af1	Brown	Black	Dark	A/A	C/C	C/C	TGC/TTC	C/C	G/G	A/G	A/A	0.028	0.094	**0.878**
Af2	Brown	Black	Dark	A/A	C/C	C/C	TGC/TTC	C/C	G/G	G/G	A/A	0.023	0.031	**0.946**
Af3	Brown	Black	Dark	A/A	C/C	C/C	TGC/TTC	C/C	A/G	G/G	A/A	0.164	0.041	**0.795**
As1	–	–	–	A/A	C/C	C/C	TTT/CTC	C/C	G/G	A/G	A/T	0.042	**0.649**	0.308
As2	–	–	–	A/A	C/C	C/C	CTC/CTC	C/C	G/G	A/G	T/T	0.020	**0.921**	0.060
As3	–	–	–	A/A	C/C	C/C	CTC/CTC	C/C	G/G	A/A	T/T	0.013	**0.964**	0.023
As4	–	–	–	A/G	C/C	C/C	TTT/CGC	C/C	A/G	A/A	A/T	0.212	**0.708**	0.080
As5	–	–	–	A/A	C/C	C/C	TTC/CGC	C/C	G/G	A/G	T/T	0.019	**0.922**	0.059
As6	–	–	–	A/A	C/C	C/C	CTC/CTC	C/G	G/G	A/A	T/T	0.119	**0.858**	0.023

E European modern sample, *Af* African modern sample, *As* Asian modern sample

[a] OCA2 diplotype correspond to markers rs7495174/rs6497268/rs11855019. OCA2 diplotype and rs12913832 genotype predictive of blue eye color phenotype are underlined

[b] Probability of being from European/Asian/African population determined using the STRUCTURE program. The greatest probability, most likely estimate of ancestry, is indicated in bold

Table 5 Genotypes obtained for the ten SNPs typed on 25 ancient samples and inferred ancestry

Sample	Culture	rs12913832 HERC2	rs1805007 MC1R	rs1805008 MC1R	OCA2 diplotype[a]	rs16891982 SLC24A2	rs1426654 SLC24A5	rs2031526 DCT	rs1545397 OCA2	Ancestry inferred from STRUCTURE[b]		
										European	Asian	African
S07	Andronovo	A/A	C/C	C/C	TTT/CTC	C/C	G/A	A/G	T/T	0.143	**0.811**	0.046
S08	Andronovo	A/A	–	–	–	C/C	G/G	A/A	A/T	0.030	**0.870**	0.100
S09	Andronovo	G/G	C/C	C/C	TGT/TGT	G/G	A/A	G/G	A/A	**0.967**	0.011	0.022
S10	Andronovo	-	C/C	C/C	TGT/TTT	G/G	A/A	A/G	A/A	**0.966**	0.017	0.017
S11	Andronovo	G/G	–	–	–	C/G	A/A	G/G	-	**0.837**	0.054	0.109
S13	Andronovo	–	C/C	C/C	TGT/TGT	G/G	A/A	A/G	A/A	**0.967**	0.016	0.017
S14	Andronovo	A/G	C/C	C/C	TGT/TGC	C/G	A/A	G/G	A/A	**0.861**	0.033	0.107
S15	Andronovo	G/G	C/C	C/C	−GT/−GT	G/G	A/A	A/A	A/A	**0.948**	0.033	0.018
S16	Andronovo	G/G	C/C	C/C	TGT/TGT	G/G	A/A	A/G	A/A	**0.964**	0.016	0.020
S18	Karasuk	G/G	C/C	–	–	G/G	A/G	G/G	-	**0.744**	0.075	0.181
S19	Karasuk	G/G	–	–	–	–	A/A	G/G	A/A	**0.942**	0.014	0.044
S21	Tagar	G/G	C/C	–	TGT/TGC	G/G	A/A	G/G	A/A	**0.967**	0.011	0.023
S22	Tagar	A/G	C/C	C/C	−GT/−GT	G/G	A/A	G/G	–	**0.967**	0.013	0.020
S23	Tagar	G/G	–	–	–	G/G	A/A	G/G	-	**0.967**	0.014	0.020
S24	Tagar	G/G	–	–	–	G/G	A/A	G/G	A/A	**0.971**	0.011	0.018
S25	Tagar	G/G	C/C	–	−GT/−GT	G/G	A/A	G/G	A/T	**0.946**	0.038	0.016
S26	Tagar	A/G	C/C	C/C	TGT/TGT	G/G	A/A	A/A	A/A	**0.953**	0.033	0.014
S28	Tagar	A/A	–	–	TGT/TGT	G/G	A/A	G/G	A/T	**0.946**	0.037	0.016
S29	Tagar	G/G	–	–	–	G/G	A/A	–	–	**0.963**	0.019	0.018
S32	Tagar	A/G	C/C	–	–	C/C	A/A	A/G	A/T	**0.491**	**0.419**	0.090
S33	Tashtyk	G/G	C/C	C/C	TGT/TGT	G/G	A/A	A/A	A/A	**0.948**	0.034	0.018
S34	Tashtyk	G/G	C/C	–	−GT/−GT	G/G	A/A	A/A	–	**0.928**	0.056	0.016
S35	Tashtyk	A/G	–	–	–	G/G	A/A	A/G	A/T	**0.920**	0.066	0.015
S36	Tashtyk	G/G	C/C	C/C	TGT/TGT	G/G	A/A	A/G	A/A	**0.962**	0.018	0.020
S37	Tashtyk	G/G	C/T	C/C	TGT/TGT	G/G	A/A	G/G	A/A	**0.969**	0.010	0.021

[a] OCA2 diplotype correspond to markers rs7495174/rs6497268/rs11855019. OCA2 diplotype and rs12913832 genotype predictive of blue eye color phenotype are underlined

[b] Probability of being from European/Asian/African population determined using the STRUCTURE program. The greatest probability, indicating the most likely individual ancestry, is indicated in bold

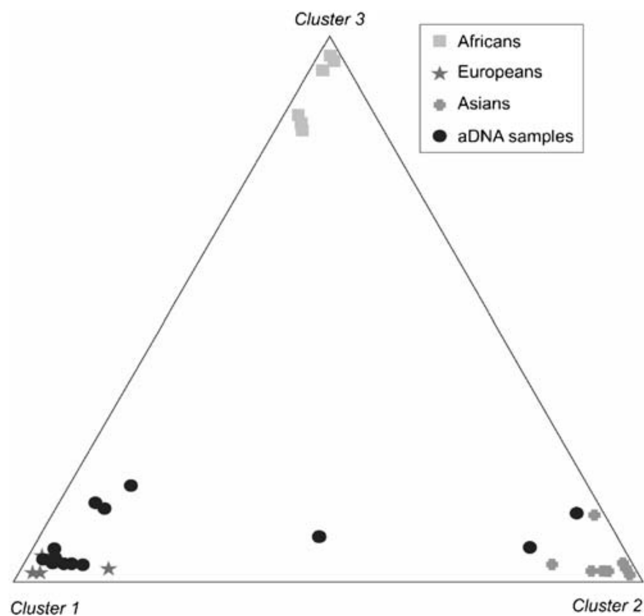

Fig. 3 Triangular plot showing the individual ancestry of each aDNA sample. This plot was computed with the STRUCTURE program (burning period=10,000; number of MCMC after burning=5,000)

probably blue- or green-eyed individuals. The remaining samples had the A/G (n=5) or A/A (n=3) genotypes, which are predictive of brown eye color phenotype.

Figure 3 shows a triangular plot of the relative autosomal contributions of European, African, and Asian populations in the gene pool of ancient samples. Using the four ancestry-informative SNPs and the program STRUCTURE, all ancient samples clearly had a high proportion of European ancestry except three. Two samples, S07 and S08, had a higher proportion of Asian ancestry, and one sample, S32, had an admixed ancestry with both European and Asian components.

These results indicate that most of the ancient specimens analyzed in this study probably exhibited pigment features typical for modern Europeans and were thus genetically and phenotypically closer to modern Europeans than to Asians.

Discussion

To our knowledge, this study is the first successful report of multiplexed typing of autosomal SNPs on DNA retrieved from ancient skeletal remains. Due to the degraded nature of the DNA investigated, ensuring that the ancient samples did not become contaminated in any part of the analysis and monitoring possible contamination were of major concerns. We believe the possibility that our data came from exogenous DNA contamination to be unlikely for the following reasons. (1) All procedures for ancient samples preceding PCR cycling were performed in a dedicated aDNA work area where extensive precautions to prevent contamination were taken. (2) Extraction and PCR negative controls were always negative. (3) Genotypes were considered reliable when results were reproduced from at least two independent PCR reactions from the same extract and from two different extracts whenever possible. (4) Reproducible mtDNA HVI sequences, mitochondrial coding SNP genotypes, STR profiles, and Y-chromosomal SNP genotypes for male samples were previously obtained from these samples, indicating that DNA was probably well preserved. (5) STR profiles obtained from new extracts were always consistent with those obtained previously and different from those of laboratory personnel, indicating that aDNA samples were not contaminated by laboratory personnel and ruling out the possibility of cross-contamination between samples. (6) Finally, the findings of this study are consistent with other genetic data (i.e., Y-chromosomal and mitochondrial haplogroups) obtained previously from the same remains [20, 21]. The results of this study were further interpreted with respect to mitochondrial and Y-chromosomal data in an attempt to reconstruct the population history of south Siberian in an article currently under review for publication. Physical anthropological records also appear to support our data. No cranial records were available for the ancient samples analyzed in this study but studies of craniometric and non-metric traits have already been performed on ancient samples from the same period and geographical region [36, 37]. These studies revealed that Bronze Age and early Iron Age Siberian people clearly exhibited Caucasoid features rather than Mongoloid ones. Thus, the genetic data obtained from analysis of various independent systems, i.e., autosomal, Y-chromosome, and mtDNA, as well as physical anthropological records are in accordance, which suggests that the findings of this study make sense and probably do not arise from contamination.

The success of nuclear SNP genotyping was probably due to the genotyping system used. The multiplex PCR amplification followed by multiplex SBE reaction using the SNaPshot Kit seems to be ideal for the analysis of degraded and low content DNA for the following reasons: (1) PCR primers can be designed to generate product lengths less than 150 base pairs, enabling amplification of highly degraded DNA; (2) multiplex PCR amplifies different DNA fragments in a single PCR reaction tube reducing the amount of template needed; and (3) SNaPshot minisequencing detection is robust and efficient on both modern and ancient samples that contain different quantities and qualities of endogenous DNA. Moreover, this typing methodology only requires an automated DNA sequencer, which is an instrument that is already available in laboratories performing classical sequencing or STR typing.

This study clearly provides evidence that valuable information on pigment phenotypes and individual bio-geographical origin of an individual can be obtained by

analyzing the ten SNPs selected in this study. Nevertheless, we believe that accurate and reliable prediction of all three pigment characteristics (i.e., eye, hair, and skin color) is currently not achievable because most of genetic variants and even genes that contribute to normal human pigmentation variation remains unknown. The question of how many genetic markers are required to reliably estimate the individual ancestry of an individual has often been raised. Increasing the number of loci will undoubtedly improve inference of individual ancestry but an increase in the degree of multiplexing will also correlate with a decrease in amplification efficiency particularly when degraded DNA is used as template. Because the aim of this study was to apply the multiplexed assay on aDNA, we were limited in the number of SNPs selected. That is why only four markers were voluntarily selected among the many AIMs that have been described in the literature.

Autosomal markers are an important source of information for aDNA studies, which is currently under-exploited. On the one hand, in studies that attempt to determine the geographical origin of ancient populations, AIMs can complement data obtained from the widely used genetic sex systems, mtDNA and MSY, that reflect the maternal and paternal ancestries of an individual. On the other hand, SNPs located within nuclear genes can be very helpful to investigate a wide range of phenotypic traits, including pigment phenotypes as was exemplified in this study. Therefore, we believe that in the future the use of uniparentally inherited markers in aDNA studies will probably be complemented by autosomal data.

Besides molecular anthropology, the technical procedure described in this study may also find a forensic interest. Challenging samples that contain degraded and/or low amounts of DNA are also commonly encountered in forensic investigations notably in mass disaster victim identification (e.g., airplane crashes, terrorist attacks, or tsunamis where people from various geographical origins are generally involved and corpses are often severely damaged) or in identification of long deceased individuals (where only skeletal remains, teeth, or hair are available). Obtaining information on pigment traits and/or individual biogeographical ancestry may aid human identification in such cases where reference samples are not always available for comparison. The combination of physical anthropological measurements, which provide useful information on height and facial features and the typing of autosomal SNPs, which provides valuable information on pigment phenotypes and geographical origin, may enable a more precise description of physical appearance of the victim and potentially a more rapid identification.

In conclusion, to our knowledge, this work demonstrates for the first time that several autosomal SNPs can be simultaneously analyzed on degraded and aged DNA. Such investigation can provide valuable information on phenotypic features from the skeletal remains of an individual, which may be of interest to both molecular anthropologists and forensic geneticists.

Acknowledgments We would like to thank all volunteers who participated to this study by giving us their DNA. We thank more particularly Sophie Lienart, Sarah Romac, Prisca Blandin, and Aurélie Marchet for their valuable technical help and Marie Lacan for pertinent comments during the redaction of this paper.

References

1. Willerslev E, Cooper A (2005) Ancient DNA. Proc R Soc Lon 272:3–16
2. Pakendorf B, Stoneking M (2005) Mitochondrial DNA and human evolution. Annu Rev Genomics Hum Genet 6:165–183
3. Green RE, Krause J, Ptak SE et al (2006) Analysis of one million base pairs of Neanderthal DNA. Nature 444:330–336
4. Krause J, Lalueza-Fox C, Orlando L et al (2007) The derived FOXP2 variant of modern humans was shared with Neandertals. Curr Biol 17:1908–1912
5. Noonan JP, Hofreiter M, Smith D et al (2005) Genomic sequencing of Pleistocene cave bears. Science 309:597–599
6. Noonan JP, Coop G, Kudaravalli S et al (2006) Sequencing and analysis of Neanderthal genomic DNA. Science 314:1113–1118
7. Poinar HN, Schwarz C, Qi J, Shapiro B et al (2006) Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. Science 311:392–394
8. Amory S, Keyser C, Crubézy E, Ludes B (2007) STR typing of ancient DNA extracted from hair shafts of Siberian mummies. Forensic Sci Int 166:218–229
9. Haak W, Brandt G, Jong HN et al (2008) Ancient DNA, strontium isotopes, and osteological analyses shed light on social and kinship organization of the Later Stone Age. Proc Natl Acad Sci USA 105:18226–18231
10. Keyser-Tracqui C, Crubézy E, Ludes B (2003) Nuclear and mitochondrial DNA analysis of a 2, 000-year-old necropolis in the Egyin Gol Valley of Mongolia. Am J Hum Genet 73:247–260
11. Ricaut FX, Keyser-Tracqui C, Crubézy E, Ludes B (2005) STR-genotyping from human medieval tooth and bone samples. Forensic Sci Int 151:31–35
12. Schultes T, Hummel S, Herrmann B (1999) Amplification of Y-chromosomal STRs from ancient skeletal material. Hum Genet 104:164–166
13. Larcombe LA, Nickerson P, Hoppa RD, Matheson C (2005) Detection of a single nucleotide polymorphism in the IL-6 promoter region of ancient nuclear DNA. Infect Genet Evol 5:117–122
14. Burger J, Kirchner M, Bramanti B, Haak W, Thomas MG (2007) Absence of the lactase-persistence-associated allele in early Neolithic Europeans. Proc Natl Acad Sci USA 104:3736–3741
15. Lalueza-Fox C, Römpler H, Caramelli D et al (2007) A melanocortin 1 receptor allele suggests varying pigmentation among Neanderthals. Science 318:1453–1455
16. Parra EJ (2007) Human pigmentation variation: evolution, genetic basis, and implications for public health. Am J Phys Anthropol Suppl 45:85–105
17. Bennett DC, Lamoreux ML (2003) The color loci of mice—a genetic century. Pigment Cell Res 16:333–344
18. Parra EJ, Kittles RA, Shriver MD (2004) Implications of correlations between skin color and genetic ancestry for biomedical research. Nat Genet 36:S54–S60

19. Branicki W, Brudnik U, Wojas-Pelc A (2005) Genetic prediction of pigmentary traits in forensic studies. Prob Forensic Sci 64:343–357
20. Bouakaze C, Keyser C, Amory S, Crubézy E, Ludes B (2007) First successful assay of Y-SNP typing by SNaPshot minisequencing on ancient DNA. Int J Legal Med 121:493–499
21. Keyser C, Romac S, Bouakaze C, Amory S, Crubézy E, Ludes B (2008) Tracing back ancient south Siberian population history using mitochondrial and Y-chromosome SNPs. Forensic Sci Int Genet Suppl Ser 1:343–345
22. Keyser-Tracqui C, Ludes B (2005) Methods for the study of ancient DNA. Methods Mol Biol 297:253–264
23. Branicki W, Brudnik U, Kupiec T, Wolańska-Nowak P, Wojas-Pelc A (2007) Determination of phenotype associated SNPs in the MC1R gene. J Forensic Sci 52:349–354
24. Duffy DL, Montgomery GW, Chen W et al (2007) A three-single-nucleotide polymorphism haplotype in intron 1 of OCA2 explains most human eye-color variation. Am J Hum Genet 80:241–252
25. Eiberg H, Troelsen J, Nielsen M, Mikkelsen A, Mengel-From J, Kjaer KW, Hansen L (2008) Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the HERC2 gene inhibiting OCA2 expression. Hum Genet 123:177–187
26. Sturm RA, Duffy DL, Zhao ZZ et al (2008) A single SNP in an evolutionary conserved region within intron 86 of the HERC2 gene determines human blue-brown eye color. Am J Hum Genet 82:424–431
27. Graf J, Hodgson R, van Daal A (2005) Single nucleotide polymorphisms in the MATP gene are associated with normal human pigmentation variation. Hum Mutat 25:278–284
28. Lamason RL, Mohideen MA, Mest JR et al (2005) SLC24A5, a putative cation exchanger, affects pigmentation in zebra fish and humans. Science 310:1782–1786
29. Soejima M, Koda Y (2007) Population differences of two coding SNPs in pigmentation-related genes SLC24A5 and SLC45A2. Int J Legal Med 121:36–39
30. Myles S, Somel M, Tang K, Kelso J, Stoneking M (2007) Identifying genes underlying skin pigmentation differences among human populations. Hum Genet 120:613–621
31. Yuasa I, Umetsu K, Harihara S et al (2007) Distribution of two Asian-related coding SNPs in the MC1R and OCA2 genes. Biochem Genet 45:535–542
32. Römpler H, Dear PH, Krause J et al (2006) Multiplex amplification of ancient DNA. Nat Protoc 1:720–728
33. Sanchez JJ, Endicott P (2006) Developing multiplexed SNP assays with special reference to degraded DNA templates. Nat Protoc 1:1370–1378
34. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959
35. The International HapMap Consortium (2003) The International HapMap Project. Nature 426:789–796
36. Kozintsev AG, Gromov AV, Moiseyev VG (1999) Collateral relatives of American Indians among the Bronze Age populations of Siberia? Am J Phys Anthropol 108:193–204
37. Moiseyev VG (2006) Nonmetric traits in early Iron Age cranial series from western and southern Siberia. Archaeol Ethnol Anthropol Eurasia 25:145–152