

Review

Genetic studies of the Roma (Gypsies): a review

Luba Kalaydjieva*^{1,2}, David Gresham¹ and Francesc Calafell³

Address: ¹Centre for Human Genetics, Edith Cowan University, Perth, Australia, ²Western Australian Institute for Medical Research, Perth, Australia and ³Unitat de Biologia Evolutiva, Facultat de Ciències de la Salut i de la Vida, Universitat Pompeu Fabra, Barcelona, Spain

E-mail: Luba Kalaydjieva* - L.Kalaydjieva@ecu.edu.au; David Gresham - D.Gresham@ecu.edu.au;

Francesc Calafell - francesc.calafell@cexs.upf.es

*Corresponding author

Published: 2 April 2001

BMC Medical Genetics 2001, 2:5

Received: 15 January 2001

Accepted: 2 April 2001

This article is available from: <http://www.biomedcentral.com/1471-2350/2/5>

(c) 2001 Kalaydjieva et al, licensee BioMed Central Ltd.

Abstract

Background: Data provided by the social sciences as well as genetic research suggest that the 8-10 million Roma (Gypsies) who live in Europe today are best described as a conglomerate of genetically isolated founder populations. The relationship between the traditional social structure observed by the Roma, where the Group is the primary unit, and the boundaries, demographic history and biological relatedness of the diverse founder populations appears complex and has not been addressed by population genetic studies.

Results: Recent medical genetic research has identified a number of novel, or previously known but rare conditions, caused by private founder mutations. A summary of the findings, provided in this review, should assist diagnosis and counselling in affected families, and promote future collaborative research. The available incomplete epidemiological data suggest a non-random distribution of disease-causing mutations among Romani groups.

Conclusion: Although far from systematic, the published information indicates that medical genetics has an important role to play in improving the health of this underprivileged and forgotten people of Europe. Reported carrier rates for some Mendelian disorders are in the range of 5 -15%, sufficient to justify newborn screening and early treatment, or community-based education and carrier testing programs for disorders where no therapy is currently available. To be most productive, future studies of the epidemiology of single gene disorders should take social organisation and cultural anthropology into consideration, thus allowing the targeting of public health programs and contributing to the understanding of population structure and demographic history of the Roma.

Introduction

The Roma (Gypsies) became one of the peoples of Europe around one thousand years ago, when they first arrived in the Balkans [1,2]. The current size of the European Romani population, around 8 million [2], is equivalent to that of an average European country (Figure 1). While human rights and socio-economic issues related to the Roma are increasingly becoming the focus

of political debate and media coverage throughout Europe, their poor health status [3-6] is rarely discussed and still awaits the attention of the medical profession.

This review of genetic studies of the Roma was prompted by two recent developments: (i) Studies conducted over the last decade have resulted in the identification of a number of novel single gene disorders and disease-caus-

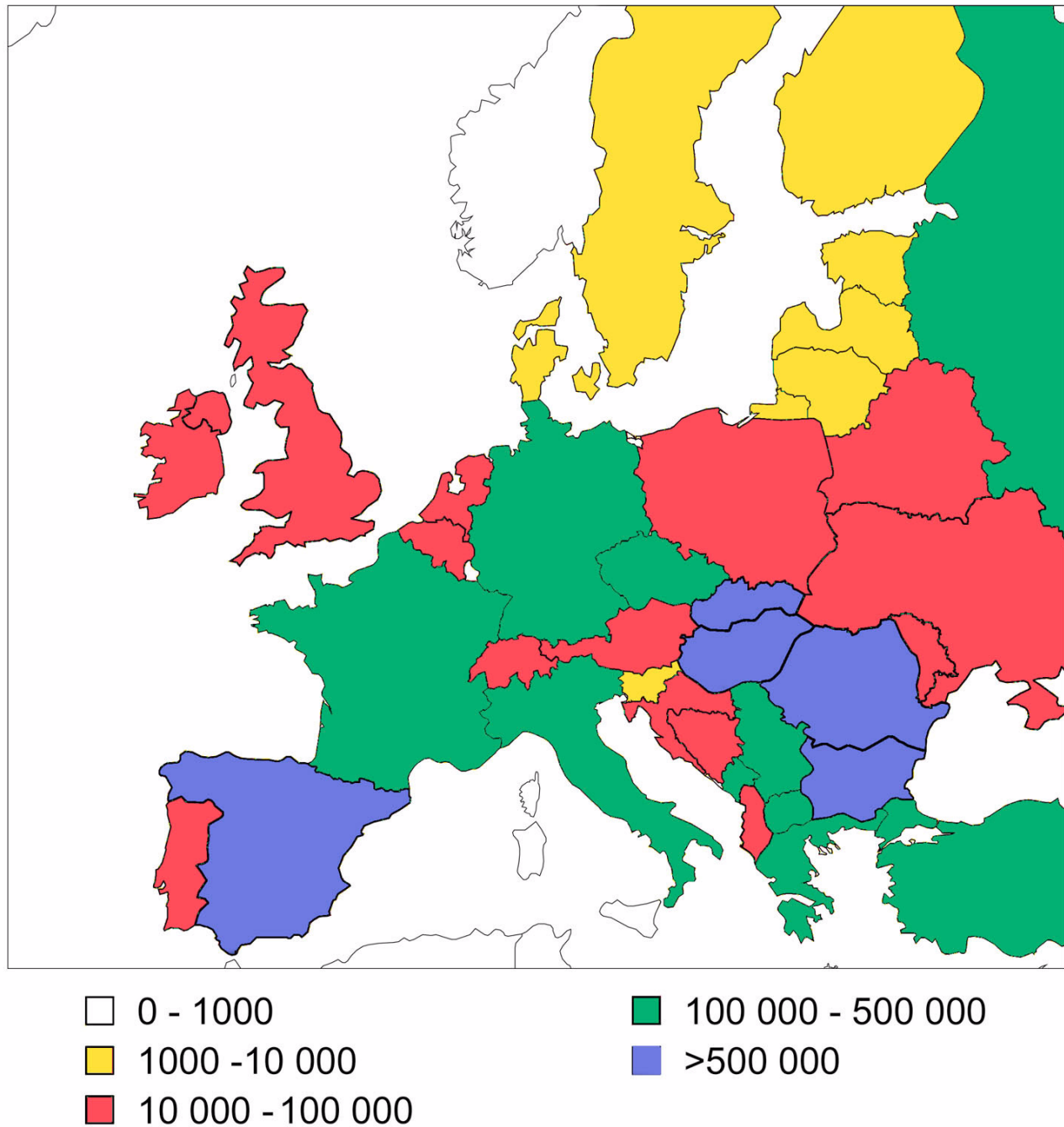


Figure 1
Romani population size in different European countries The collection of this type of data depends on declared ethnic identity which, in the case of the Roma, can be affected by a number of political and social circumstances. The estimates in the figure are the average of the numbers provided by different sources, such as census data, ministries of internal affairs and human rights organizations [2].

ing mutations. The accumulating data are already sufficient to outline a pattern and draw conclusions about public health policies and future research. (ii) The eco-

nomie and political changes in Eastern Europe and the wars in former Yugoslavia have led to the west-bound migration of large numbers of Roma [7,8], changing the

traditional demographic profile of Gypsy minorities across Europe. A predictable consequence of this new diaspora is that medical practitioners in many countries will encounter Romani patients with previously unknown or very rare disorders. A summary of the available information should facilitate diagnostic investigations and counselling in these affected families and stimulate international collaboration.

Materials and Methods

Literature searches were performed using the U.S.A National Library of Medicine PubMed/MEDLINE databases for the period 1960 to December 2000. Database searches using the keyword "Gypsies" identified 297 articles whilst the keyword "Gypsy" produced 573 articles. The discrepancy is due mainly to the inclusion of articles about the "gypsy retransposable element" and the "gypsy moth". Searches using the terms "Roma", "Romani" and "Romany" yielded results that were not relevant to the topic (eg. Roma, the capital of Italy) or else incomplete.

The majority of the 297 articles dealt with issues beyond the focus of this review, namely social problems related to the health of the Roma (28.6%), or general medical problems (29.6%). The remainder were reports on genetic research, of which 41 studies (13.8%) were in the field of clinical genetics, 44 (14.8%) were molecular studies of genetic disorders, and 39 (13.1%) covered population genetic research. In the clinical and molecular genetics fields, we have given preference to publications which were not limited to single case descriptions, and dealt with disorders with public health impact. Population genetics papers were selected on the basis of the compatibility of study design, specifically the analysis of comparable polymorphic systems.

Complementary data on history, linguistics, cultural anthropology and demography were found through standard library and bibliographic searches, and included publications recommended by consulting experts in Romani studies (Drs. Elena Marushiakova and Vesselin Popov from the Bulgarian Academy of Sciences and Dr. Ian Hancock from the University of Texas at Austin).

The "Track Record" of Genetics

Genetic studies of the Roma have been conducted for over 70 years, with thousands of individuals sampled across Europe. During the years of the Third Reich, Gypsies, together with Jews, attracted the special attention of German geneticists [9]. A grant proposal signed by Nobel prize winner Ferdinand Sauerbruch and funded by the Deutsche Forschungsgemeinschaft designed the "genetic and medical research" at the death camp in Auschwitz [9]. The Race Hygiene and Population Biology Research Centre, established in 1936, organised thorough records

of Jewish and Romani pedigrees and provided "the scientific basis" for the "final solution", the annihilation of millions of Jews and Roma in the concentration camps of Nazi-occupied Europe.

Post-war genetic research has been preoccupied with the Indian origins of the Roma [10–16], pursuing the "Indian connection" even in studies meant to focus on severe genetic disorders [17]. Most studies have remained in the realm of scientific exploration, away from the health needs of the Roma. Many publications display judgemental and paternalistic attitudes, that would be considered unacceptable if used with regard to other populations.

This historical "track record", the persisting practices of discrimination and marginalisation [3–6], and the fact that, unlike the Jews, the Finns and the French Canadians, the Roma are still the "object" of investigations conducted by outsiders, are all likely to impact on the attitudes of the Roma towards genetics. Building up the trust and collaboration necessary for both public health programs and research, should become a goal of the health care systems of Europe.

Population Genetics

Population genetic studies have used mostly "classical" polymorphisms to investigate Romani individuals from different European countries and address three main questions: (i) similarity between Roma and Indians; (ii) relatedness to European populations; (iii) affinities between Romani populations from different countries [10–24]. Single locus comparisons have resulted in controversy, with some pointing to close genetic affinity between Roma and Indians, and others indicating that the Roma are indistinguishable from Europeans. Heterogeneity between countries has become apparent and has led to the conclusion that the European Roma are composed of two different populations, characterised respectively by a high and a low frequency of blood group B [23], or defined as East and West European Roma, with the former closely related to Indian populations [16]. Heterogeneity of Romani populations within the same country has been suggested by the very small number of studies addressing this issue [19,21,25,26].

In an attempt to summarise the existing data, we have conducted a multilocus re-analysis of several marker systems using comparable studies of the Roma in different European countries [11,22,23,24,26,27], Europeans [28–33] and north Indian populations [34–36]. The comparisons (Figure 2, Table 1) provide a general indication that most of the Roma are genetically closer to Indians than to European populations, hardly surprising in view of the linguistic theories on the Indian origins of the

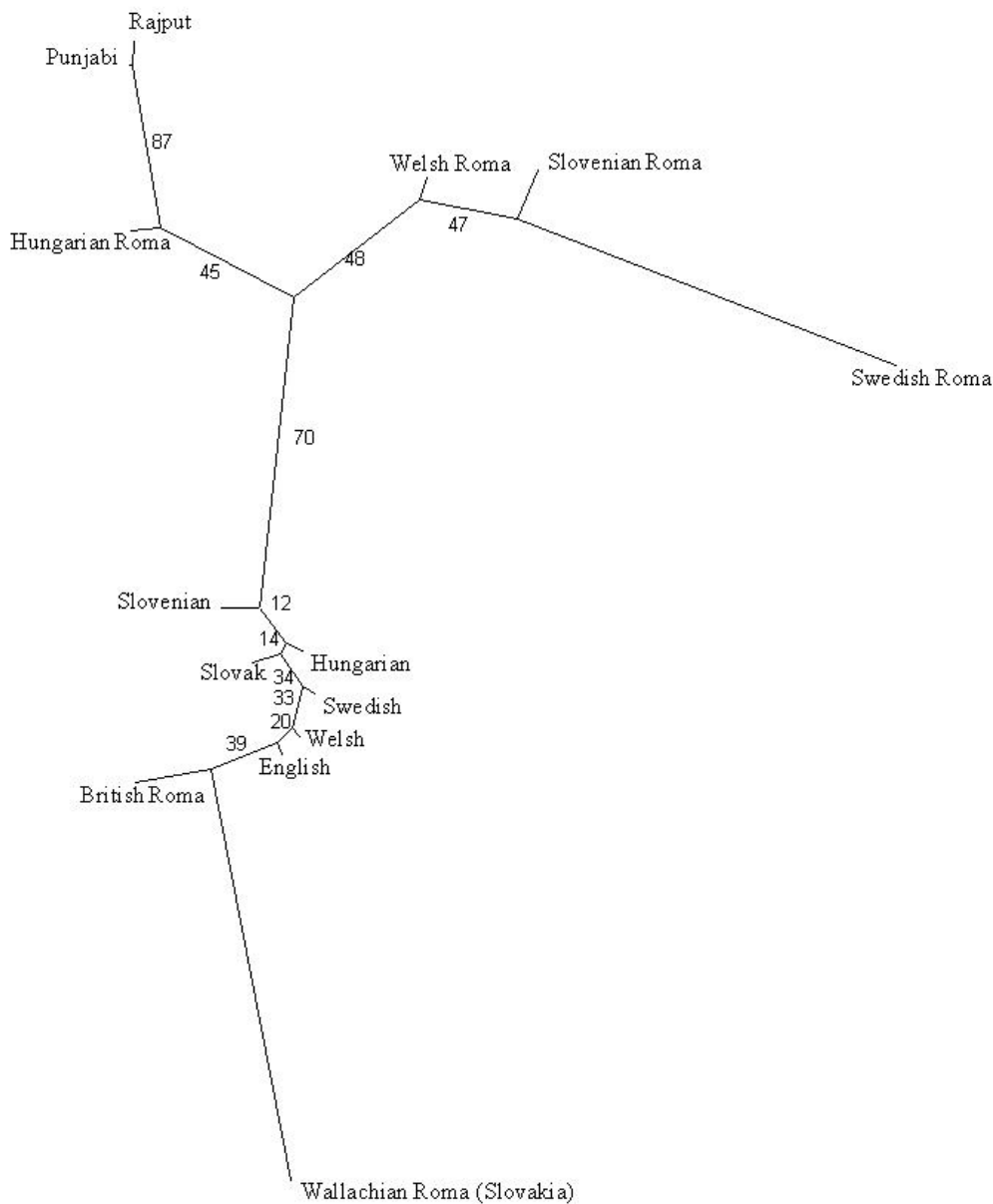


Figure 2
Multilocus comparison between Romani populations from different European countries, autochthonous European populations and populations from north India The polymorphic systems included in the analysis comprised A₁A₂BO, MN, haptoglobin and Rh (CDE), with a total of 11 independent alleles. Information on these markers was available for the Roma in Slovakia (n = 350) [26], Hungary (n = 507) [11], England (n = 109) [23], Slovenia (n = 350) [27], Sweden (n = 115) [24] and Wales (n = 84) [22], for non-Roma Europeans (n = 5169) and for two north Indian populations, Rajput (n = 175) [34,35] and Punjabi (n = 140) [35,36]. Genetic distances between pairs of populations were computed by means of Reynold's coancestry coefficient [84] and displayed as a neighbour-joining tree [85]. The robustness of the branches in the tree was assessed with a bootstrap approach [86]. The analysis was conducted using the PHYLIP 3.57c package [87].

Roma proposed two centuries ago [1]. More importantly, the analysis highlights the internal diversity of the Roma, who appear to be genetically far more heterogeneous than autochthonous European populations.

Table 1: Multi-locus reanalysis of previously published data on European Roma

Populations	Proportion of the variance explained by differences	
	between groups	within groups
All Roma (n = 1287) versus non-Roma Europeans (n = 5169)	1.81 ± 1.45%	0.58 ± 0.29%
Roma (n = 1287) versus North Indians (n = 315)	0.36 ± 0.69%	2.8 ± 0.39%
Between Romani populations (n = 1287) in Europe	3.47 ± 0.46%	
Between European Populations (n = 5169)	0.19 ± 0.12%	

The populations and references included in the comparison are as indicated in figure 2. Genetic variance was apportioned between and within populations and between and within groups of populations by means of the Analysis of Molecular Variance [88], as implemented in the Arlequin 1.1 package [89].

Genetic Disorders of the Roma Diseases and mutations identified to-date

As a result of traditionally low socio-economic status and limited access of the Roma to health care, their unique genetic heritage has long escaped the attention of European medicine and is now being randomly "discovered".

To date, nine Mendelian disorders caused by private "Romani" mutations have been described (Table 2).

The list includes three novel neurological disorders, namely hereditary motor and sensory neuropathies type Lom (HMSN-L) [37–39] and type Russe (HMSN-R) [40], and the congenital cataracts facial dysmorphism neuropathy syndrome (CCFDN) [41,42].

In addition, a number of previously known but rare disorders have been identified and shown to be caused by novel private mutations (Table 2). Examples include limb-girdle muscular dystrophy type 2C (LGMD2C) [43], galactokinase deficiency [44], primary congenital glaucoma [45], and congenital myasthenia [46].

In view of the lack of systematic studies, the list cannot be comprehensive and is likely to represent the biases

and interests of individual medical researchers working in this field. Data in the literature, particularly from the Spanish Collaborative Study of Congenital Malformations [47], point to the existence of a number of additional rare single gene disorders, whose molecular basis is still to be identified. These include hereditary idiopathic torsion dystonia (ITD) [48], epidermolysis bullosa [49], albinism [49], and some rare autosomal recessive malformation syndromes, such as Bowen-Conradi, Jarcho-Levin, Meckel, Smith-Lemli-Opitz, and Fraser [47,49].

A third group of Mendelian disorders includes common conditions, where the mutation prevalent in the surrounding or in global populations is likely to have been introduced by admixture, for example cystic fibrosis and delF508 [50], phenylketonuria and the R252W and IVS10nt546 mutations [51,52], and medium chain acyl-coenzyme A dehydrogenase (MCAD) deficiency and G985 [53].

Molecular genetic findings

With the exception of phenylketonuria, Mendelian disorders have been described as genetically homogeneous, with a single mutation accounting for all affected individuals and related polymorphic haplotypes unambiguously indicating a common origin and founder effect [37–40,42–46].

At the same time, many studies emphasise the small size of the conserved region of homozygosity and the diversity of disease haplotypes observed even within single affected kindreds [37,40,42,44,54] (Figure 3). Haplotype diversification, generated by numerous historical recombinations and marker mutations [39] as a consequence of the old age [37,43] and high frequencies of disease-causing mutations, has important implications for gene mapping studies: (i) Homozygosity mapping, relying on consanguinity in the affected families, is not applicable in studies using the standard genetic maps and can result in spuriously negative results [54]. (ii) The diversity of historical recombinations becomes a powerful tool in the subsequent refined genetic mapping and positional cloning of disease genes [55,39].

Epidemiological data

Reported gene frequencies are high for both private and "imported" mutations, and often exceed by an order of magnitude those for global populations. For example, galactokinase deficiency whose worldwide frequency is 1:150,000 to 1:1,000,000 [56,57] affects 1 in 5,000 Romani children [44]; autosomal dominant polycystic kidney disease (ADPKD) has a global prevalence of 1:1000 individuals worldwide [58] and 1:40 among the Roma in some parts of Hungary [17]; primary congenital glauco-

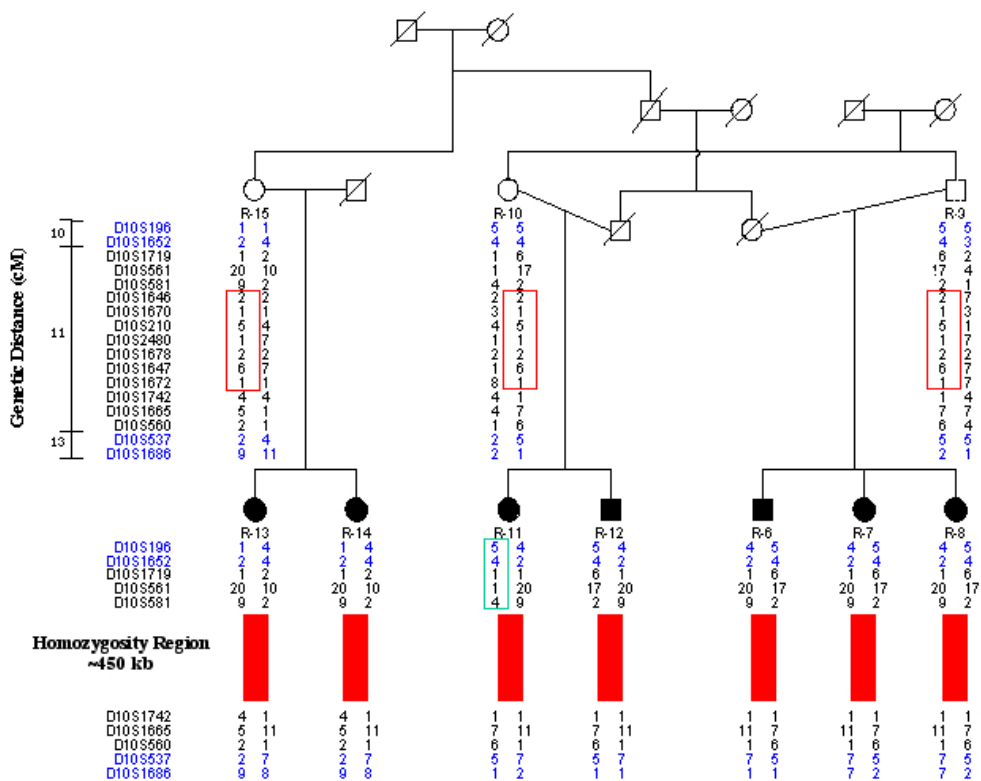


Figure 3
Genetic mapping of hereditary motor and sensory neuropathy type Russe (HMSN-R): findings in the region of linkage on chromosome 10q23 This affected family originates from a closed endogamous Romani group where consanguineous marriages are common. The linkage study was conducted using the ABI Prism Linkage Mapping Sets LMS and LMS version 2 (PE Biosystems), with an average intermarker distance of 10 cM. The ABI panel markers flanking the HMSN-R region (shown in blue) presented with different alleles in the affected individuals. Haplotype heterozygosity, resulting from historical recombinations and a recent cross-over event (individual R-11), extended into the 10 cM interval containing the gene and could have resulted in exclusion of the region if homozygosity mapping had been used. In the set of affected families included in the original study [40], the conserved region of homozygosity (red bars) was found to span only <500 kb. Courtesy of Dr. Tamara Rogers.

ma ranges between 1:5,000 and 1:22,000 worldwide [59,60] and about 1:400 among the Roma in Central Slovakia [61,62].

Carrier rates for a number of disorders have been estimated to be in the 5 to 20% range (Table 3).

Although incomplete, the available data already lead to some practical conclusions: (i) What may appear to be a novel disorder confined to a single family, could in fact be an indication of a common problem affecting large numbers of individuals. Research should therefore extend beyond case descriptions and aim at more comprehensive epidemiological information. (ii) The emphasis

on consanguinity in affected families displaces the focus from an obvious need for public health intervention to patterns of personal behaviour. In the face of the reported high gene frequencies, consanguinity is no more relevant than it would be as a cause of beta-thalassemia in Mediterranean countries. (iii) High gene frequencies may result in the parallel segregation of phenotypically similar but genetically distinct disorders within the same kindred [40,42]. This clustering should be borne in mind in diagnostic studies, where assumptions based on pedigree structure should be avoided and independent clinical and genetic assessment should be conducted in all cases.

Table 2: Mendelian disorders of the Roma caused by private founder mutations

Disorder	OMIM*	Inheritance	Map Location	Gene	Mutation	Ref.
Primary congenital Glaucoma	231300	AR	2p21	CYP1B1	E387K	45,54
Galactokinase Deficiency	230200	AR	17q24	GK1	P28T	44
Polycystic kidney Disease	173900	AD	4q21-q23	PKD2	R306X**	90
Hereditary motor and Sensory neuropathy-Lom	601455	AR	8q24	NDRG1	R148X	37,39
Hereditary motor and Sensory neuropathy-Russe	605285	AR	10q23			40
Congenital cataracts facial dysmorphism neuropathy	604168	AR	18qter			42
Limb girdle muscular dystrophy type 2C	253700	AR	13q12	SGCG	C283Y	43,65, 91
Congenital myasthenia	254210	AR	17p13	CHRNE	I267delG	46
Glanzmann Thrombasthenia	273800	AR	17q21	ITGA2B	IVS15DS, G-A+I	64,92

* Using the OMIM numbers, detailed clinical information can be obtained at <http://www3.ncbi.nlm.nih.gov/Omim/> ** The R306X mutation in PKD2 has been identified in Romani families from Bulgaria. It has not been confirmed in the Hungarian ADPKD families, but appears probable because of a reported common migration history of all affected groups.

Table 3: Reported carrier rates for single gene disorders among the Roma

Disorder	Country	General Roma	High-risk groups	Ref.
Primary congenital glaucoma	Slovakia	5%	*11%	45,54
Galactokinase Deficiency	Bulgaria	2%	*4%-5%	44
Autosomal dominant polycystic kidney disease	Hungary		2.4%	17
Hereditary motor and sensory neuropathy-Lom	Bulgaria	*2%	*20%	37,39
Limb girdle muscular dystrophy type 2C	**Bulgaria	2%	6%	93,66
MCAD deficiency	***Spain		*2.5%-10%	53
Phenylketonuria	Czecho slovakia	6%		94
Oculocutaneous albinism	Spain	3.4%		49
Fraser syndrome	Spain	2.7%		47
Epidermolysis bullosa	Spain	2.4%		49

Most estimates are based on prevalence figures. *Carrier rates determined through direct mutation detection are indicated in red. **The LGMD2C carrier rates for the general Romani population of Bulgaria are probably an overestimate since the screening was conducted in a geographical region where the high risk groups are clustered. ***The screening for the G985 mutation in Spain, performed in Gypsy groups residing in different parts of the country, revealed substantial differences between groups.

Research into Mendelian disorders has provided ample evidence of genetic stratification, with mutations occurring at high frequencies in some Romani communities and altogether absent in others, located in close geographic proximity. In some cases, such as Glanzmann thrombasthenia [63,64], LGMD2C [65,66], galactokinase deficiency [44], CCFDN [42] and HMSN-R [40], the identity of the affected groups has been specified. Other studies, for example of congenital glaucoma [61,62] and ADPKD [17] provide only an indication of the area of residence of the affected communities. In the few cases where gene frequencies can be compared between high-risk groups and the general Romani population of the same country, substantial differences become apparent (Table 3).

At the same time, founder mutations have spread with the Romani diaspora and are shared by affected individuals throughout Europe (Figure 4). International collaboration has already made a substantial contribution to the study of disease phenotypes in large samples of genetically homogeneous patients [46, 67–71] as well as to the refined mapping of disease genes [55]. Such collaboration will be essential for future research into new disorders, founder mutations and factors modifying disease severity, and for understanding the epidemiology of genetic diseases of the Roma. The first steps to European collaboration have been made, with the founding of the Gypsy Genetic Heritage Consortium in 1997, and the forthcoming ENMC workshop on neuromuscular disorders in Gypsies.

Historical demographic data are limited, however tax registries and census data give an approximate idea of population size and rate of demographic growth through the centuries (Table 4). A small size of the original population is suggested by the fact that although most of the migrants arriving in Europe in the 11th-12th century remained within the limits of the Ottoman Empire [1,75], the overall number of Roma in its Balkan provinces in the 15th century was estimated at only 17,000.

During its subsequent history in Europe, this founder population split into numerous socially divided and geographically dispersed endogamous groups, with historical records from different parts of the continent consistently describing the travelling Gypsies as "a group of 30 to 100 people led by an elder" [1,2]. These splits, a possible compound product of the ancestral tradition of the *jatis* of India, and the new social pressures in Europe (e.g. Gypsy slavery in Romania [76] and repressive legislation banning Gypsies from most western European countries [1,2]), can be regarded as secondary bottlenecks, reducing further the number of unrelated founders in each group. The historical formation of the present-day 8 million Romani population of Europe is

Discussion

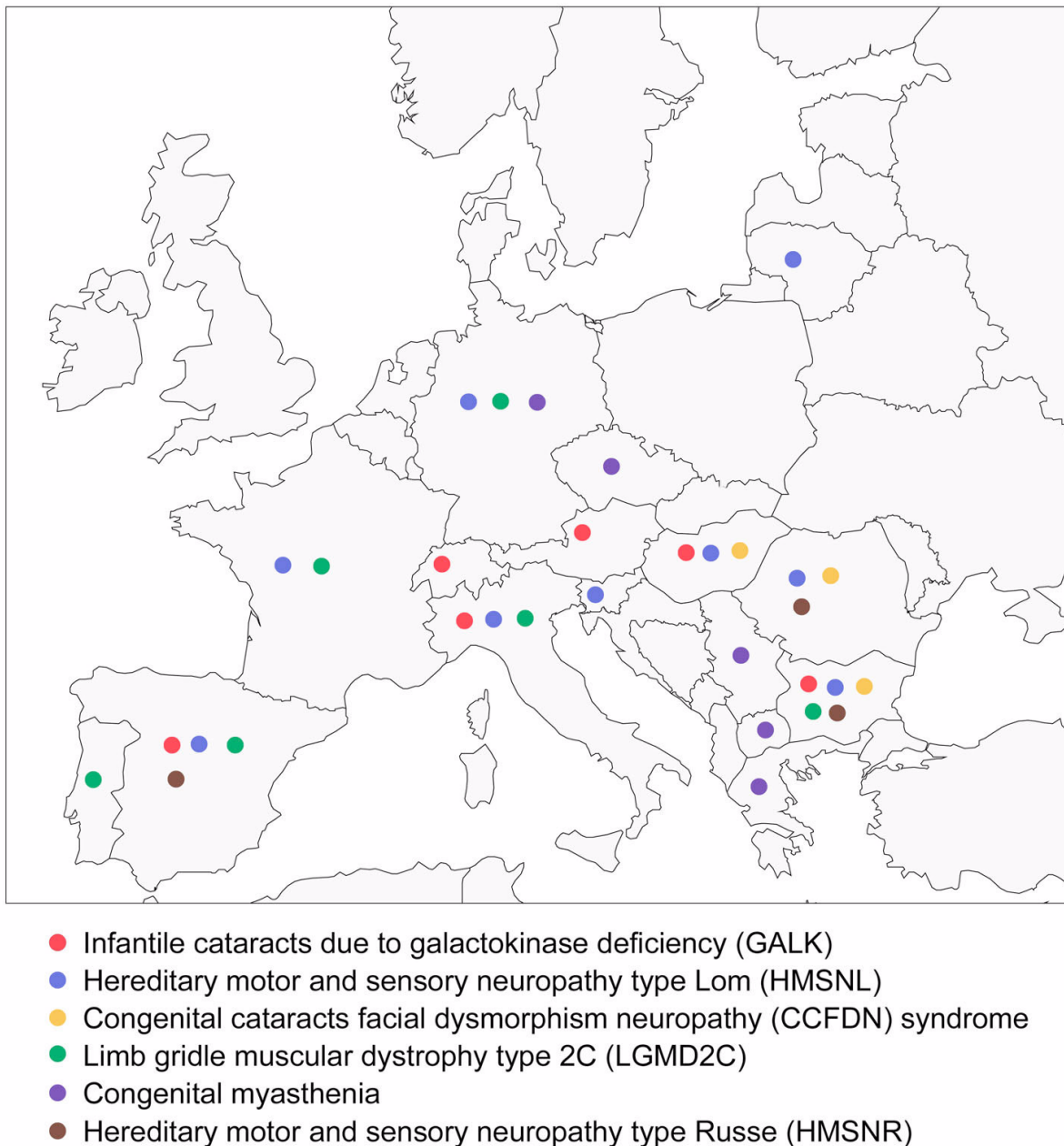
The pattern emerging from genetic research is that of a conglomerate of founder populations which extend across Europe but at the same time differ within individual countries, and whose demographic history, internal structure and relationships are poorly understood. An insight is provided by the social sciences.

The 18th century theory on the Indian origins of the Roma [reviewed in 1], is based on the similarities between Romani and languages spoken in the Indian subcontinent and is supported by genetic evidence. However the lack of close relationship to any specific language or dialect has left unresolved the question of the original ethnic composition of the proto-Roma, with both single [72,73] and diverse [74] origins proposed by linguists. Translated into the language of genetics, this is a relevant question related to the homogeneity or diversity of the founding population.

Inferred from linguistic influences retained in all Romani dialects, the major migration routes pass through Persia, Armenia, Greece and the Slavic-speaking parts of the Balkans [1]. The first documents pointing to the arrival of the Roma in the Balkans date from the 11th-12th century [1,75]. By the 15th century, mention of their presence can be found in historical records from all parts of Europe [1,2].

therefore the product of the complex initial migrations of numerous small groups, superimposed on which are two large waves of recent migrations from the Balkans into Western Europe, in the 19th - early 20th century, after the abolition of slavery in Rumania [1,2,76] and over the last decade, after the political changes in Eastern Europe [7,8].

The Group is still the primary building block of the social organisation of the Roma [1,2]. Group identity and the ensuing divisions and rules of endogamy are based on tradition, customs and organs of self-rule, language and dialects, trades, history of migrations, and religion. Individual groups can be classified into major metagroups [1,2,75]: the Roma of East European extraction; the Sinti in Germany and Manouches in France and Catalonia; the Kaló in Spain, Ciganos in Portugal and Gitans of southern France; and the Romanichals of Britain [1]. The greatest diversity is found in the Balkans, where numerous groups with well defined social boundaries exist. The 700-800,000 Roma in Bulgaria belong to three metagroups, comprising a large number of smaller groups [75].

**Figure 4**

Distribution of reported founder mutations in Europe The figure is based on available information referring to the following disorders: • Infantile cataracts due to galactokinase deficiency in Bulgaria [44], Austria [95], Switzerland [96], Italy [97], Hungary and Spain [our unpublished findings]. • Hereditary motor and sensory neuropathy - Lom in Bulgaria [37,38], Italy [68], Slovenia [69], Germany [70], Spain [71], France and Romania [55] and Hungary [our unpublished findings]. • Congenital cataracts facial dysmorphism neuropathy syndrome in Bulgaria [41,42], Romania, Hungary and the United States [our unpublished findings]. • Limb girdle muscular dystrophy type 2C in France, Spain, Italy, Germany [43], Portugal [91] and Bulgaria [65,66,93]. • Congenital myasthenia in Serbia, Macedonia, Greece, Bohemia and Germany [46]. • Hereditary motor and sensory neuropathy - Russe in Bulgaria [40], Romania and Spain [our unpublished findings]. The existing data are the product of *ad-hoc* collaborative studies and are not likely to represent the true spread of Romani founder mutations. The distribution of LGMD2C in Western Europe and in Bulgaria leads to the prediction that the disorder occurs and awaits detection along the entire European migration route, spanning the Balkans and Central Europe. Filling the gaps in the map will be particularly useful in the case of treatable disorders which are strong candidates for newborn screening, such as galactokinase deficiency and congenital myasthenia.

Table 4: Data on the historical demography of the Roma [ref. 1,75,98]

Source	Period	Region	Population size
Ottoman Empire Tax Registries	15 th century	Balkan Provinces	17,000
	16 th century	Balkan Provinces	65,000
	16 th century	Bulgaria only	5,700
Ottoman Empire Army List	17 th century	Balkan Provinces	11-15,000 males of military age
Austro-Hungarian Empire	1772	Transylvania	39,000
Census Data	1837	Transylvania	53,000
	1893	Transylvania	105,000
	1881-1885	Bulgaria	62,324

Linguistics, history and cultural anthropology suggest two major, equally plausible historical scenarios that could lead to a "jigsaw puzzle" of founder populations: (i) a genetically substructured ancestral population, where the old social traditions of strict endogamy have been retained and subsequent splits of the comprising groups have enhanced the original genetic differences; (ii) a small homogeneous ancestral population spawning numerous subgroups where strong drift effects have resulted in substantial genetic divergence. Genetic research has indeed faced the "jigsaw puzzle" and has thus far been unable to resolve it. The genetic data provide evidence of population stratification, however a closer examination is precluded by the random cross-section sampling design of most population genetic studies, where the traditional social organisation and self-identity of the Roma have been ignored and subjects classified on the basis of the political boundaries of Europe. The relationship between social organisation and genetic structure does not appear to be straightforward and is still to be addressed in population genetic research based on the long standing identity of Gypsy groups. The issue is of relevance to public health policies and the targeted prevention of mendelian disorders, as well as to future studies of genetically complex disorders.

The existing information on single gene disorders is certainly not exclusive to the Roma. The phenomenon of clustering of rare disorders and private founder mutations has been studied in detail in well characterised founder populations, such as the Jews [77,78], Finns [79,80] and French Canadians [81]. Unlike the above examples however, genetic studies of the Roma have failed to take the immediate benefits of research back to the individuals and families that have been the object of research. Yet by now it should be obvious that genetics has an important role to play in improving the quality of health care for the Roma. Treatable disorders such as ga-

lactokinase and MCAD deficiency, with an expected incidence of affected births in the range of 1:1,000 to 1:5,000, meet the standard criteria for newborn screening more than does phenylketonuria, with its average incidence of 1:10,000. Adding the simple, sensitive and specific mutation tests to existing newborn screening programs would be technically simple and highly efficient due to the homogeneous genetic basis of the disorders.

Carrier testing should be made available to Romani communities at high risk for severe untreatable disorders. Information on the identity of affected Romani populations is important for public health intervention since it would allow the planning and facilitate the implementation of targeted prevention programs, especially in the Eastern European countries where economic resources are limited. The importance of the educational component of such programs has already been demonstrated by the highly successful prevention of Tay-Sachs disease among Ashkenazi Jews [82] and the failure of sickle-cell screening among Afro-Americans [83]. This component would be particularly important for a population like the Roma, which has been subject to racism and persecution throughout its co-existence with European societies.

The attention of geneticists is increasingly attracted by genetically isolated populations in the third world. In terms of living standards and the major health indicators, the Roma are much closer to the developing world than to their European neighbours [3]. This forgotten people of Europe can be regarded as a test case for the capacity of genetics to provide better health.

Acknowledgements

We thank the Romani families and communities, and the numerous colleagues in different countries who have made our research into the genetics

of the Roma possible, the members of the Gypsy Genetic Heritage Consortium Prof. J.-C. Kaplan, Prof. A. Urtizberea, Prof. J.-P. Liegeois, Drs. M. Jeanpierre and L. Merlini for their commitment to international collaboration, and Drs. E. Marushiakova, V. Popov and I. Hancock for enlightening discussions of the ethnology, history and linguistics of the Roma. Special thanks to the research team at the Centre for Human Genetics of Edith Cowan University.

L.K. wishes to acknowledge funding from The Wellcome Trust, The National Health and Medical Research Council of Australia, The Muscular Dystrophy Association of the US, L'Association Française contre les Myopathies, The Australian Research Council and Edith Cowan University.

References

- Fraser A: **The Gypsies**. Oxford: Blackwell Publishers, 1992,
- Liegeois J-P: **Roma, Gypsies, Travellers**. Strasbourg: Council of Europe Press, 1994,
- Braham M: **The untouchables: a survey of the Roma people in Central and Eastern Europe**. Geneva: UNHCR, 1993,
- Corretger JM, Fortuny C, Botet F, Valls O: **Marginalidad, grupos étnicos y salud**. *An Esp Pediatr* 1992, **Suppl 48**:115-117
- Binnie GAG: **The health of Gypsies. Problem of caring for travellers is British, not just European**. *BMJ* 1998, **316**:1824-1825
- Hajioff S, McKee M: **The health of the Roma people: a review of the published literature**. *J Epidemiol Community Health* 2000, **54**:864-869
- Reyniers A: **Gypsy populations and their movements within central and eastern Europe and towards some OECD countries**. In *International Migration and Labour Market Policies: Occasional Papers No 1*, Paris 1995,
- Romani east-west migrations: strangers in anybody's land**. *Cambridge Review of International Affairs* 2000, Spring-Summer Issue.
- Fings K, Heuss H, Sparing F: **From "Race Science" to the Camps. The Gypsies during the Second World War**. Hartfield: University of Hertfordshire Press, 1997,
- Avcin M: **Gypsy isolates in Slovenia**. *J Biosoc Sci* 1969, **1**:221-233
- Rex-Kiss B, Szabo L, Szabo S, Hartmann E: **ABO, MN, Rh blood groups, Hp types and Hp level, Gm(I) factor investigations on the Gypsy population of Hungary**. *Hum Biol* 1973, **45**:41-61
- Bartsocas CS, Karayanni C, Tsipouras P, Baibas E, Bouloukos A, Papadatos C: **Genetic structure of the Greek gypsies**. *Clin Genet* 1979, **15**:5-10
- Sivakova D: **Distribution of three red-cell enzyme polymorphisms (ACP, PGMI and AK) in gypsies from Slovakia (Czechoslovakia)**. *Ann Hum Biol* 1983, **10**:449-452
- Tauszik T, Friss A, Gyodi E, Santora Z, Takacs S, Kotvasz A, Toth AM, Horvath M, Tarjan L, Petranyi G, et al: **Genetic polymorphisms of the Gypsy population in Hungary as based on studies of red blood cell antigens**. *Haematologia (Budap.)* 1985, **18**:205-217
- de Pablo R, Vilches C, Moreno ME, Rementeria MC, Solis R, Kreisler M: **Distribution of HLA antigens in Spanish Gypsies: a comparative study**. *Tissue Antigens* 1992, **40**:187-196
- Mastana SS, Papiha SS: **Origin of Romany Gypsies - genetic evidence**. *Z Morphol Anthropol* 1992, **79**:43-51
- Forrai G, Tauszik T, Auszik N, Moh T, Tunyog M, Holics C, Bankovi G, Gal I: **A high incidence of PKD in a large geographic area of south-western Hungary: A medical genetic study**. In *Genetics of Kidney Disorders*. Edited by Bartsocas C. New York: Alan R. Liss, Inc., 1989,
- Bernasovsky I, Suchy J, Bernasovska K, Vargova T: **Blood groups of Roms (Gypsies) in Czechoslovakia**. *Am J Phys Anthropol* 1976, **45**:277-280
- Sivakova D, Sieglava Z, Lubyova B, Novakova J: **A genetic profile of a Romany (Gypsy) subethnic group from a single region in Slovakia**. *Gene Geogr* 1994, **8**:109-116
- Guglielmino CR, Beres J: **Genetic structure in relation to the history of hungarian ethnic groups**. *Hum Biol* 1996, **68**:335-355
- Gyodi E, Tauszik T, Petranyi G, Kotvasz A, Palaffy G, Takacs, Nemark P, Hollan SR: **The HLA antigen distribution in the Gypsy population in Hungary**. *Tissue Antigens* 1981, **18**:1-12
- Harper PS, Williams EM, Sunderland E: **Genetic markers in Welsh Gypsies**. *J Med Genet* 1977, **14**:177-182
- Clarke VA: **Genetic factors in some British Gypsies**. In *Genetic Variation in Britain*. Edited by Roberts DF, Sunderland E. London: Taylor and Francis, 1973,
- Beckman L, Takman J: **On the anthropology of a Swedish Gypsy population**. *Hereditas* 1965, **53**:272-280
- Galikova J, Vilimova M, Ferak V, Mayerova A: **Haptoglobin types in Gypsies from Slovakia (Czechoslovakia)**. *Hum Hered* 1969, **19**:480-485
- Bernasovsky I, Halko N, Biros I, Sivakova D, Jurickova J: **Some genetic markers in Valachian (Olachian) Gypsies in Slovakia**. *Gene Geogr* 1994, **8**:99-107
- Hocevar M: **Die Verteilung Blutgruppen bei einem Zigeunerisolat**. In *Proceedings of the 10th Congress of the International Society on Blood Transfusion, Stockholm, 1965*, ii:312-319
- Mourant AE, Kopec AC, Domaniewska-Sobczak K: **The distribution of the human blood groups and other polymorphisms**. London: Oxford University Press, 1976,
- Sanger R, Race RR: **The combination of blood groups in a sample of 250 people**. *Annals of Eugenics* 1950, **15**:77-90
- Fisher RA, Race RR: **Rh gene frequencies in Britain**. *Nature* 1946, **157**:48-49
- Beckman LA, Takman H, Arfords KE: **Distribution of blood and serum groups in a Swedish gypsy population**. *Acta Genetica (Basel)* 1965, **15**:134-139
- Smars G, Beckman L, Book JA: **Osteogenesis imperfecta and blood groups**. *Acta Genet Stat Med* 1961, **11**:133-136
- Watkin IM: **The Welsh element in the South Wales coalfield**. *Journal of the Royal Anthropological Institute* 1965, **95**:104-114
- Tiwari SC, Bhasin MK: **The blood groups of the Brahmins and Rajputs of Garwal**. *Hum Biol* 1968, **40**:386-395
- Cavalli-Sforza LL, Menozzi P, Piazza A: **History and geography of human genes**. Princeton: Princeton University Press, 1994,
- Papiha SS, Roberts DF, Wig NN, Singh S: **Red cell enzyme polymorphisms in Punjabis in north India**. *Am J Phys Anthropol* 1972, **37**:293-299
- Kalaydjieva L, Hallmayer J, Chandler D, Savov A, Nikolova A, Angelicheva D, King RHM, Ishpekova B, Honeyman K, Calafell F, et al: **Gene mapping in Gypsies identifies a novel demyelinating neuropathy on chromosome 8q24**. *Nat Genet* 1996, **14**:214-217
- Kalaydjieva L, Nikolova A, Tournev I, Petrova J, Hristova A, Ishpekova B, Petkova I, Shmarov A, Stancheva S, Middleton L, et al: **Hereditary motor and sensory neuropathy - Lom, a novel demyelinating neuropathy associated with deafness in Gypsies. Clinical, electrophysiological and nerve biopsy findings**. *Brain* 1998, **121**:399-408
- Kalaydjieva L, Gresham D, Gooding R, Heather L, Baas F, de Jonge R, Blechschmidt K, Angelicheva D, Chandler D, Worsley P, et al: **N-myc downstream regulated gene 1 is mutated in hereditary motor and sensory neuropathy - Lom**. *Am J Hum Genet* 2000, **67**:47-58
- Rogers T, Chandler D, Angelicheva D, Thomas PK, Youl B, Tournev I, Gergelcheva V, Kalaydjieva L: **A novel locus for autosomal recessive peripheral neuropathy in the EGR2 region on 10q23**. *Am J Hum Genet* 2000, **67**:664-671
- Tournev I, Kalaydjieva L, Youl B, Ishpekova B, Guerguelcheva V, Kamenov O, Katarzova M, Kamenov Z, King RHM, Romanski K, et al: **Congenital cataracts facial dysmorphism neuropathy syndrome, a novel complex genetic disease in Balkan Gypsies: clinical and electrophysiological observations**. *Ann Neurol* 1999, **45**:742-750
- Angelicheva D, Turnev I, Dye D, Chandler D, Thomas PK, Kalaydjieva L: **Congenital cataracts facial dysmorphism neuropathy syndrome: a novel developmental disorder in Gypsies maps to 18qter**. *Eur J Hum Genet* 1999, **7**:560-566
- Piccolo F, Jeanpierre M, Leturcq F, Dode C, Azibi K, Toutain A, Merlini L, Jarre L, Navarro C, Krishnamoorthy R, et al: **A founder mutation in the γ -sarcoglycan gene of Gypsies possibly predating their migration out of India**. *Hum Mol Genet* 1996, **5**:2019-2022
- Kalaydjieva L, Perez-Lezaun A, Angelicheva D, Onengut S, Dye D, Bosshard NU, Jordanova A, Savov A, Yanakiev P, Kremensky I, et al: **A founder mutation in the GK1 gene is responsible for galactokinase deficiency in Roma (Gypsies)** *Am J Hum Genet* 1999, **65**:1299-1307
- Plasilová M, Stoilov I, Sarfarazi M, Kadasi L, Feráková E, Ferák V: **Identification of a single ancestral CYP1B1 mutation in Slovak Gypsies (Roms) affected with primary congenital glaucoma**. *J Med Genet* 1999, **36**:290-294

46. Abicht A, Stucka R, Karcagi V, Vherczegfalvi A, Horváth R, Mortier W, Schara U, Ramaekers V, Jost W, Brunner J, et al: **A common mutation 1267delG in congenital myasthenic patients of Gypsy ethnic origin.** *Neurology* 1999, **53**:1564-1569
47. Martinez-Frias ML: **Analisis del riesgo que para defectos congenitos tienen diferentes grupos etnicos de nuestro pais.** *Am Esp Pediatr* 1998, **48**:395-400
48. Gimenez-Roldan S, Delgado G, Marin M, Villanueva A, Mateo D: **Hereditary torsion dystonia in Gypsies.** In *Advances in Neurology* 1988, **50**:73-81
49. Martinez-Frias ML, Bermejo E: **Prevalence of congenital anomaly syndromes in a Spanish Gypsy population.** *J Med Genet* 1992, **29**:483-486
50. Angelicheva D, Calafell F, Savov A, Jordanova A, Kufardjieva A, Galeva I, Nedkova V, Ivanova T, Yankova P, Konstantinova D, et al: **Cystic fibrosis mutations and associated haplotypes in Bulgaria: a comparative population genetic study.** *Hum Genet* 1997, **99**:513-520
51. Kalanin J, Takarada Y, Kagawa S, Yamashita K, Ohtsuka N, Matsuoka A: **Gypsy phenylketonuria: a point mutation of the phenylalanine hydroxylase gene in Gypsy families from Slovakia.** *Am J Med Genet* 1994, **49**:235-239
52. Desviat LR, Perez B, Ugarte M: **Phenylketonuria in Spanish Gypsies: prevalence of the IVS10nt546 mutation on haplotype 34.** *Hum Mutat* 1997, **9**:66-68
53. Martinez G, Garcia-Lozano JR, Ribes A, Maldonado MD, Baldellou A, de Pablo R, Nunez-Roldan A: **High risk of medium chain acyl-CoA dehydrogenase deficiency among Gypsies.** *Pediatr Res* 1998, **44**:83-84
54. Plasilová M, Feráková E, Kádasi L, Poláková E, Gerinec A, Ott J, Ferák V: **Linkage of autosomal recessive primary congenital glaucoma to the GLC3A locus in Roms (Gypsies) from Slovakia.** *Hum Hered* 1998, **48**:30-33
55. Chandler D, Angelicheva D, Heather L, Gooding R, Gresham D, Yanakiev P, de Jonge R, Baas F, Dye D, Karagoyozov L, et al: **Hereditary motor and sensory neuropathy - Lom: Refined genetic mapping in Romani (Gypsy) families from several European countries.** *Neuromuscul Disord* 2000, **10**:584-591
56. Gitzelmann R, Hansen RG: **Galactose metabolism, hereditary defects and their clinical significance.** In *Inherited Disorders of Carbohydrate Metabolism.* Edited by Burman D, Holton JB, Pennock CA. Lancaster: MTP, 1980,
57. Levy HL: **Screening for galactosemia.** In *Inherited Disorders of Carbohydrate Metabolism.* Edited by Burman D, Holton JB, Pennock CA. Lancaster: MTP, 1980,
58. Dalggaard OZ: **Bilateral polycystic disease of the kidneys: a follow-up of two hundred and eighty four patients and their families.** *Acta Med Scand* 1957, **328**:1-255
59. d'Epainy SL, Remé CH: **Ausgewählte Aspekte des kongenitalen Glaukoms.** *Klin Mbl Augenheilk* 1977, **170**:249-259
60. Travers JP: **The presentation of congenital glaucoma.** *Br J Ophthalmol* 1979, **56**:241-242
61. Gencikova A, Gencik A: **Congenital glaucoma in Gypsies from Slovakia.** *Hum Hered* 1982, **32**:270-273
62. Gencik A, Gencikova A, Ferak V: **Population genetical aspects of primary congenital glaucoma. I. Incidence, prevalence, gene frequency and age of onset.** *Hum Genet* 1982, **61**:193-197
63. Levy JM, Mayer G, Sacrez R, Ruff R, Francfort J-J, Rodier L: **Thrombasthénie de Glanzmann-Naegeli. Etude d'un groupe ethnique à forte endogamie.** *Ann Pediatr (Paris)* 1971, **18**:381-389
64. de la Salle C, Schwartz A, Baas M-J, Lanza F, Cazenave J-P: **Detection by PCR and HphI restriction analysis of a splice site mutation at the 5' end of intron 15 of the platelet GPIIb (IIb integrin) gene responsible for Glanzmann's thrombasthenia type I in Gypsies originating from the Strasbourg area.** *Thromb Haemost* 1995, **74**:990-991
65. Tournev I, Aneva L, Kamenov O, Ishpekova B, Katzarova V, Gerguelcheva V, Angelicheva D, Kalaydjieva L: **Gamma-sarcoglycan deficiency in Bulgarian Gypsies.** *Muscle Nerve* 1998, **Suppl 7**:136-136
66. Gresham D, Tournev I, Angelicheva D, Avena L, Kamenov O, Jeanpierre M-P, Kalaydjieva L: **Limb-girdle muscular dystrophy in a Xoroxane Roma population.** *European Research Conferences: Inherited Disorders and Their Genes in Different European Populations, Obernai, France, 1999,*
67. Merlini L, Kaplan J-C, Navarro C, Barois A, Bonneau D, Brasa J, Echenne B, Gallano P, Jarre L, Jeanpierr M, et al: **Homogeneous phenotype of the gypsy limb-girdle MD with the gamma-sarcoglycan C283Y mutation.** *Neurology* 2000, **54**:1075-1079
68. Merlini L, Villanova M, Sabatelli P, Trogu A, Malandrini A, Yanakiev P, Maraldi NM, Kalaydjieva L: **Hereditary motor and sensory neuropathy Lom type in an Italian Gypsy family.** *Neuromuscul Disord* 1998, **8**:182-185
69. Butinar D, Zidar J, Leonardis L, Popovic M, Kalaydjieva L, Angelicheva D, Sininger Y, Keats B, Starr A: **Hereditary auditory, vestibular, motor and sensory neuropathy in a Slovenian Roma (Gypsy) kindred.** *Ann Neurol* 1999, **46**:36-44
70. Baethmann M, Gohlich-Ratmann G, Schröder JM, Kalaydjieva L, Voit T: **HMSNL in a 13-year-old Bulgarian girl.** *Neuromuscul Disord* 1998, **8**:90-94
71. Colomer J, Iturriaga J, Kalaydjieva L, Angelicheva D, King RHM, Thomas PK: **Hereditary motor and sensory neuropathy LOM (HMSNL) in a Spanish family: clinical, electrophysiological, pathological and genetic studies.** *Neuromuscul Disord* 2000, **10**:578-583
72. Sampson J: **Notes on Professor R.L.Turner's "The position of Romani in Indo-Aryan".** *Journal of the Gypsy Lore Society* 1927, **6**:57-68
73. Turner RL: **The position of Romani in Indo-Aryan.** *Journal of the Gypsy Lore Society* 1926, **5**:145-189
74. Hancock I: **The emergence of Romani as a koine outside of India.** In *Scholarship and the Gypsy Struggle: Commitment in Romani Studies.* Edited by Acton TA. Hatfield: University of Hertfordshire Press, 2000,
75. Marushiakova E, Popov V: **Gypsies (Roma) in Bulgaria.** In *Studien zur Tsiganologie und Folkloristik.* Frankfurt am Main: Peter Lang, 1997,
76. Hancock I: **The pariah syndrome.** *Ann Arbor: Karoma Publishers Inc., 1987,*
77. Goodman RM: **Genetic disorders among the Jewish people.** *Baltimore: John Hopkins University Press, 1978,*
78. Motulsky AG: **Jewish diseases and origins** *Nat Genet.* 1995, **9**:99-101
79. de la Chapelle A: **Disease gene mapping in isolated human populations: the example of Finland.** *J Med Genet* 1993, **30**:857-865
80. Peltonen L, Jalanko A, Varilo T: **Molecular genetics of the Finnish disease heritage.** *Hum Mol Genet* 1999, **8**:1913-1923
81. Heyer E: **One founder/one gene hypothesis in a new expanding population: Saguenay (Quebec, Canada).** *Hum Biol* 1999, **71**:99-109
82. Kaback M, Lim-Steele J, Dabholkar D, Brown D, Levy N, Zeiger K: **Tay-Sachs disease- carrier screening, prenatal diagnosis, and the molecular era. An international perspective, 1970 to 1993. The International TSD Data Collection Network.** *JAMA* 1993, **270**:2307-2315
83. Markel H: **The stigma of disease: implications of genetic screening.** *Am J Med* 1992, **93**:209-215
84. Reynolds J, Weir BS, Cockerham CC: **Estimation of the coancestry coefficient: basis for a short term genetic distance.** *Genetics* 1983, **105**:767-779
85. Saitou N, Nei M: **The neighbor-joining method: a new method for reconstructing phylogenetic trees.** *Mol Biol Evol* 1987, **4**:406-425
86. Felsenstein J: **Confidence limits on phylogenies: an approach using the bootstrap.** *Evolution* 1985, **39**:783-791
87. Felsenstein J: **PHYLIP - Phylogeny Inference Package (Version 3.2).** *Cladistics* 1989, **5**:164-166
88. Excoffier L, Smouse PE, Quattro JM: **Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data.** *Genetics* 1992, **131**:479-491
89. Schneider S, Kueffer J-M, Excoffier L: **Arlequin ver 1.1: A software for population genetic data analysis.** *Genetics and Biometry Laboratory, University of Geneva, Switzerland, 1997,*
90. Velthuisen B, Saris JJ, de Haij S, Hayashi T, Reynolds DM, Mochizuki T, Elles R, Fossdal R, Bogdanova N, van Dijk MA, et al: **A spectrum of mutations in the second gene for autosomal dominant polycystic kidney disease (PKD2).** *Am J Hum Genet* 1997, **61**:574-555
91. Lasa A, Piccolo F, De Diego C, Jeanpierre M, Colomer J, Rodriguez MJ, Urtizberea JA, Baiget M, Kaplan J-C, Gallano P: **Severe limb girdle muscular dystrophy in Spanish Gypsies: further evidence for a founder mutation in the gamma-sarcoglycan gene.** *Eur J Hum Genet* 1998, **6**:396-399

92. Schlegel N, Gayet O, Morel-Kopp MC, Wyler B, Hurtaud-Roux MF, Kaplan C, MacGregor J: **The molecular basis of Glanzmann thrombasthenia (GT) in a Gypsy population in France. Identification of a new mutation of the IIb gene.** *Blood* 1994, **84**:477a-
93. Todorova A, Ashikov A, Beltcheva O, Tournev I, Kremensky I: **C283Y mutation and other C-terminal nucleotide changes in the γ -sarcoglycan gene in the Bulgarian Gypsy population.** *Hum Mutat* 1999, **14**:40-44
94. Blehova J, Daneslova J, Grec L, Hajeck F, Matousek M, Vojtik V: **Vyskyt fenylketonurie cechach a na Morave.** *Cheskoslovenska Pediatrie* 1959, **14**:498-503
95. Thalhammer O, Gitzelmann R, Pantlischko M: **Hypergalactosemia and galactosuria due to galactokinase deficiency in a newborn.** *Pediatrics* 1968, **42**:441-445
96. Gitzelmann R: **Hereditary galactokinase deficiency.** *Citation Classics Current Contents* 1987, **30**:14-
97. Bolgiani MP, Gallenca M, Barocelli PC: **Su un caso di galattosemia da deficit di galattochinasi.** *Pediat Med Chir* 1984, **6**:333-336
98. Achim V: **Tigani in Istoria Romaniei.** In *Colectia "Biblioteca enciclopedica de istorie a Romaniei"*. Bucuresti: Editura Enciclopedica, 1998,

Pre-publication history

The pre-publication history for this paper can be accessed here:

<http://www.biomedcentral.com/content/backmatter/1471-2350-2-5-b1.pdf>

Publish with **BioMedcentral** and every scientist can read your work free of charge

"BioMedcentral will be the most significant development for disseminating the results of biomedical research in our lifetime."

Paul Nurse, Director-General, Imperial Cancer Research Fund

Publish with **BMC** and your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours - you keep the copyright



BioMedcentral.com

Submit your manuscript here:

<http://www.biomedcentral.com/manuscript/>

editorial@biomedcentral.com