



# Reconstructing the Origins and Migrations of Diasporic Populations: The Case of the European Gypsies

**ABSTRACT** The sociohistorical study of Gypsies/Sinti/Roma has been dominated by the “Gypsy-lorist” paradigm. According to this paradigm, these itinerant people belong to a unitary ethnic group of South Asian origin whose cultural practices have been preserved over the centuries. Recently, this perspective has come under criticism for perpetuating the image of Gypsies as an isolate within the wider context of the development of European societies, and, in particular, for placing too much importance on the external origin of Gypsy cultural and linguistic practices. This article attempts to place the available biological anthropological data for Roma origins and population history (from molecular genetic and clinical studies) in the wider ethnohistorical and linguistic context, and assesses their potential impact for an integrationist approach to Gypsy studies. These data suggest that, while the “Gypsy-lorist” paradigm is problematic, Gypsy populations share a common biological origin, a reality that should not be ignored. [Keywords: Diaspora, Gypsy, Roma, mtDNA, Y-chromosome]

## MOLECULAR ANTHROPOLOGICAL PERSPECTIVES ON DIASPORIC POPULATIONS

Diasporic populations occupy a special place within four-field anthropology because of the complex and multifaceted approaches that are required to study them. The maintenance of cultural, linguistic, and, to a lesser extent, biological identities (or lack thereof) is influenced both by internal mechanisms and by interactions with the so-called host populations. The complexity of the patterns that may emerge out of the diasporic condition also raises the very interesting problem of the covariance of genes, language, and culture, one that has recently come to the forefront of research in the field of anthropology. At the same time, the relative lack of an adequate comparative body of knowledge on the matter—four-fields research having been limited, until recently, largely to Jewish populations (e.g., Motulsky 1995; Thomas et al. 2000)—has stunted the growth of a potentially important area of academic inquiry. It is for this reason that we have attempted to synthesize the available literature about a rather different diasporic group, the Roma (Gypsies)<sup>1</sup> of Europe.

## THE “GYPSY-LORISTS” AND POPULATION BIOLOGY

The Roma/Sinti/Gypsies are an itinerant people without a written history of their own. As an almost inevitable conse-

quence of this fact, historical and anthropological accounts of them have come exclusively from scholars belonging to the dominant non-Gypsy majority. This has undoubtedly contributed to a series of research biases in the field that have only recently come under closer scrutiny and critique (e.g., Hancock 1987; Lucassen et al. 1998; Okely 1983). In particular, the synthetic work of the 19th-century German historian H. M. G. Grellmann (1753–1804), who was influenced by contemporary Herderian ideas, has had a lasting impact on a scholarly tradition that emphasized national character based on language and customs (Grellmann 1787). This tradition later became institutionalized with the founding of the *Journal of the Gypsy Lore Society* (JGLS) in 1888, which provided an outlet for a considerable number of scholarly papers on the Gypsies.

Although the JGLS itself has undergone several changes in theoretical and methodological orientation and has recently changed its name to *Romani Studies*, several authors have made the claim that the “Gypsy-lorist” paradigm has not been sufficiently eradicated from modern academic discourse on the Roma. Leo Lucassen et al. (1998) state that the scholarly construction of Gypsies as a unitary isolated ethnic group with a common origin and inherited South Asian cultural practices contributed to the centuries of labeling and stigmatizing that ultimately led to the “criminal-biological” portrait that accompanied the

systematic killing of Gypsies in National Socialist Germany. Furthermore, much like Edward Said's Orientalist body of scholarship (1978), the "Gypsy-lorist" work is presented by Wim Willems (1998) as part of a trend in Western learning to exert European authority over a constructed "Other." These scholars urge the academic community to move away from such notions and toward a more integrated study of Gypsies in the context of developing European societies since the Middle Ages.

While finding this goal more than admirable, we believe that abandoning the concept of a certain unity in our object of study, at least with respect to the genetic component, would be a mistake. In this article, we will review the available genetic evidence for the origin and spread of the Gypsies in the context of the linguistic and historical evidence, as well as explore its implications for the nature of various interactions with the local populations. As will be shown below, these data indicate that Gypsy populations share a common biological origin, while also having undergone various degrees of biological and cultural assimilation with non-Gypsy groups in different parts of Europe.

#### CONTEMPORARY EUROPEAN DEMOGRAPHICS: THE SAMPLING PROBLEM

Although obtaining representative biological samples for molecular studies of other diasporic populations (e.g., Jews) is sometimes difficult, the obstacles involved in working with Gypsies are much greater. These problems have arisen in part because of their marginalized place in traditional European and Asian history and their relative lack of political organization (until relatively recently). Unfortunately, there are no reliable statistical and demographic data for the distribution or internal subdivisions of Gypsy societies in Europe. Many Roma individuals in Central and Eastern Europe simply do not want to identify themselves as such and, thus, claim non-Roma identities. Others are of Roma descent but genuinely identify themselves with the majority group. Still others simply cannot understand the survey questions. For these reasons, most authors (e.g., Guy 2001; Liégeois 1994) agree that the number of Gypsies in each country should be counted as about three times that reported in the official census. Even so, these estimations may run low: The Czech Gypsy population is estimated at about 300,000 people, despite the fact that only 32,903 declared themselves as Roma; in Slovakia, there may be 500,000 rather than the official 83,988; and, in Romania, 1.5–2.5 million rather than the 409,723 (Guy 2001 and references therein). In Bulgaria, although only 313,396 people were designated by the census as Roma, the Ministry of the Interior approximated a population size of 700,000–800,000 and Roma leaders claimed more than one million.

Aside from the obvious problem involving the unknown numbers of the Roma, the issues surrounding the complexity of the various ethnonyms and language groups still plagues much of the research on Gypsy populations. Some names such as Kalderaš or Zlatara or Kazandži are

derived from various occupations. Others are derived from names given them by *gadžé* (non-Gypsies), such as Pletoși (Romanian [hereafter Rom.] *pletoși*, "long-haired"), or by other Gypsy groups (e.g., *Romungri*, a term meaning "Hungarian Rom," mistakenly applied by Lovara to Polish and Russian Gypsies [Matras 2002]). These self-designations are often confusing because Gypsy groups probably tended to appropriate *gadžé* terms into their group names (e.g., Polska Roma, Polish Roma) for convenience, social strategy for acceptance, or as a sign of resistance.

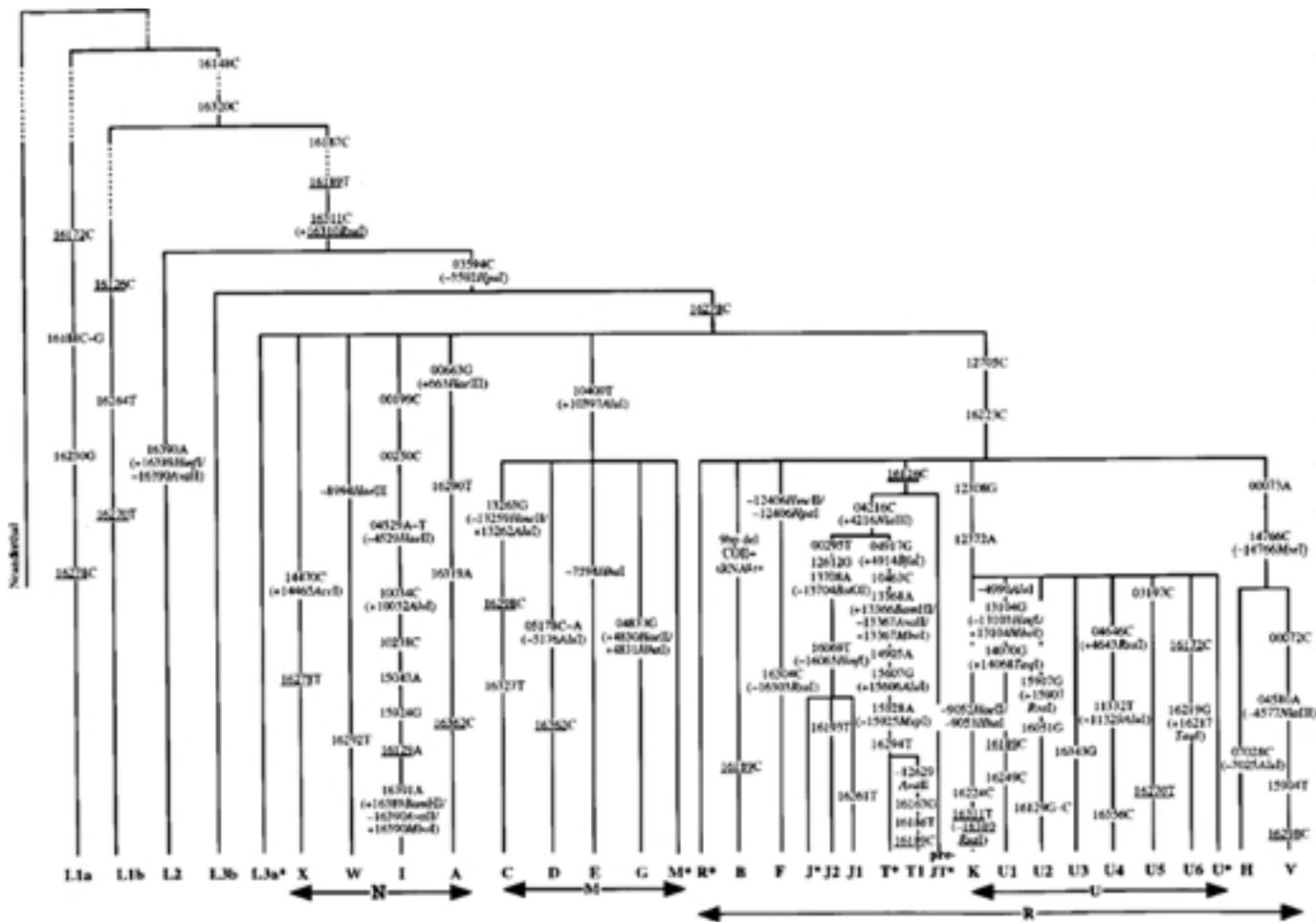
Moreover, Western European Romani-speaking groups such as the Sinti and Manuš, as well as non-Romani-speaking ones such as the Spanish Gitanos, often dissociate themselves from the more recently arrived Roma from the east (Gay y Blasco 1999, 2002; Williams 1984). Judith Okely (1983), Jean-Pierre Liégeois (1994), and others have pointed out the problems associated with analyzing contemporary Roma groups in terms of classical ethnic markers (social group, language, "culture"). In various forms, the arguments presented aim to demonstrate that Gypsies cannot be thought of as a global diaspora, and, instead, that "Gypsy-ness" is a way of life more than an ethnic group defined through foreign origins in the "genetic" sense.

At the same time, as indicated by the available genetic evidence to be described below, Gypsies are not entirely culturally defined. Their history reveals them to be founder populations that have undergone limited gene flow with other Europeans, mostly due to their endogamous marriage practices. While this endogamy is cultural in nature (depending, among other factors on class and position in the social space), it can effectively act as a biological barrier. This is the major reason that the Gypsies pose an interesting problem for population biological studies, which have mostly dealt with culturally more homogeneous and self-conscious entities, such as the Jews.

#### THE GENETIC DATA ON GYPSY ORIGINS AND MIGRATIONS

The history of genetic research on the Roma begins with the National Socialist German programs supported by the Deutsche Forschungsgemeinschaft and the Race and Population Biology Research Centre. These programs evidently attracted eminent biologists, such as the Nobel laureate Ferdinand Sauerbruch, who signed the first grant proposal for "genetic and medical research" to be carried out at the concentration camp in Auschwitz (Kalaydjieva et al. 2001b). By contrast, postwar genetic research was mainly carried out to answer questions relating to the purported Indian origin of Roma populations, the extent of admixture with and relatedness to European populations, and their intragroup diversity.

In this article, we have divided the genetic studies on Gypsies into two categories. The first deals with novel single gene disorders, disease-causing mutations, and "classical" polymorphic systems (e.g., blood groups such as ABO and Rhesus). The second group specifically investigates Gypsy population history through single-locus marker



**FIGURE 1.** A mtDNA phylogeny showing the major maternal haplogroups that are present in modern human populations (from Macaulay et al., *American Journal of Human Genetics* 64, 232–249. ©1998 by the American Society of Human Genetics. All rights reserved).

systems (nonrecombinant Y-chromosome [NRY] and mitochondrial DNA [mtDNA]). We discuss these studies in the context of previous linguistic and historical data to evaluate a series of hypotheses about Gypsy origins and later expansions.

### **Population and Medical Genetic Studies**

In a recent study, Luba Kalaydjieva et al. (2001b) performed a multilocus comparison between Romani populations from different European countries (Slovakia, Hungary, England, Slovenia, Sweden, and Wales), autochthonous European populations, and populations from North India (Rajputs and Punjabis), using polymorphic blood group systems such as A<sub>1</sub>A<sub>2</sub>BO, MN, haptoglobin, and Rhesus (Rh). They concluded that, in general, most Roma were genetically closer to Indian than to European populations, which was the expected result given the linguistic phylogeny of Romani dialects (see discussion below). In addition, an analysis of molecular variance (AMOVA) showed that the Roma populations exhibited a higher degree of internal (between-group) diversity than did the majority of surrounding European populations (Kalaydjieva et al. 2001b).

Studies over the last decade have also resulted in the identification of a number of novel single gene disorders and disease-causing mutations in Gypsies. To date, nine Mendelian disorders caused by "private" Roma mutations have been described (table 2 and figure 4 in Kalaydjieva et al. 2001b). Originally found in Bulgarian populations, they were later identified in Spanish, Portuguese, French, and German individuals. These Mendelian disorders have been described as genetically homogeneous, with a single mutation accounting for all affected individuals and related polymorphic haplotypes, indicating a common origin and founder effect (Kalaydjieva et al. 2001b). Among these novel neurological disorders are two types of hereditary motor and sensory neuropathy in the Lom (Chandler et al. 2000; Colomer et al. 2000) and the Russe (Thomas et al. 2001), and the congenital cataracts facial dysmorphism neuropathy syndrome (Tournev et al. 1999). Other examples include limb-girdle muscular dystrophy type 2C, galactokinase deficiency, primary congenital glaucoma, and congenital myasthenia. These biomedical data imply that Gypsies have experienced some degree of isolation from other ethnic groups following their establishment as founding populations in various regions of Europe. Otherwise, one would not expect to observe so many private disease-causing mutations in these groups. Furthermore, similar patterns of genetic diseases have been seen in other populations with a history of isolation and endogamy (Heyer 1999; Motulsky 1995; Peltonen et al. 1999).

### **Molecular Genetic Studies**

The unique genetic heritage of the Gypsies shown by these disorders and classical blood group marker data stimulated additional research into Romani origins using the mtDNA and Y-chromosome (Gresham et al. 2001; Kalaydjieva et al. 2001a). Unfortunately, the work on nonrecombinant

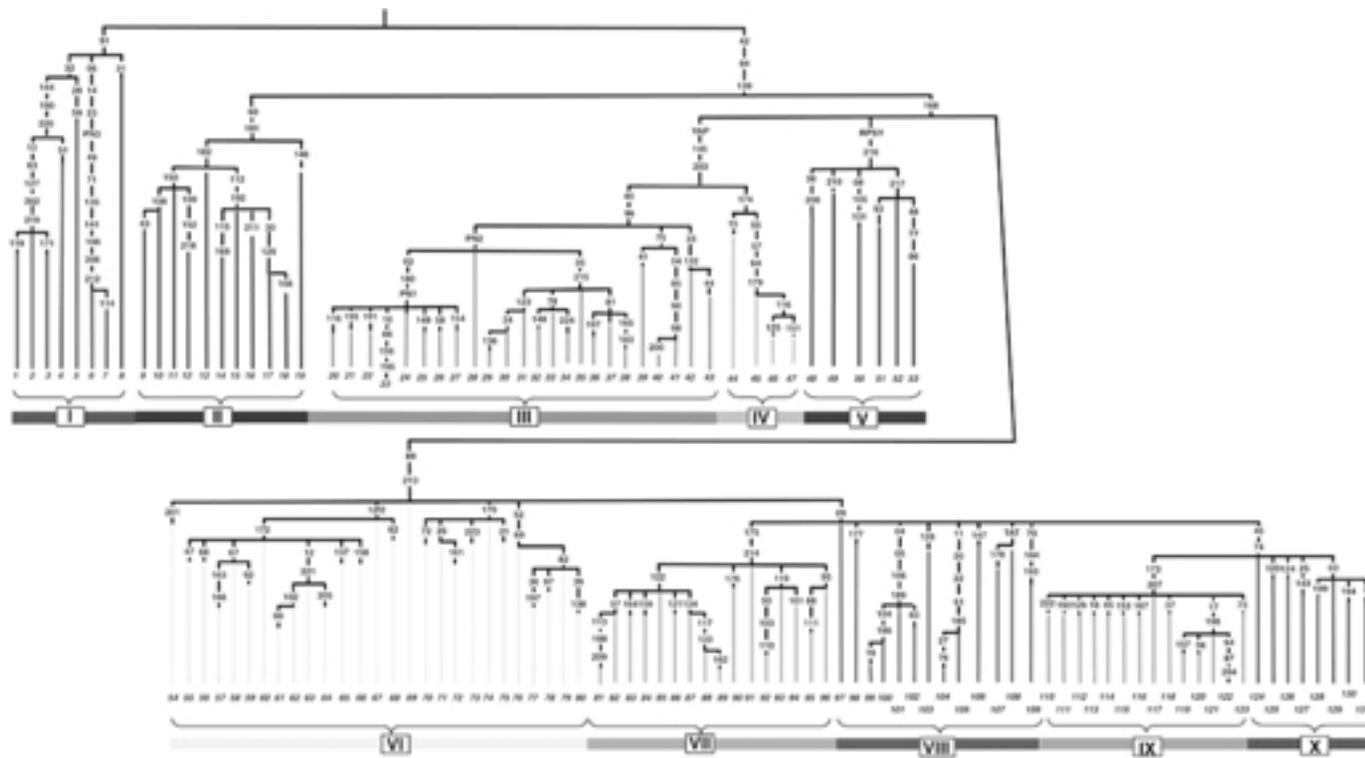
DNA marker systems like the mtDNA and Y-chromosome amounts to only two studies. For this reason, and because of the small sample sizes and geographic biases of the sampling, all results summarized here must be regarded as somewhat preliminary, despite their provocative conclusions.

Before describing the patterns of genetic variation in the Roma, we will first define the units of analysis in these molecular studies. In genetics parlance, a *haplotype* is the combination of genetic markers or polymorphisms that are present in a gene, gene segment or genome, such as the mtDNA, and that are inherited together as a unit. A *haplogroup* is a set of haplotypes that share a specific set of polymorphisms (genetic markers), which, in turn, reflects their genetic relatedness. In general, we use these terms to talk about different portions of a genetic genealogy, or phylogeny—specifically, either the large branches (haplogroups) that have more ancient origins or the smaller branches (haplotypes) which reflect the process of genetic diversification within these haplogroups over time.

We can identify both haplotypes and haplogroups through several different methods. One method that has been extensively used with the mtDNA is RFLP analysis. RFLP analysis utilizes the ability of restriction enzymes to cut DNA at specific locations (combinations of nucleotide bases called *recognition sequences*) to detect variation in DNA sequences, as point mutations will either eliminate existing sites or create new ones. The presence of new mutations is determined by comparing the sample RFLPs with those of a standard reference DNA, such as the published human sequence (Anderson et al. 1981). These RFLPs are used to construct the general architecture of mitochondrial phylogenetic trees, such that shown in Figure 1, which shows all of the main branches identified in modern human populations.

The mtDNA genome also has a noncoding control region (CR) in which its bidirectional replication is initiated. The CR mutates more rapidly than the coding sequences in the rest of the mtDNA genome, perhaps two-to-ten times faster, with two hypervariable regions (HVR-I and HVR-II) being the most rapidly evolving portions of it (Horai et al. 1995; Ingman et al. 2000). Because of this feature, one can obtain a very detailed assessment of mutational changes that have occurred within a single mtDNA, as well as delineate the finer branches of a mitochondrial phylogeny, using direct sequencing.

In contrast, single-nucleotide polymorphisms (SNPs) are used for NRY haplogroup typing and the more rapidly mutating Y-chromosome short tandem repeats (STRs) used to define individual haplotypes. SNPs are unique event substitutions (i.e., point mutations, insertions, or deletions) that occur in the nonrecombining portion of the Y-chromosome. Their ancestral states are easily inferred from hominoid ape Y-chromosome sequences. These biallelic (presence/absence) markers define the major branches of the human NRY phylogeny and also preserve population-specific haplotype information over the span of human history (Figure 2).



**FIGURE 2.** A Y-chromosome SNP phylogeny showing the major paternal haplogroups that are present in modern human populations (from Underhill et al., *Annals of Human Genetics* 65, 43±–62. ©1998 by The Annals of Human Genetics. All rights reserved).

An STR, or microsatellite, is a genetic marker composed of multiple copies of a short nucleotide sequence (tandem repeats) such as GATA at a single locus. Allele size differences for individual STRs are based on the number of tandem repeats at a particular locus (e.g., 6, 7, or 8). Haplotypes from the same SNP haplogroup may have several different alleles for a particular STR locus because the higher mutation rate of the STR produces a large number of alleles that distinguish otherwise identical Y-chromosomes from each other. For this reason, several to many different STR loci are usually analyzed for every Y-chromosome, as it is possible to observe in these loci the directionality of allele size changes occurring within them, hence, how they have evolved (expanded or contracted in allele size) within a paternal lineage or haplogroup. This cumulative information gives a clearer picture of the genealogical relationships of the NRY haplotypes in which the STR markers occur and allows the construction of the finer branches of the phylogenetic trees.

Having described the biological units of analysis, we will now turn our attention to the genetic data from Gypsy populations. The most extensive study explored the genetic structure of 14 well-defined Roma populations, using both NRY and mtDNA markers (Gresham et al. 2001). Almost all of the populations sampled were from Bulgaria, with the remaining two coming from outside the Balkans—one from Madrid and the other from Vilnius, Lithuania. The Bulgarian samples represented nine different ethnonyms (Turgovtzi, Feredjelli, Kalaidjii, Koshnichari, Lom, Monteni, Intreni, Lingurari, and Kalderash) and five linguistic entities (Balkan Romanes, Turkish, Old Vlax Romanes, Macedo-Romanian [Aroumanian], and New Vlax Romanes). They could further be historically divided into three major mi-

gration events: namely, (1) the early settlement of Bulgaria; (2) the 17th- and 18th-century migration from Moldo-Wallachia; and (3) the “great Kelderara” invasion of the 19th century.

The NRY diversity was expressed mostly in three haplogroups, namely VI-68, VI-52, and VI-56, which represented approximately 80 percent of all Roma Y-chromosomes (Table 1; Figure 2). Within these haplogroups, four haplotypes (VI-68A, VI-68B, VI-52A, and VI-56A) accounted for 57 percent of Roma Y-chromosomes. VI-68 was by far the most common haplogroup, as it was found in all 14 populations and was considered a good candidate for the founding paternal lineage of the Roma (Gresham et al. 2001). When compared with non-Romani haplogroup VI-68 Y-chromosomes, those of the Roma represented a limited subset of closely related haplotypes within the overall diversity of this haplogroup in Asia. The most common paternal haplotype, VI-68A, was shared by 80 individuals and accounted for 71 percent of haplogroup VI-68 chromosomes, along with 32 percent of all Roma chromosomes. Moreover, it was one mutational step away from the second most common lineage, VI-68B, which occurred mainly in the Lom and the Lithuanian Roma. Assuming that VI-68A was its founding haplotype, a coalescence time of 992 years was estimated for haplogroup VI-68 (Gresham et al. 2001).

Several other haplogroups occurring in the Roma also seemed to have an Asian origin. For example, haplogroup VI-56 is seen mostly in Central Asia and the Middle East, but in Europe appears in a single individual from Sardinia (Underhill et al. 2000). In addition, haplogroups VI-57 and IX-108 occurred mainly or exclusively in Asian populations. On the other hand, haplogroups VI-52 and IX-104 (Eu7

**TABLE 1.** Y-Chromosome SNP Haplotype Frequencies in World Populations

Population	n	NRY SNP Haplotypes								
		VI-68	VI-52	VI-56	IX-104	VI-71	VI-57	V-52	IX-108	III-36
Roma	252	44.9	22.2	12.7	6.7	3.6	2.0	1.6	0.4	3.6
Sudan	40					10.0				
Ethiopia	88				1.1	4.5				6.8
Mali	44				2.3					
Morocco	28				17.9	50.0				
Central Africa	37									
Khoisan	39									10.3
South Africa	53									1.9
Europe	60		13.3		48.3	15.0		13.3	5.0	
Sardinia	22			1.7		9.1				
Basques	45		2.2		57.8	2.2		2.2		
Middle East	24	4.2		8.3	8.3	33.3				
Central Asia/Siberia	184	2.2		1.1	5.4	9.2	0.5		16.3	
Pakistan/India	88	2.3			6.8	8.0	4.5		31.8	
Hunza	38				7.9	2.6			28.9	
Japan	23									
China	20									
Taiwan	74									
Cambodia/Laos	18								5.5	
New Guinea	23									
Australia	7				14.3					
Americas	108				4.6	0.9		0.9		

Note: A total of 2% of the Roma Y-chromosomes could not be assigned to a known SNP haplogroup. The data were taken from Gresham et al. (2001) and Underhill et al. (2001).

**TABLE 2.** mtDNA Haplogroup Frequencies in Indian and Roma Populations

Haplogroup	Dravidic	Hindi	Roma
<i>EAST EURASIAN</i>	65.7	55.7	26.5
A	—	0.3	—
B	—	—	—
F	—	0.7	—
M-Total	65.7	52.7	26.5
M2	?	9.9*	—
M3	6.6	6.0 (10.5*)	—
M4	?	5.5*	—
M5	?	13.4*	25.8
M-C	—	0.7	—
M-D	—	1.0	—
M-E	0.8	—	—
M-G	0.4	—	—
<i>WEST EURASIAN</i>	14.5	27.4	69.8
H	1.2	2.3	35.6
I	—	1.3	1.8
J	0.4	0.7	5.5
K	0.4	—	1.5
N1b	?	?	1.8
T	1.7	1.7	2.2
U1	?	2.3 <sup>ˆ</sup>	0.4
U2i	10.3	15.3	—
U3	?	— <sup>ˆ</sup>	10.2
U4	?	4.7 <sup>ˆ</sup>	—
U5	?	1.2 <sup>ˆ</sup>	2.2
U7	?	12.7 <sup>ˆ</sup>	—
V	—	—	—
W	0.4	3.7	1.1
X	—	0.7	7.6
<i>OTHER</i>	19.8	16.7	—

*Note:* Data taken from Kivisild et al. (1999), Bamshad et al. (2001), and Gresham et al. (2001). The proportion of haplotypes that belong to East Eurasian [Asian], West Eurasian, and Other categories are specified by the subheaders in italics. The Other category may contain stem haplotypes from the N and R nodes of the human mtDNA phylogeny, such as N1b in the Roma, or those belonging to haplogroup U and its subbranches. The asterisk (\*) indicates approximate subhaplogroup frequencies taken from Figure 2 in Bamshad et al. (2001). These were placed in the Hindi column because it was not possible to ascertain their exact frequencies in the Dravidic and the Hindi populations. The carat (ˆ) indicates subhaplogroup frequencies for haplogroup U as presented in Kivisild et al. (1999); these frequencies were not specified for either the Dravidic or Hindi populations in Bamshad et al. (2001).

and Eu18 in Semino et al. 2000) are common in Europe, although IX-104 also appears in Central Asia, Pakistan, and India (Underhill et al. 2000). The presence of these haplogroups in the Roma may reflect a certain amount of European admixture, although it is also possible that IX-104 was part of the ancestral Gypsy gene pool.

The remaining haplogroups exhibited different patterns of distribution. Unlike the other haplogroups, IV-71 appeared in populations from across the world and showed no specific geographic association. By contrast, haplogroup III-36 had previously been seen in only Ethiopia and South Africa (Underhill et al. 2000, 2001). Among the Roma, this haplogroup occurred at a low frequency but appeared only in Bulgarian populations.

From a mtDNA standpoint, the major haplogroups identified in the Roma were M and H (Tables 2 and 3; Figure 1). Haplogroup M is common in Asia, rare in Europe, and also appears in eastern Africa, where it may have originated (Quintana-Murci et al. 1999). However, the HVR-I sequences from Roma haplogroup M were missing the African motif, indicating that they had an Asian or Indian origin (Gresham et al. 2001). In fact, subhaplogroup M5 was observed in 11 of 14 populations, and accounted for 97

percent of all Roma haplogroup M mtDNAs. Phylogenetic analysis of Roma and Indian haplogroup M sequences from Toomas Kivisild et al. (1999) showed that all but two of the Gypsy sequences clustered together as a small subset of the diverse types present within Indian populations. At the same time, the Roma lacked mtDNAs from subhaplogroups (M2-M4) that typically occur in Indian (Hindu) populations (Bamshad et al. 2001) (Table 2). Taken together, the coalescence time of Roma M HVR-I sequences was estimated to be 4,625 years before present (ybp).

Haplogroup H was the most frequent mtDNA lineage among the Roma, detected in 13 out of the 14 sampled populations. This haplogroup is also the most common mtDNA lineage in Europe and the Near East (Richards et al. 2000) but is also observed in India (Kivisild et al. 1999; see also Tables 2 and 3). Even so, two of the Roma HVR-I sequences from this haplogroup have not yet been detected in Near Eastern and European individuals (Richards et al. 2000).

Interestingly, haplogroup X was also present in 7.6 percent of the Roma samples. In fact, three out of five of its sequences had not been identified in populations from Europe or the Near East, where this haplogroup is

**TABLE 3.** Haplogroup Frequencies Based on HVR-I Data [from Simoni et al. 2000]

POPULATION	FREQUENCY OF HAPLOTYPE											
	H	I	J	K	T	U3	U4	U5	V	W	X	Oth*
ROMA	.356	.018	.055	.015	.022	.102	.000	.022	.000	.011	.076	.287
Albania	.524	.071	.009	.000	.000	.000	.000	.143	.000	.000	.000	.119
Austria	.325	.034	.000	.103	.085	.009	.043	.068	.009	.009	.009	.188
Belgium	.406	.000	.030	.125	.031	.000	.063	.031	.094	.000	.000	.156
England:												
Cornwall:	.348	.058	.000	.029	.087	.000	.058	.043	.014	.000	.000	.145
Mainland:	.350	.030	.011	.100	.070	.000	.040	.080	.030	.000	.030	.140
Wales:	.478	.033	.000	.076	.043	.000	.000	.043	.033	.000	.011	.130
Bulgaria:	.233	.000	.067	.133	.100	.100	.067	.033	.000	.000	.067	.167
Caucasus:	.220	.060	.000	.020	.140	.140	.040	.080	.000	.000	.000	.260
Denmark:	.344	.000	.000	.031	.063	.000	.000	.063	.031	.031	.000	.250
Estonia:	.214	.000	.000	.000	.107	.000	.071	.179	.000	.071	.000	.250
Finland:	.278	.063	.044	.051	.051	.000	.025	.139	.089	.076	.000	.127
France:	.450	.018	.009	.045	.099	.000	.000	.054	.027	.036	.000	.180
Georgia:	.178	.022	.000	.111	.244	.044	.111	.067	.000	.022	.044	.089
Germany:												
Northern	.250	.009	.178	.093	.083	.019	.009	.074	.065	.009	.009	.269
Southern	.309	.028	.000	.052	.088	.024	.052	.092	.048	.016	.004	.177
Iceland:	.226	.019	.014	.038	.075	.057	.019	.113	.019	.000	.000	.245
Israel/Druze:	.244	.044	.000	.156	.044	.000	.000	.000	.000	.000	.178	.178
Italy:												
Alps:	.184	.026	.032	.088	.219	.018	.035	.061	.061	.000	.000	.193
Sardinia	.384	.027	.067	.055	.110	.000	.000	.082	.027	.014	.014	.219
Sicily	.397	.016	.027	.048	.016	.032	.016	.000	.079	.048	.032	.190
Southern	.378	.027	.041	.027	.108	.000	.000	.081	.027	.054	.027	.108
Tuscany	.224	.041	.000	.061	.061	.000	.041	.061	.000	.020	.041	.224
Karelia:	.313	.024	.000	.024	.072	.000	.084	.181	.060	.036	.000	.120
Kurdistan:	.276	.069	.095	.172	.034	.069	.000	.000	.000	.034	.000	.276
Near East:	.095	.071	.000	.000	.119	.024	.000	.000	.000	.000	.095	.190
Norway:	.467	.067	.019	.133	.000	.000	.000	.100	.033	.033	.000	.133
Portugal:	.444	.000	.000	.074	.111	.019	.056	.000	.037	.000	.019	.130
Saami:	.033	.025	.019	.000	.004	.000	.000	.529	.346	.000	.000	.058
Spain:												
Basques	.509	.000	.000	.047	.047	.000	.000	.104	.066	.000	.019	.179
Catalunya	.267	.000	.000	.000	.000	.000	.133	.000	.267	.133	.067	.067
Central	.522	.041	.004	.033	.022	.000	.011	.043	.033	.022	.011	.185
Galicia	.243	.011	.000	.054	.095	.014	.027	.081	.054	.041	.027	.203
Sweden:	.250	.000	.027	.031	.125	.000	.094	.063	.063	.000	.000	.281
Switzerland:	.278	.014	.000	.069	.028	.000	.056	.083	.056	.014	.000	.292
Turkey:	.221	.032	.000	.032	.074	.053	.032	.021	.032	.042	.032	.168
Volga-Finnic:	.176	.029	.032	.029	.118	.000	.147	.118	.029	.000	.000	.176

Note: Data taken from Simoni et al. (2000) and Gresham et al. (2001). The "Other" category of haplotypes contains the M, N1b, and U1 frequencies for the Roma. For European populations, this category includes any M and additional East Eurasian haplogroups present in them, which are rare, as well as the HVR-I sequences that could not easily be assigned to known haplogroups. If analyzed for informative coding region RFLPs, then these sequences would likely be grouped into H, HV, N, R and U haplogroups. See Richards et al. (1996, 1998, 2000), Torroni et al. (1996), Macaulay et al. (1999), Helgason et al. (2000), and references therein for additional estimates of haplogroup frequencies in European populations.

widely distributed, albeit at low frequencies (approximately two percent) (Richards et al. 1998; Richards et al. 2000; Simoni et al. 2000; Torroni et al. 1996; see also Tables 2 and 3). These findings suggest that this haplogroup was part of the Indian founding maternal lineages (Kivisild et al. 1999).

In terms of the genetic structure of Roma populations, the 13 paternal and 25 maternal haplotypes were differentially distributed amongst these groups. Principal components analyses (PCA) using both Y-chromosome and mtDNA-haplogroup frequencies consistently identified two clusters: The first was formed by the Monteni, Intreni, Lingurari (Romanian-speakers), Kalderaš (New Vlax speakers), and Lom (Old Vlax speakers), and the second by

the Feredjelli and Turgovtzi (Turkish, Balkan Romanes) (Gresham et al. 2001). In addition, the Spanish and Lithuanian Roma clustered together in the mtDNA PCA analysis, whereas the Kalaidjii North and South groups clustered together in the Y-chromosome PCA analysis (Gresham et al. 2001).

The Y-chromosome STR and mtDNA HVR-I data were also subjected to AMOVA to determine whether variation at these loci was structured by country of residence, place of residence, or trade-based ethnonym. No significant intergroup differences were observed for either set of markers except when the history of Gypsy migrations was taken into consideration. When haplotypic variation was partitioned by language, there was a significant intergroup difference

only with respect to mtDNA (Gresham et al. 2001). This finding could possibly reflect the strict endogamy rules practiced by the Romanian-speaking Roma males. More significantly, ten percent of the variance for the Y-chromosome and five percent for the mtDNA was attributable to the three major migrations of Roma populations that have taken place over the past several hundred years (see above).

These results were concordant with those from a much more localized study on the Vlax Roma in Bulgaria (Kalaydjieva et al. 2001a). In this study, both Y-chromosome and mtDNA diversity suggested the recent common origin for the three groups being studied (Kalderaš, Monteni, and Lom, who speak New Vlax Romanes, Aroumanian, and Old Vlax Romanes, respectively). They also highlighted the limited genetic diversity of this young founder lineage, whose historical migration event occurred in the 19th to early 20th century from Moldo-Wallachia.

## GYPSY POPULATION HISTORY IN LIGHT OF GENETIC DATA

### *Indian Origins*

The origin of Gypsy populations in the Indian subcontinent has been proposed under various hypotheses stemming from historical linguistics (e.g., Fraser 1992b; Hancock 2000; Miklosich 1872; Sampson 1926; Tálos 1999; Turner 1926). A comprehensive review of the history of linguistic theories on the origin of Romani can be found in Yaron Matras's (2002) recent book *Romani: A Linguistic Introduction*. While nearly all scholars agree on the idea that the Romani lexicon originated within the Indo-Aryan family, there is quite a divergence of opinion when it comes to identifying the social context of the first migrations of proto-Romani speakers. Various explanations have been proposed in relation to the terms *Rom*, *Lom*, and *Dom*, which are frequently seen as a Gypsy ethnonyms, especially because *dom* refers to a caste-type affiliation that, in several Indian languages, describes itinerant groups that specialize in service-providing trades, including crafts and entertainment. Although this hypothesis is appealing because of its social implications (also indicating at least a common social, if not ethnic and linguistic, origin), this hypothesis is very difficult to test and remains speculative.

Nevertheless, several activist scholars have recently revived 19th-century hypotheses regarding the polyphyletic emergence of Romani in the Central Asian subcontinental area (e.g., Hancock 2000; Kochanowski 1990). In an attempt to justify a warrior image for Gypsy ancestors, Ian Hancock (2000) has suggested a link between the defeat of the Rajputs by the Muslim Ghaznavids in the 12th century. This idea is difficult to reconcile with historical data that places Gypsies in the Byzantine empire before this time (around the 10th–11th centuries; see Soulis 1961), as well as with linguistic evidence of reasonably long-term contact with Persian and Armenian before this (Matras 2002). This evidence also points to a southern Black Sea route through Anatolia

and into Greece and Europe (Matras 1996; 2002:25), which would explain the virtual absence of Arabic (and Turkish) and the presence of a few items of Georgian and Ossetian origin in Romani. At the same time, a more northerly route through the Caucasus and Crimea to Eastern Europe is unlikely, given that the first evidence of Gypsies in Russia appears in the 16th century, probably coming from the Romanian Principalities.

The Indian Origin hypothesis seems to be supported by the classical marker, Y-chromosome and mtDNA data. Multilocus comparison of classical genetic markers in Roma groups and other world populations show most of them to have strong affinities with Rajput and Punjabi populations.<sup>2</sup> The majority of Roma NRY haplogroups also appear in Central Asian, South Asian, and Middle Eastern populations, suggesting they could have been part of the ancestral Gypsy gene pool. In addition, Roma haplogroup M mtDNAs are essentially a subset of Indian haplotypes, although the Roma lack subhaplogroups M2–M4 that are commonly seen in Indian populations. Thus, the genetic data suggest that Gypsies trace much of their ancestry to an Indian/Central Asian source population(s) and underwent a population bottleneck that led to the expansion of certain haplogroups in descendant Roma groups.

### *Gypsy History/Diaspora in Europe*

Based on these limited but informative genetic data, it appears that the cultural and linguistic diversity of contemporary (European) Gypsy groups masks a common genetic background (shared founder Y-chromosome haplotype VI-68A; closely related M haplotypes). The genetic information also reveals some common patterns of population history for the Roma, such as the bottleneck → diversification → migration [bottleneck] sequence. However, the larger amount of genetic diversity present in Gypsy groups compared to Europeans requires further explanation. It could reflect their origins in a genetically heterogeneous yet closely related set of subpopulations (such as might be suggested by an origin based on occupations and lifeways rather than the traditional ethnic groups). Alternately, it could have resulted from genetic drift and founder effects in subdivided small populations subsequent to a significant bottleneck at the founding of the lineages (as attested to by the presence of many unique NRY and mtDNA haplotypes). Clearly, the nature and frequency of unusual genetic diseases in Gypsies, as well as the appearance of unique haplotypes in these groups, indicates that they have experienced considerable endogamy and isolation over the past several centuries.

Of course, admixture with non-Gypsies has occurred in Roma groups. This is suggested by the presence of mtDNA haplogroups H, T, and U5 in Roma populations, and by frequencies of NRY haplogroups VI-52 and IX-104 similar to those seen in European populations. However, if admixture was extensive, one would expect Roma groups to have much higher frequencies of West Eurasian mtDNA (H-K,

T-X; Macaulay et al. 1999; Richards et al. 2000, 2002; Torroni et al. 1996) or NRY (VI, IX; Underhill et al. 2001) haplogroups that are ubiquitous in European populations, as well as share specific haplotypes from those haplogroups with them. Thus, only a very detailed comparative genetic study of "host" populations in countries where admixture is suspected to have taken large proportions can determine its exact nature. The publication of comprehensive reviews of mtDNA and NRY haplotype diversity in Europe involving very large sample sizes will also aid in this effort (e.g., Richards et al. 2002; Rosser et al. 2000).

Additional insights into Gypsy population structure have also been gained through genetic studies. Interestingly, it is more language than geography that accounts for the patterns of mtDNA and Y-chromosome variation in Roma populations (Gresham et al. 2001). This observation implies that Roma cultural and linguistic traditions have helped to shape genetic variation in the Roma. These results contrast with the patterns of NRY diversity in Europe as a whole, where geography more than language seems to account for the patterns (Rosser et al. 2000). Furthermore, the difference in the levels of diversity in the two different markers (with NRY sequences being more diverse than those of the mtDNA) was observed on a wider European scale as well and may have resulted for some of the same reasons, such as patrilocality leading to differential sex-mediated gene flow (in which females travel more than males to find mates and take up residence). However, when comparing genetic and linguistic diversity in Roma groups, only mtDNA haplogroup frequencies were significantly related to dialectic/language differences. Such results again point to the effects of endogamy in the Roma and suggest that Gypsy women may have played an important role in the transmission of cultural traditions in Roma populations through time.

### **Slavery, Admixture, and Assimilation**

Having reviewed the different processes that can be used to reconstruct Gypsy population history from its genetic structure, we will now attempt to analyze these data within a historical-linguistic context. The phenomenon of slavery in the autonomous Danubian principalities of Wallachia and Moldavia is of particular interest to Gypsy population history. In this context, both the laws restricting contact between Romanians and Gypsies, and the possibly reactive Gypsy customs regarding this issue, probably had a great impact on the genetic structure of these populations. Although, initially, princes donated Gypsy slaves to monasteries, by the 16th century, the slave system had become more complex. At this time, it included three types of Gypsies: (1) the Vătrași (Rom. *vatră*, "hearth" or "home") or Țigani boierești, who were directly enslaved by the nobility (Rom. *boieri*, "boyars"); (2) the Țigani mănăstirești, who belonged to monasteries and were under independent ecumenical control; and (3) the Lăieși, or "free" Gypsies, who remained itinerant but paid tribute and were considered Țigani

domnești (Rom. *domn*, "lord" or "prince"), or Crown Gypsies, who technically belonged to the state (Kogalniceanu 1837; Panaitescu 1941). This firmly established system underwent negligible changes until the 1830s, and it was not until 1856, after about twenty years of intellectual debates surrounding emancipation, that Gypsies were freed in Moldo-Wallachia.

In this context, there are a couple of points that merit emphasizing. First, the slaves that were owned by the monasteries and nobility were absolutely at the disposal of their masters and possessed no personal rights. Moreover, they were forced to settle on the outskirts of villages and were often buried in different cemeteries despite being Christians (Achim 1999). Second, the laws permitted masters to sell, exchange, or give away any Gypsies or their children, and often it would be the case that children would be separated from their families and given away to other masters. At the same time, in a recent treatment of Gypsy history in the Romanian Principalities, Viorel Achim (1999) has produced many documents to show that the Romanian slavery system was extremely flexible and pragmatic, allowing for great freedom even in comparison with the technically free Romanian serfs.

While initially blind to intermarriage, laws became prohibitive as mixed unions became more common (e.g., the Sobornicescul Hrisov of 1785 in Moldavia and the Pravilniceasca Condiță of 1780 in Wallachia, see Achim 1999:42–43). This change was partly a result of the fact that the old laws dictated that, if a Gypsy were to marry a non-Gypsy, the latter would become a slave him- or herself. However, when the demand for slaves declined, nobles and monasteries realized that serfs were more economically useful. As a result, prohibitive laws began to be enacted to prevent the conversion of free peasants into slaves. Thus, beginning in the late 18th century, even if laws prohibited mixed marriages, the children from such unions would most likely have become free and were probably assimilated into the local Romanian population.

In view of the above history, it is not entirely clear just how much admixture there is in the currently self-identified Roma population in the former Principalities, or in the Roma populations that migrated subsequent to emancipation in 1856. Assimilation was usually accompanied by the loss of Gypsy customs and language, and the wholesale adoption of Romanian ones (Achim 1999), whereas the nonassimilated populations remained in a state of limited and controlled nomadism (e.g., the *Kalderas*). Thus, the issue of admixture in Roma populations clearly requires further investigation, using both biological and cultural approaches.

### **Migrations during the 19th and 20th Centuries**

The migrations subsequent to the emancipation acts of the 1850s in Romania, known as the "great Kelderara invasion," constitute the second largest single event in the history of Roma populations in Europe. Although, as Patrick Williams (1984) notes, the arrival of these newcomers into Western



trends in their strategies of dealing with the Gypsies: mainly assimilation, both cultural and biological. Nomadism was forbidden in most European states, with the exception of France, where settling was forbidden (Guy 2001). However, most laws were contradictory and poorly enforced, leaving room for flexibility for the Gypsies, who were at times sometimes obliged through force to settle and at times prohibited from doing so. The same was true for mixed marriages.

In the Habsburg domains, for instance, under the “enlightened” rule of Maria-Theresa and Joseph II, a series of decrees were given beginning with 1761 and ending in 1783. The most significant one was the 1773 decree that forbade marriage between Gypsies and ordered that Gypsy children over the age of five should be taken away and brought up in non-Gypsy families. While these measures were not generally enforced, there was a limited supply of *gadžé* men and women who were willing to marry Gypsies, and Gypsy children frequently fled back to their parents (Fraser 1992a). However, the Burgenland (Austria) townships took the imperial decrees quite seriously, and, thus, one would expect a certain degree of admixture to have occurred there as a consequence.

Otherwise, endogamy seems to have been common for Gypsy groups throughout Central and Eastern Europe (Clébert 1969), despite examples to the contrary (e.g., an integrated *gadžé* husband among the *Parižoske Rom*; Williams 1984). The latter are generally isolated cases in which the children grew up in either one of the two environments, with future generations following that same pattern. The end result was that, after a few generations, they no longer identified with either one of the two cultural affiliations. This pattern is not surprising in the above context, both as a practice of resistance from the part of the Gypsies themselves and as a result of discrimination from the local populations.

The most obvious candidate for a behavioral adaptation to social and political pressures is the institution of the bride-price, which has been extensively documented ethnohistorically (e.g., Cohn 1973). Werner Cohn’s data comes from North American Gypsies but his results are generalizable in view of what has been recorded elsewhere (e.g., Bereczkei 1998). The bride-price was paid by the groom’s parents to the bride’s parents (the kin term for the relation between the two sets of parents is *xanamik*, pl. *xanamika*) in loosely arranged marriages, in which both sets of parents usually consulted the would-be bride and groom and gave them time to get to know each other. In practice, however, these arrangements often worked poorly, and elopement (*našimos*) sometime constituted the only way out of such a contract. Bride swapping (*parruimos*) is also ethnographically recorded, especially in the case of first-cousin marriage (Cohn 1973:56; Williams 1984). Cohn argues that the institution of the bride-price made it financially profitable to marry roughly within the family, in which the risk was lowered and the bride-price was either symbolic or dropped altogether.

In a series of demographic studies of Hungarian Gypsy and non-Gypsy populations, Tamás Bereczkei (1993, 1998) has shown that historically different reproductive strategies, and not socioeconomic and educational conditions, were responsible for a significant difference in fertility rates between the two ethnic groups. Bereczkei’s (1998) most recent study examined fertility in four rural villages (Magyarmecske, Besence, Bogádmindszent, and Gilvanfa) inhabited by communities of both Gypsies and non-Gypsy Hungarians of similar socioeconomic status. He observed that the difference in fertility rates ( $3.93 \pm 2.14$  newborns per Gypsy female v.  $2.32 \pm 1.24$  per non-Gypsy female; Bereczkei 1998:288) was not, as previously thought, because of the lack of education or to the relative differences in the use of contraceptives but, instead, to a difference in the use of kinship networks.

In this regard, the combination of an *r*-selected reproduction strategy—specifically, one that favors the production of many offspring—with the extension of kinship networks is viewed as an adaptive social mechanism for an economically depressed, marginalized, and largely endogamous population such as the Gypsies (Bereczkei 1993, 1998). In modern Hungarian Gypsy families, there is a lack of funds that would allow the concentration of resources in a few socioeconomically successful offspring (as is the case in non-Gypsy Hungarian families). These families compensate for this problem by having a large network of uncles, aunts, and cousins who are willing to contribute both financially and personally to child rearing. Overall, the presence of such reproductive adaptations in Gypsy/Roma populations is important to take into account when discussing diversification following multiple bottlenecks and founder effects.

## CONCLUDING REMARKS

The genetic data, together with the historical and linguistic evidence, suggest that the Roma or Gypsies do, indeed, seem to be descendent communities of an ancient population that arrived in Europe from Asia some 1,000 years ago, and one that underwent complex population biological processes including significant drift, admixture, and founder effects. The high frequency of the putative founder haplogroup IV-68 in the Roma also parallels another highly resolved male lineage that occurs predominantly in Jewish populations, the Cohen Modal Haplotype (Thomas et al. 2000). As such, these data generally contradict the hypothesis that the Gypsies are a primarily a socioculturally defined group of European descent (cf. Lucassen et al. 1998; Okely 1983; Wexler 1997).

Nevertheless, their complex history of migrations is not yet fully understood. This is partly a consequence of limited molecular genetic studies being carried out with Roma groups, and a significant sampling problem, which is intimately tied to current attitudes toward Roma in most European countries where racial and ethnic discrimination continues to be a problem. However, the idea that

contemporary Roma populations might have descended from multiple sources both within India and perhaps elsewhere in South/Central Asia must be subjected to further inquiry.

In this regard, the level of admixture within local populations is difficult to assess, especially in the Near East and southern Central Asia. However, molecular data differentiate Gypsy populations from their European counterparts. These findings suggest that, even if some coalescence of Roma and non-Roma peripatetic social groups did occur as a result of a common marginalized fate (as suggested in Lucassen et al. 1998), then it must have taken place on a limited scale.

An integrated population history approach to admixture and relations with the host populations is especially interesting for the case of the Gypsies, who, unlike Jews, have not kept an easily identifiable, homogeneous core of their Indian cultural practices. Future and more detailed studies of diversity within European and Asian Roma could provide important data for understanding changing cultural practices that may have affected the relations of Roma with their neighbors. While not abandoning the common origin and similarities among Gypsy groups, the tracing of the social construction of their various cultural identities from the point of entry into Europe until today will greatly contribute to a multidisciplinary approach to Roma studies.

Overall, the molecular genetic studies of the Roma unambiguously support the linguistic theory of an Indian, or at least a largely (South) Asian, origin for the Gypsies. They also support historical data for three major recent migrations of Gypsies in Europe: (1) the early migrations into North/West Europe; (2) the 17th–18th-century expansions from Moldo-Wallachia and Hungary; and (3) the 19th–early-20th-century exodus from the Romanian Old Kingdom following the abolition of slavery. These studies further reveal intriguing patterns of endogamy (and admixture) based on the distribution of maternal and paternal lineages in these groups, and the occurrence of rare Mendelian genetic disorders in Roma populations. The analysis of other genetic loci, such as those present on the X-chromosome (Schaffner 2004), will provide additional power to resolve the issues of population origins and admixture.

Unfortunately, an accurate description of the genetic profile of the proto-Roma remains elusive, partly because of the strong effects of genetic drift on the small population sizes with frequent fission events. It has been proposed that this pattern could be explained as a social adaptation to pressures from various host populations (Hancock 1987). However, this idea cannot be directly tested at this time because the samples were drawn from a relatively small number of individuals, with the Western European sample being especially sparse.

Given the social anthropological realities of groups such as the Spanish Gitanos (e.g., Gay y Blasco 2002), whose identity is not linguistically based and whose probable arrival in Europe took place 700 years ago, more detailed ge-

netic studies are required to assess the degree to which their genetic lineages overlap with those in Central and Eastern Europe. Some overlap seems likely due to the recent discovery of Gypsy-specific neuropathies in a family with Gypsy background in Spain (Thomas et al. 2001). In this regard, it would be greatly beneficial to systematically sample Roma groups for an extensive genetic study, as well as adjacent Romanian villager communities in the former Danubian Principalities from whence most of today's Gypsies are known to have migrated in the last three centuries.

Finally, we do not believe that understanding the Gypsies/Roma as a population primarily derived from a single or multiple sources in Asia would ultimately result in the proliferation of racial-criminal discourses. On the contrary, for a population that is overall both economically and socially marginalized, knowledge about certain biological characteristics of Gypsies is of key medical importance. This knowledge would allow for earlier detection and treatment of genetic diseases, as is currently done with Ashkenazi Jews and Tay-Sachs disease or with African Americans and sickle-cell anemia.

---

**RADU P. IOVIȚĂ** Department of Anthropology, University of Pennsylvania, Philadelphia, PA 19104-6398

**THEODORE G. SCHURR** Department of Anthropology, University of Pennsylvania, Philadelphia, PA 19104-6398

## NOTES

*Acknowledgments.* The authors would like to thank the anonymous reviewers for their insightful comments and helpful suggestions for revising this manuscript.

1. Recently, the term *Gypsy* has come to be regarded as derogatory, with many Gypsy/Roma/Sinti activists preferring the term *Roma*. This stance has come about largely through political campaigns for civil rights in various European countries but may actually not be representative of the way most European Gypsies identify themselves today (for a discussion of Roma activism and "Gypsiness," see Gay y Blasco 2002; Liégeois 1994). For this reason, we have chosen to refer to the populations as "Gypsy" when speaking in general and by preferred self-identifier (e.g., Roma, Sinti, Manuș, Čurara) when speaking about particular groups. In addition, we have chosen to preserve differences in spelling of Romani words that are because of differences in the original spelling in the cited works, so as to facilitate the readers' examination of original papers or books.
2. The choice of populations for sampling was probably somewhat influenced by Ian Hancock's theories (2000) as well as by the availability of the genetic data from India (e.g., Bamshad 2001; Kivisild 1999).

## REFERENCES CITED

- Achim, Viorel  
1999 *Tigani în Istoria României*. Bucharest: Editura Enciclopediei.
- Anderson, S., A. T. Bankier, B. G. Barrell, et al.  
1981 Sequence and Organization of the Human Mitochondrial Genome. *Nature* 290(5806):457–465.
- Bamshad, Michael, Toomas Kivisild, W. Scott Watkins, et al.  
2001 Genetic Evidence on the Origins of Indian Caste Populations. *Genome Research* 11(6):994–1004.
- Berczkei, Tamás  
1993 R-Selected Reproductive Strategies among Hungarian Gypsies: A Preliminary Analysis. *Ethology and Sociobiology* 14(2):71–88.

- 1998 Kinship Network, Direct Childcare, and Fertility among Hungarians and Gypsies. *Evolution and Human Behavior* 19(5):283–298.
- Chandler, David, Dora Angelicheva, Lisa Heather, et al.  
2000 Hereditary Motor and Sensory Neuropathy-Lom (HMSNL): Refined Genetic Mapping in Romani (Gypsy) Families from Several European Countries. *Neuromuscul Disord* 10(8):584–591.
- Clébert, Jean-Paul  
1969 *The Gypsies*. C. Duff, trans. Baltimore: Penguin.
- Cohn, Werner  
1973 *The Gypsies*. Reading, MA: Addison-Wesley Publishing.
- Colomer, Jaume, Cristina Iturriaga, Luba Klaydjieva, et al.  
2000 Hereditary Motor and Sensory Neuropathy-Lom (HMSNL) in a Spanish Family: Clinical, Electrophysiological, Pathological and Genetic Studies. *Neuromuscul Disord* 10(8):578–583.
- Fosztó, László, and Marian-Viorel Anastasoae  
2001 Romania: Representations, Public Policies and Political Projects. *In Between Past and Future: The Roma of Central and Eastern Europe*. W. Guy, ed. Pp. 351–370. Hatfield: University of Hertfordshire Press.
- Fraser, Angus  
1992a *The Gypsies*. Oxford: Blackwell.  
1992b Looking into the Seeds of Time. *Tsiganologische Studien* 1:135–166.
- Gay y Blasco, Paloma  
1999 *Gypsies in Madrid: Sex, Gender, and the Performance of Identity*. Oxford: Berg.  
2002 Gypsy/Roma Diasporas: A Comparative Perspective. *Social Anthropology* 10(2):173–188.
- Grellmann, Heinrich Moritz Gottlieb  
1787 *Die Zigeuner: Ein historischer Versuch über die Lebensart und Verfassung, Sitten und Schicksale dieses Volks in Europa, nebst ihrem Ursprunge*. Göttingen: Johann Christian Dieterich.
- Gresham, David, Bharti Morar, Peter A. Underhill, et al.  
2001 Origins and Divergence of the Roma (Gypsies). *American Journal of Human Genetics* 69:1314–1331.
- Guy, Will, ed.  
2001 *Between Past and Future: The Roma of Central and Eastern Europe*. Hatfield: University of Hertfordshire Press.
- Hancock, Ian  
1987 *The Pariah Syndrome*. Ann Arbor, MI: Karoma Publishers.  
2000 The Emergence of Romani as a Koine outside of India. *In Scholarship and Gypsy Struggle: Commitment in Romani Studies*. T. Acton, ed. Pp. 1–13. Hatfield: University of Hertfordshire Press.
- Heyer, Evelynne  
1999 One Founder/One Gene Hypothesis in a New Expanding Population: Saguenay (Quebec, Canada). *Human Biology* 71(1):99–109.
- Horai, Satoshi, Kenji Hayasaka, Rumi Kondo, et al.  
1995 Recent African Origin of Modern Humans Revealed by Complete Sequences of Hominoid Mitochondrial DNAs. *Proceedings of the National Academy of Science U.S.A.* 92(2):532–536.
- Ingman, Max, Henrik Kaessmann, Svante Pääbo, et al.  
2000 Mitochondrial Genome Variation and the Origin of Modern Humans. *Nature* 408(6813):708–713.
- Kalaydjieva, Luba, Francesc Calafell, Mark A. Jobling, et al.  
2001a Patterns of Inter- and Intra-Group Genetic Diversity in the Vlax Roma as Revealed by Y Chromosome and Mitochondrial DNA Lineages. *European Journal of Human Genetics* 9:97–104.
- Kalaydjieva, Luba, David Gresham, and Francesc Calafell  
2001b Genetic Studies of the Roma (Gypsies): A Review. *BMC Medical Genetics* 2:5–18.
- Kivisild, Toomas, Michael J. Bamshad, K. Kaldma, et al.  
1999 Deep Common Ancestry of Indian and Western-Eurasian Mitochondrial DNA Lineages. *Current Biology* 9:1331–1334.
- Kochanowski, Vania de Gila  
1990 Migrations aryennes et indo-aryennes. *Diogenes* 149(119–141).
- Kogalniceanu, Mihail  
1837 *Esquisse sur l'histoire, les moeurs et la langue des Cigains, connus en France sous le nom de Bohémiens: suivie d'un recueil de sept cents mots cigains*. Berlin: B. Behr.
- Liégeois, Jean-Pierre  
1994 *Roma, Gypsies, Travellers*. Strasbourg: Council of Europe.
- Lucassen, Leo, Wim Willems, and Annemarie Cottaar  
1998 *Gypsies and Other Itinerant Groups: A Socio-Historical Approach*. London: Macmillan Press.
- Macaulay, Vincent, Martin Richards, Eileen Hickey, et al.  
1999 The Emerging Tree of West Eurasian mtDNAs: A Synthesis of Control-Region Sequences and RFLPs. *American Journal of Human Genetics* 64:232–249.
- Matras, Yaron  
1996 Review of Boretzky and Iglá 1994b. *Zeitschrift für Balkanologie* 32:214–224.  
2002 *Romani: A Linguistic Introduction*. Cambridge: Cambridge University Press.
- Miklosich, Franz  
1872 *Über die Mundarten und die Wanderungen der Zigeuner Europas*. 12 vols. Wien: In commission bei K. Gerold's sohn.
- Motulsky, Arno G.  
1995 Jewish Diseases and Origins. *Nature Genetics* 9(2):99–101.
- Okely, Judith  
1983 *The Traveller-Gypsies*. Cambridge: Cambridge University Press.
- Panaiteescu, P. N.  
1941 The Gypsies in Wallachia and Moldavia: A Chapter of Economic History. *Journal of the Gypsy Lore Society* 3(20):58–72.
- Peltonen, Leena, Anu Jalanko, and Teppo Varilo  
1999 Molecular Genetics of the Finnish Disease Heritage. *Human Molecular Genetics* 8(10):1913–1923.
- Quintana-Murci, Lluís, et al.  
1999 Genetic Evidence of an Early Exit of Homo Sapiens Sapiens from Africa through Eastern Africa. *Nature Genetics* 23:437–441.
- Richards, Martin, Vincent Macaulay, Eileen Hickey, et al.  
2000 Tracing European Founder Lineages in the Near Eastern mtDNA Pool. *American Journal of Human Genetics* 67:1251–1276.
- Richards, Martin B., Vincent A. Macaulay, Hans-Jürgen Bandelt, et al.  
1998 Phylogeography of Mitochondrial DNA in Western Europe. *Annals of Human Genetics* 62(part 3):241–260.
- Richards, Martin, Vincent Macaulay, Antonio Torroni, et al.  
2002 In Search of Geographical Patterns in European Mitochondrial DNA. *American Journal of Human Genetics* 71(5):1168–1174.
- Rosser, Zoë H., Tatiana Zerjal, Matthew E. Hurles, et al.  
2000 Y-Chromosomal Diversity in Europe Is Clinal and Influenced Primarily by Geography, Rather Than by Language. *American Journal of Human Genetics* 67(6):1526–1543.
- Said, Edward  
1978 *Orientalism*. New York: Routledge and KeganPaul.
- Samson, John  
1926 *The Dialect of the Gypsies of Wales: Being the Older Form of British Romani Preserved in the Speech of the Clan of Abram Wood*. Oxford: Clarendon Press.
- Schaffner, Stephan F.  
2004 The X Chromosome in Population Genetics. *Nature Reviews Genetics* 5:43–51.
- Semino, Ornella, Giuseppe Passarino, Peter J. Oefner, et al.  
2000 The Genetic Legacy of Paleolithic Homo Sapiens Sapiens in Extant Europeans: A Y Chromosome Perspective. *Science* 290:1155–1159.
- Simoni, Lucia, Francesc Calafell, Davide Pettener, et al.  
2000 Geographic Patterns of mtDNA Diversity in Europe. *American Journal of Human Genetics* 66(1):262–278.
- Soulis, George Christos  
1961 *The Gypsies in the Byzantine Empire and the Balkans in the Late Middle Ages*. *Dumbarton Oaks Papers* 15:142–165.
- Tálos, Endre  
1999 *Etymologica Zingarica*. *Acta Linguistica Hungarica* 46:215–268.
- Thomas, Mark G., Tudor Parfitt, Deborah A. Weiss, et al.  
2000 Y Chromosomes Traveling South: The Cohen Modal Haplotype and the Origins of the Lemba—The “Black Jews

- of Southern Africa." *American Journal of Human Genetics* 66:674–686.
- Thomas, Peter Kynaston, Luba Kalaydjieva, Brian Youl, et al.  
2001 Hereditary Motor and Sensory Neuropathy-Russe: New Autosomal Recessive Neuropathy in Balkan Gypsies. *Annals of Neurology* 50(4):452–457.
- Torrioni, Antonio, Kirsi Huoponen, Paolo Francalacci, et al.  
1996 Classification of European mtDNAs from an Analysis of Three European Populations. *Genetics* 144(4):1835–1850.
- Tournev, Ivailo, Luba Kalaydjieva, Bryan Youl, et al.  
1999 Congenital Cataracts Facial Dysmorphism Neuropathy Syndrome, a Novel Complex Genetic Disease in Balkan Gypsies: Clinical and Electrophysiological Observations. *Annals of Neurology* 45(6):742–750.
- Turner, Ralph L.  
1926 The Position of Romani in Indo-Aryan. *Journal of the Gypsy Lore Society* 3(5):145–189.
- Underhill, Peter A., Giuseppe Passarino, Alice A. Lin, et al.  
2000 Y Chromosome Sequence Variation and the History of Human Populations. *Nature Genetics* 26:358–361.
- Underhill, Peter A., Giuseppe Passarino, Alice A. Lin, et al.  
2001 The Phylogeography of Y Chromosome Binary Haplotypes and the Origins of Modern Human Populations. *Annals of Human Genetics* 65:43–62.
- Wexler, Paul  
1997 Could There Be a Rotwelsch Origin for the Romani Lexicon? 3rd International Conference on Romani Linguistics, Prague, December 1996.
- Willems, Wim  
1998 Ethnicity as a Death-Trap: The History of Gypsy Studies. *In* *Gypsies and Other Itinerant Groups: A Socio-Historical Approach*. L. Lucassen and W. Willems, ed. Pp. 17–34. New York: MacMillan.
- Williams, Patrick  
1984 *Mariage Tsigane: Une cérémonie de fiançailles chez les Rom de Paris*. Paris: L'Harmattan.