ARTICLE

Founder effect and estimation of the age of the French Gypsy mutation associated with Glanzmann thrombasthenia in Manouche families

Mathieu Fiore*,[1], Xavier Pillois[1], Paquita Nurden[1], Alan T Nurden[1] and Frédéric Austerlitz[2]

The c.1544+1G>A substitution at the 5′ splice donor site of intron 15 of the *ITGA2B* gene, called the French Gypsy mutation, causes Glanzmann thrombasthenia, an inherited hemorrhagic disorder transmitted as an autosomal recessive trait and characterized by an altered synthesis of the platelet αIIbβ3 integrin. So far, this mutation has only been found in affected individuals originating from French Manouche families, strongly suggesting a founder effect. Our goal was to investigate the origin of the French Gypsy mutation. We estimated the age of the mutation by a likelihood-based method that uses the length of the shared haplotypes among a set of patients. For this, we genotyped 23 individuals of Manouche origin; consisting of 9 Glanzmann thrombasthenia patients homozygous for the French Gypsy mutation, 6 heterozygous carriers and 8 homozygous wild-type individuals. They were genotyped for four single-nucleotide polymorphisms using high-resolution melting curve analysis, and for two CA repeats in the *BRCA1* and *THRA* genes at chromosome 17, using fragment analysis gels. We found that a haplotype of five polymorphic loci covering a 4-cM region was strongly associated with the French Gypsy mutation, suggesting a founder effect. The estimated age of this founder mutation was 300–400 years (range 255–552 years). Thus, all carriers of the French Gypsy mutation c.1544+1G>A at intron 15 descended from a common ancestor 300–400 years ago.

INTRODUCTION

Glanzmann thrombasthenia (GT) is a rare inherited bleeding disorder because of a lack or a qualitative deficiency of the αIIbβ3 integrin expressed at the platelet surface.[1] The *ITGA2B* gene encodes for the αIIb subunit while the *ITGB3* gene encodes for β3; they are both located on chromosome 17q21–23. As GT is an autosomal recessive disease, in non-consanguineous populations, patients are often compound heterozygotes. More than 150 mutations causing GT have already been reported. They are distributed on almost all the exons of *ITGA2B* and *ITGB3* genes. An Internet database that includes clinical and biological data, as well as mutation analyses on reported patients is regularly updated (http://med.mssm.edu/glanzmanndb).

GT occurs in high frequency in certain ethnic populations with an increased incidence of consanguinity, such as Iraqi Jews or Jordanian Arabs.[2,3] The unusual high frequency of GT among these groups is largely the result of founder mutations. Specific gene mutations involved in rare disorders and of private founder mutations have been studied in detail in well-characterized founder populations, such as Israeli Jews, Finns and French Canadians.[2,4,5]

In France, GT is particularly frequent in the Gypsy population, mainly represented by the Manouche tribe.[6,7] Their clinical phenotype is the same as for classic type I disease, with a moderate and occasional severe bleeding syndrome, absent platelet aggregation and absent clot retraction. Flow cytometry analysis has shown that their platelet surface expresses <5% of normal αIIbβ3 levels, while western blotting showed no αIIb and only trace amounts of β3. They are also susceptible to develop platelet iso-antibodies after platelet transfusions.[8,9]

Two groups independently localized the disease-causing mutation: a homozygous c.1544+1G>A substitution at the splice donor site of intron 15 of the *ITGA2B* gene (Figure 1).[10,11] Interestingly, this mutation was detected in all GT patients coming from the Manouche community and originating from different regions of France, the so-called French Gypsy mutation.[12,13] Moreover, this mutation appears limited to this tribe, which led us to hypothesize the existence of a founder mechanism in this population.[14–16]

In this study, we examined whether GT patients within different families originating from the Manouche tribe and carrying the French Gypsy mutation share a common ancestor, and if so, to estimate the age of the founding event.

MATERIALS AND METHODS
Sample selected for haplotyping
Our study sample was composed of 23 individuals from 16 families originating from the French Manouche community, with 9 affected, 6 heterozygous carriers and 8 normal individuals. The male to female ratio was 0.43 (10 males and 13 females). These patients, originating from different regions, were registered in the French Reference Centre for Platelet Diseases. To the best of our knowledge, all families were unrelated. DNA was isolated using a standard protocol from peripheral blood using the QIAamp DNA Blood Mini Kit

[1]Centre de Référence des Pathologies Plaquettaires, Plateforme Technologique d'Innovation Biomédicale, Hôpital Xavier Arnozan, Pessac, France; [2]Laboratoire d'Eco-Anthropologie et Ethnobiologie, UMR 7206 CNRS/MNHN/Universite Paris 7, Museum National d'Histoire Naturelle, CP 139, 57 rue Cuvier, F-75231 Paris Cedex 05, France
*Correspondence: Dr M Fiore, Centre de Référence des Pathologies Plaquettaires, Plateforme Technologique d'Innovation Biomédicale, Hôpital Xavier Arnozan, 33604 Pessac, France. Tel: +33 55 710 2850; Fax: +33 55 710 2864; E-mail: mathieu.fiore@chu-bordeaux.fr
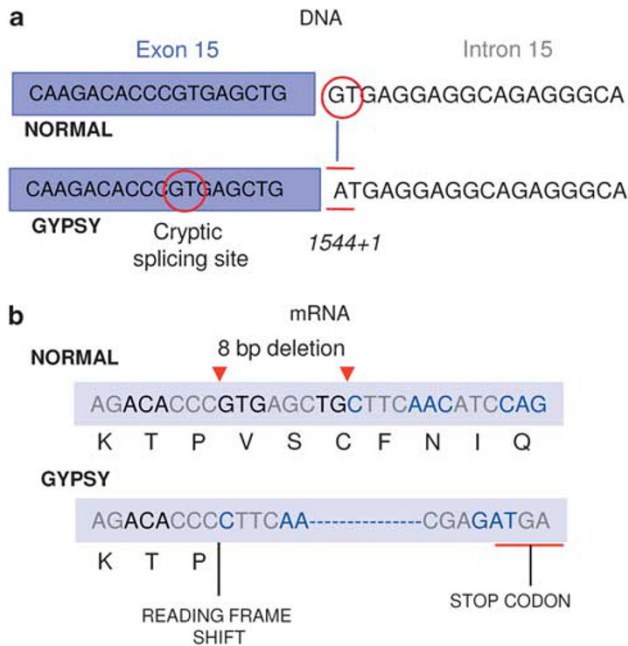
Figure 1 c.1544+1G>A mutation affects the splicing and expression of αIIb. (**a**) Schematic representation of the exon 15/intron 15 boundary of the *ITGA2B* gene showing the position of the c.1544+1G>A mutation (marked by red lines); (**b**) mRNA sequence showing the splice products. Eight base pair deletion causes a frameshift, which results in premature stop codon.

(Qiagen, Courtaboeuf, France). These individuals were previously tested for the presence of the French Gypsy mutation in *ITGA2B* intron 15 by an allelic discrimination assay with high-resolution melting (HRM) curve analysis.[17] Parents and/or siblings of patients were genotyped and the correct segregation of the mutation in the pedigree was verified, when DNA was available. All patients signed the informed consent for genetic analysis before inclusion, and national ethics authorization for DNA analysis was obtained under the promotion of the French National Institute of Health and Medical Research (INSERM : ANR-08-GENO-028–03).

Haplotype analysis

Four single-nucleotide polymorphisms (SNPs) were selected for the haplotype reconstruction: rs5910, rs5918, rs5919 and rs11870252. These SNPs were chosen based on their chromosomal position and on their allelic frequencies, according to data available from the GenBank database (National Center for Biotechnology Information). They are located on chromosome 17, in the *ITGA2B* and *ITGB3* genes, outlining a genomic fragment of approximately 3 cM from rs11870252 to rs5910 (Table 1). SNPs rs5918, rs5919 and rs11870252 are respectively located in exons 3, 6 and intron 11 of the *ITGB3* gene. On the other hand, SNP rs5910 is located in exon 30 of the *ITGA2B* gene. The distance from rs5910 to the French Gypsy mutation is about 8000 bp. Two short tandem repeats (STRs) flanking the *ITGA2B* gene on chromosome 17 were also selected in *THRA* and *BRCA1* genes.[18] These STRs consist of tandemly repeated dinucleotide (CA) sequences. Alleles were named using an arbitrary scale for the observed fragment length of the microsatellites.

Genotyping

SNP genotyping was performed by HRM (see Supplementary Table). In brief, 20–60 ng of DNA were amplified in a Light Cycler 480 system (Roche Diagnostics, Meylan, France) in a final volume of 20 μl with 2 mM MgCl₂, 0.2 μM of each primer, and 10 μl of PCR mix containing LC Green (Roche Diagnostics). Amplification was performed by 45 cycles of 95 °C for 10 s, annealing for 15 s (see Supplementary Table) and 72 °C for 15 s followed by a melt from 65 to 95 °C at 4 °C/s. Microsatellites were amplified using fluorescently labelled forward primers in a 25 μl reaction mix, containing 50 ng of

genomic DNA, 1.25 units of Taq polymerase (Applied Biosystems, Courtaboeuf, France), 2.5 mM of MgCl₂, 400 nM of each primer, and 50 μM of each dNTP. After 7 min at 95 °C, 40 cycles of amplification (95 °C for 30 s; 54 °C for 30 s; 72 °C for 40 s) were performed, followed by 5 min at 72 °C. PCR products were loaded onto a Beckman Coulter CEQ 8000 sequencer with the Beckman Coulter DNA Size Kit 400 (Beckman Coulter, Roissy, France).

Statistical analysis

The age of the most recent common ancestor of the 15 carriers of the mutant allele was estimated using a likelihood-based method.[19] This maximum-likelihood algorithm requires first a reconstruction of the haplotypes of the carriers of the mutant allele for the surrounding neutral markers. It uses as input the percentage of recombinant haplotypes, as well as the carrier frequency of the mutant allele in the population. It then estimates the number of generations (with 95% CI) elapsed since the most recent common ancestor introduced the mutation into the population and the growth rate of this mutation since then. As this method requires reconstructing the haplotypes carried by the patients, we genotyped individuals related to these patients (when available) to determine the phase. A number of GT patients could not be allocated to any pedigree. We used PHASE v2.1 software (University of Washington, Seattle, WA, USA) to infer the haplotypes for these isolated patients.[20] This method is based on the use of a maximum likelihood to estimate haplotype frequencies and to infer haplotype pairs. Finally, distances in base pairs were transformed into distances in cM by using the correspondence between length in base pairs and in cM provided by the International HapMap project.[21]

The current number of copies of the mutant allele in the study population can be computed from the population size and the allelic frequency of the mutant allele. The basic characteristic of the Gypsies is their nomadic way of life, which makes it difficult to estimate accurately their population sizes. In France, statistical institutions evaluate the number of travelling people as approximately 250 000 individuals, belonging in majority to the Gypsy community.[20,22] There is little data on the size of the French Manouche population, which despite fluctuations has remained small. As the Manouche tribe is a major contributor to the French Gypsies we estimate their number as between 80 000 and 120 000; we have therefore used a total value of 100 000 individuals. Although we did not know the effective population size, we also considered two empirical effective values of 10 000 and 25 000 individuals.

Regarding the allelic frequency of the mutant allele, its estimation requires an unbiased sample of the population. However, patients who refer to Hospital Departments do so because they require treatment, thus introducing a recruitment bias. Mutant allele frequencies reported by other studies in similar founder populations mostly range from 0.002 to 0.05.[19] Thus, we assumed here that the mutant allele frequency in our population ranged between 0.001 and 0.01.

RESULTS

The obtained genotypes are shown in Figure 2 and Table 2. It was not possible to accurately determine the haplotype associated with the French Gypsy mutation in patients M_{11} and M_{15} because of limited information from family members. The haplotype reconstruction with PHASE was ambiguous for these two patients and we discarded them from further analyses.

For the remaining individuals, 10 chromosomes positive for the c.1544+1G>A transversion in 6 different families shared a common haplotype for all SNPs and for the 2 microsatellite repeats (haplotype H1=5, 3, C, C, T, T, respectively, for *THRA/BRCA1* STRs and SNPs at nucleotides c.3063C>T, c.176T>C, c.882T>C and c.1914–19T>C) (Table 3). The STR in the *BRCA1* gene of one patient (M_5) showed a different allele than haplotype H1 (haplotype H2=5, 9, C, C, T, T), possibly corresponding to a single recombination event. Three additional individuals differed at the microsatellite repeat located in the *THRA* gene, suggesting another crossing-over event (haplotype H3=7, 3, C, C, T, T) (Table 3). The frequency of haplotypes H1, H2 and H3 in these French Gypsy mutation carriers were respectively 68% (15/22 chromosomes), 5% (1/22) and 27% (6/22) (Table 3).

Table 1 French Gypsy mutation-linked microsatellite and SNP genotypes

Gene	Nucleotide variations	SNP/STR marker	Distance to c.1544+1G>A (bp)	Allele	Frequency[a]		
					European	Asian	African
THRA	(CA)n repeats	—	(−) 4 227 753	—	—	—	—
BRCA1	(CA)n repeats	—	(−) 1 257 753	—	—	—	—
ITGA2B	c.3063C>T	rs5910	(−) 7964	C	0.698	0.533	0.559
	c.1544+1G>A	French Gypsy mutation	0	—	—	—	—
ITGB3	c.176T>C	rs5918	(+) 2 902 977	T	0.850	1.000	0.867
	c.882T>C	rs5919	(+) 2 906 787	T	0.958	0.756	0.833
	c.1914-19T>C	rs11870252	(+) 2 920 072	T	0.938	0.750	0.870

[a]The allele frequency is given for each SNP and was based on the dbSNP 130.

To assess whether these strong associations of specific haplotypes were indeed the result of a founder effect, eight related and unrelated healthy individuals not carrying the French Gypsy mutation (M_3, M_4, M_{12}, M_{13}, M_{17}, M_{19}, M_{20} and M_{22}) and coming from the Manouche community were genotyped for the whole set of SNP loci. The most common allele in these individuals for the SNP rs5918 was the T allele (15/16: 94 %), which is never associated with c.1544+1G>A among the disease chromosomes (Table 3). Moreover, the only healthy individual (M_{19}) carrying the C allele for rs5918 did not carry the C allele found for all patients at the rs5910 locus. Thus, the haplotypes H1, H2 and H3 were absent in our sample of healthy individuals, which clearly indicates that these haplotypes are specifically linked to the French Gypsy mutation.

So, all patients but one shared a smaller unique homozygous haplotype (3, C, C, T, T) in the region made by the following five markers: BRCA1, rs5910, rs5918, rs5919 and rs11870252. Thus, this defines a common ancestral core haplotype shared by almost all patients. This strong linkage disequilibrium with the French Gypsy mutation in ITGA2B is a strong indication of a founder effect. Globally our results suggest that the mutation in the Manouche families is located on a founder chromosome and that crossing-over event that occurred at BRCA1 and THRA resulted in the apparition of haplotypes H2 and H3. All this suggests that the mutation was transmitted by a relatively recent common ancestor.

The method to estimate the age of the most recent common ancestor, which considers the linkage disequilibrium across all markers, is implemented in a specific software.[19] We provided the program with an estimate of 100 000 for the number of individuals currently present in the Manouche community. Three separate analyses were performed, each using a different disease allele frequency. Only markers situated on the left side of the c.1544+1G>A could be used, as the markers situated on the right side were completely linked with the disease gene. Using these left-side markers, we estimated an age of the mutation ranging from 12.1 to 15.3 generations (Table 4), associated with a growth rate ranging from 1.33 to 2.2, depending on the assumed disease allele frequency. Assuming that a generation spans 25 years, the most recent common ancestor was estimated to have lived about 300 to 400 years ago (95% confidence interval; 255 to 552.5 years). Also, considering an empirical effective size of 10 000 and 25 000 individuals, we obtained older estimates for the age, but the differences were quite small (Table 4).

DISCUSSION

We report here a large group of patients analyzed for the French Gypsy mutation. This mutation in a homozygous state results in the absence of $\alpha IIb\beta 3$ protein because of an abnormal splicing site, a frameshift and a premature stop codon in the ITGA2B gene.[10,11] Clusters of GT families in the French Manouche community and a relative over-representation of the disease-causing mutation are indicative of a founder effect. This hypothesis was further supported by the report of nine homozygous patients for the French Gypsy mutation who were tested for the human platelet antigen-1 a/b (HPA), a SNP in the ITGB3 gene.[14–16] Jacquelin et al[14] noted that the French Gypsy mutation was strongly linked to HPA-1b, the minority allele in a normal population, suggesting that these patients carry a homozygous common ancestral haplotype at the ITGA2B locus. Furthermore, among several hundred GT patients who have been genotyped worldwide, to our knowledge, none has ever been found to carry this mutation, which rules out the hypothesis that it is caused by a mutation 'hotspot'.

Genotyping a set of four SNPs and two polymorphic STR markers near the French Gypsy mutation locus allowed us to characterize three different haplotypes in eight unrelated patients, from which a minimal ancestral common core haplotype was defined at markers BRCA1, rs5910, rs5918, rs5919 and rs11870252, spanning about 4 Mb. The large size of this core haplotype is a clear indication of a recent founder effect.

By applying the method of a maximum likelihood to the microsatellite haplotype data, we were able to calculate the age of the common ancestor of all carriers of the mutation. This estimated age of the French Gypsy mutation ranged from 12.1 to 15.3 generations (302.5–382.5 years, assuming a 25-year generation time). The study would probably have a better accuracy with a higher sample size for the French Gypsy mutation carriers, but we were limited in part by the fact that GT is a rare disease. Nevertheless, confidence intervals are narrow enough so that our conclusions are reliable, and our sample of 15 phased chromosomes from unrelated patients is of the same order of magnitude as in other studies.[19]

Founder effects have also been previously identified for GT in 11 unrelated Palestinian patients carrying the same 13-bp deletion in the ITGA2B gene, and in 6 Jordanian families carrying a C549R mutation in $\beta 3$.[23,24] The Jordanian C549R mutation probably occurred more recently because the haplotype associated with the C549R mutation was conserved over a distance of 4.5 Mb, while the common haplotype associated with the Palestinian 13-bp deletion was only of 350 kb. Analyzing the closest loci showing recombination events with the mutations yielded an age estimate of 300–600 years for the 13-bp deletion in Palestinian patients and 120–150 years for the C549R in Jordanian ones.

Our study illustrates how recent developments in population genetics allow insights into demographic and population history. Records on linguistics, cultural anthropology and demography
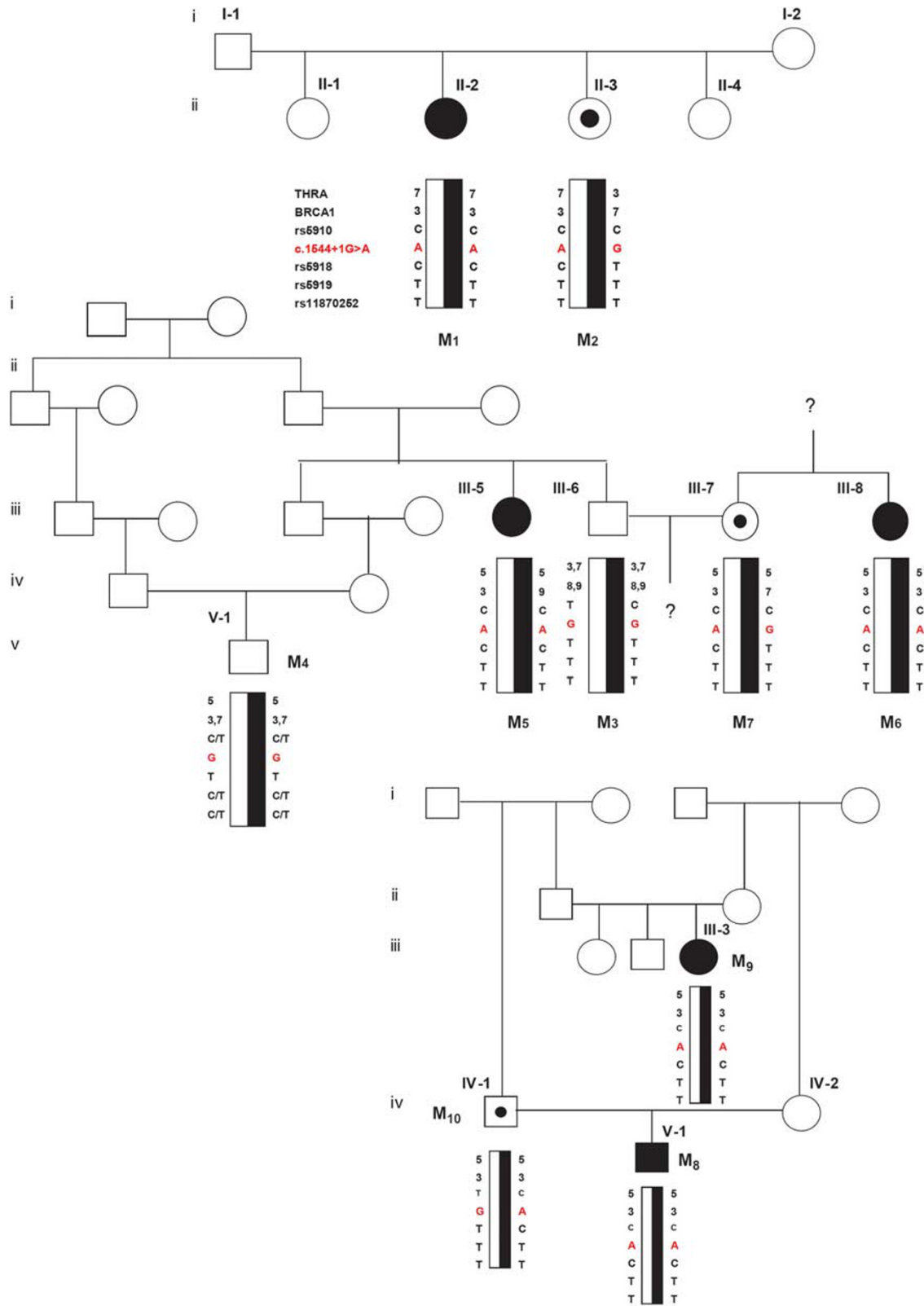
Figure 2 Pedigrees of studied cases of GT. Homozygous GT subjects are represented by solid squares or circles; heterozygous individuals by solid squares or circles within open frames. The bars represent the haplotype constructed from four SNPs and two STRs around the French Gypsy mutation of the *ITGA2B* gene.

describe the Gypsies as a population with Indian origins, with an initial exodus into Byzantium Empire dated approximately to the eleventh century, and a subsequent dispersal throughout Europe by the fifteenth century (Figure 3).[25–29] The migration of Gypsies in France took two routes: the first, through the North of Europe, and the second from Persia through North-Africa. Gypsies recognize

Table 2 STRs and SNPs genotypes for the 23 individuals (M_x) from the 16 Manouche families carrying or not the c.1544+1G > A mutation in ITGA2B gene

					Locus and nucleotide position (bp)			
		THRA[a]	BRCA1	rs5910	French Gypsy mutation c.1544+1G> A	rs5918	rs5919	rs11870252
Patients	Pedigree	35 501 011	38 458 367	39 805 315	39 813 279	42 715 729	42 719 539	42 732 824
M_1	A	7.7	3.3	C/C	A/A	C/C	T/T	T/T
M_2	A	3.7	3.7	C/C	A/G	C/T	T/T	T/T
M_3	B	3.7	8.9	C/T	G/G	T/T	T/T	T/T
M_4	B	5.5	3.7	C/T	G/G	T/T	C/T	C/T
M_5	B	5.5	3.9	C/C	A/A	C/C	T/T	T/T
M_6	B	5.5	3.3	C/C	A/A	C/C	T/T	T/T
M_7	B	5.5	3.7	C/C	A/G	C/T	T/T	T/T
M_8	C	5.5	3.3	C/C	A/A	C/C	T/T	T/T
M_9	C	5.5	3.3	C/C	A/A	C/C	T/T	T/T
M_10	C	5.5	3.3	C/T	A/G	C/T	T/T	T/T
M_11[b]	—	5.5	3.4	C/C	A/G	C/T	T/T[b]	T/T
M_12	—	5.5	6.8	C/T	G/G	T/T	T/T	T/T
M_13	—	3.3	5.5	T/T	G/G	T/T	T/T	T/T
M_14	—	5.5	3.7	C/T	A/G	C/C	T/T	T/T
M_15[b]	—	5.7	3.3	C/C	A/G	C/T	T/T[b]	T/T
M_16	—	5.5	3.3	C/C	A/A	C/C	T/T	T/T
M_17	—	7.7	5.5	C/C	G/G	T/T	T/T	T/T
M_18	—	7.7	3.3	C/C	A/A	C/C	T/T	T/T
M_19	—	7.7	5.7	T/T	G/G	C/T	T/T	T/T
M_20	—	5.7	3.7	C/T	G/G	T/T	T/T	T/T
M_21	—	5.7	3.3	C/C	A/A	C/C	T/T	T/T
M_22	—	5.7	3.7	C/T	G/G	T/T	T/T	T/T
M_23	—	5.5	3.3	C/C	A/A	C/C	T/T	T/T

[a]Fragment size designations: THRA=176 bp=1; BRCA1=157 bp=1.
[b]Not phased.

Table 3 Identification and frequency of c.1544+1G > A haplotypes

			Markers				Haplotype
			rs5910	rs5918	rs5919	rs11870252	frequency[a]
Haplotype	THRA	BRCA1	ITGA2B, E30[b]	ITGB3, E3	ITGB3, E6	ITGB3, IVS11	All c.1544+1G> A (22)
H1	5	3	C	C	T	T	15 (68%)
H2	5	9	C	C	T	T	1 (5%)
H3	7	3	C	C	T	T	6 (27%)

[a]Proportion of each haplotype carrying the French Gypsy mutation.
[b]E, exon ; IVS, intron.

Table 4 Estimation of the age of the French Gypsy mutation

Effective population size	Disease allele frequency[a]	Non-recombinants frequency[b]	Population growth rate	Age (generations)	Age (years)
100 000	0.001	0.7	1.33 (1.2, 1.7)	15.3 (11, 22.1)	382.5 (275, 552.5)
100 000	0.01	0.7	1.7 (1.5, 2.2)	13.1 (10.3, 17.9)	327.5 (257.5, 447.5)
100 000	0.1	0.7	2.2 (1.9, 2.8)	12.1 (10.2, 15.7)	302.5 (255, 392.5)
25 000	0.001	0.7	1.19 (1.08, 1.46)	17.5 (12.7, 25.6)	437.5 (317.5, 460)
25 000	0.01	0.7	1.45 (1.27, 1.87)	14.2 (10.6, 20.1)	355 (265, 502.5)
25 000	0.1	0.7	1.86 (1.62, 243)	12.7 (10.3, 16.9)	317.5 (257.5, 422.5)
10 000	0.001	0.7	1.14 (1.06, 1.37)	18.9 (14, 27.5)	472.5 (350, 687.5)
10 000	0.01	0.7	1.33 (1.2, 1.7)	15.3 (11, 22.1)	382.5 (275, 552.5)
10 000	0.1	0.7	1.7 (1.5, 2.2)	13.1 (10.3, 17.9)	327.5 (257.5, 447.5)

[a]Estimated.
[b]Calculated from the percentage of recombinant haplotypes.
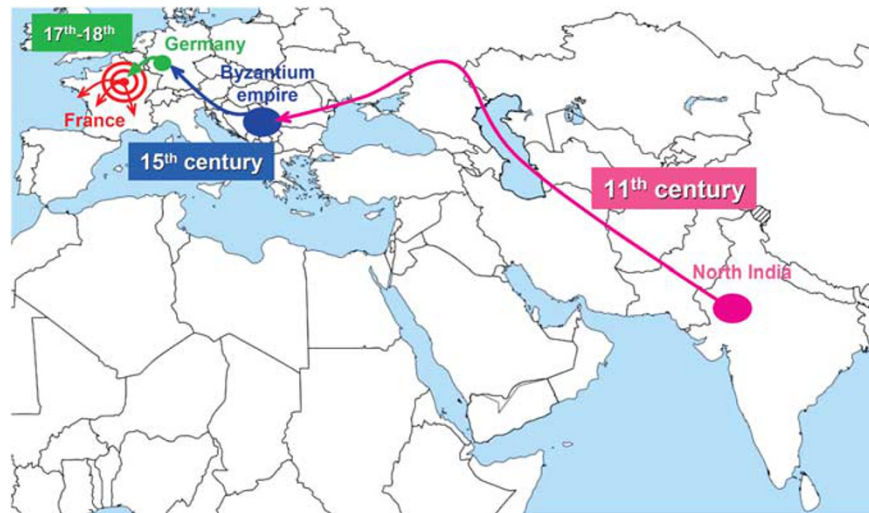
Figure 3 Origin of the French Gypsy mutation. Origin of Gypsies ●, European migration in Balkans ●, migration in Germany ●, introduction of the French Gypsy mutation in France and founder effect ●.

sub-populations among themselves based in part on territorial, cultural and dialect differences. The main branches in France are the Gitans, mostly in Southern France (Spain had been for a long time their favored country); and the Manouches, cousins of the Sinties (who stayed in Germany), who bear Germanic names.[10,11,15]

On the basis of our results, it is possible to infer the migration history of these Manouche families in France (Figure 3). It is likely that they moved from Germany to the North of France between the seventeenth and eighteenth century. Our results support a founder effect at that time and that this c.1544+1G>A mutation was brought by one founder (or a small group of related founders). Then, the number of carriers of this mutation increased within the Alsace region.[6] Limited social interactions occurred between the autochthonous population and the foreigners, because of a strong family history of consanguineous unions and maintenance of family structure. It also seems reasonable to assume that, at the time of the appearance of the mutation, the restricted movements of these Manouche families were linked to favorable local agro-ecological conditions.[30] Files from registries of the Alsace region suggest a population well integrated into the local agrarian society.[30] The high rate of consanguinity in this isolated Manouche group has resulted in an over-representation of the French Gypsy mutation. For a long time, the movements of these Manouches remained confined within the North of France. Demographic explosion, search of new outlets and also perhaps the taste of adventure pushed the Manouches to disseminate to other regions of France during the nineteenth century, propagating the mutation.[30]

The results of this study provide additional support to the historical and demographic parameters occurring in these Manouche families. Although our demographic model does not infer a definitive migrational history, this study provides additional insights into Manouche population movements, through the analysis of this founder mutation.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

1 Nurden AT: Glanzmann thrombasthenia. *Orphanet J Rare Dis* 2006; **1**: 10.
2 Seligsohn U, Rososhansky S: A Glanzmann's thrombasthenia cluster among Iraqi Jews in Israel. *Thromb Haemost* 1984; **52**: 230–231.
3 Awidi AS: Increased incidence of Glanzmann's thrombasthenia in Jordan as compared with Scandinavia. *Scand J Haematol* 1983; **30**: 218–222.
4 Piippo K, Laitinen P, Swan H *et al*: Homozygosity for a HERG potassium channel mutation causes a severe form of long QT syndrome: identification of an apparent founder mutation in the Finns. *J Am Coll Cardiol* 2000; **35**: 1919–1925.
5 Heyer E: One founder/one gene hypothesis in a new expanding population: saguenay (Quebec, Canada). *Hum Biol* 1999; **71**: 99–109.
6 Levy JM, Mayer G, Sacrez R, Ruff R, Francfort JJ, Rodier L: Glanzmann-Naegeli thrombasthenia. Study of a strongly endogamous ethnic group. *Ann Pediatr (Paris)* 1971; **18**: 129–137.
7 Levy JM, Stoll C, Gardea A, Bigel P: Clinics and genetics of Glanzmann's thrombasthenia (author's translation). *Nouv Rev Fr Hematol* 1976; **16**: 399–405.
8 d'Oiron R, Menart C, Trzeciak MC *et al*: Use of recombinant factor VIIa in 3 patients with inherited type I Glanzmann's thrombasthenia undergoing invasive procedures. *Thromb Haemost* 2000; **83**: 644–647.
9 Martin I, Kriaa F, Proulle V *et al*: Protein A sepharose immunoadsorption can restore the efficacy of platelet concentrates in patients with Glanzmann's thrombasthenia and anti-glycoprotein IIb-IIIa antibodies. *Br J Haematol* 2002; **119**: 991–997.
10 Schlegel N, Gayet O, Morel-Kopp MC *et al*: The molecular genetic basis of Glanzmann's thrombasthenia in a gypsy population in France: identification of a new mutation on the alpha IIb gene. *Blood* 1995; **86**: 977–982.
11 de la Salle C, Schwartz A, Baas MJ, Lanza F, Cazenave JP: Detection by PCR and HphI restriction analysis of a splice site mutation at the 5' end of intron 15 of the platelet GPIIb (alphaIIb integrin) gene responsible for Glanzmann's thrombasthenia type I in Gypsies originating from the Strasbourg area. *Thromb Haemost* 1995; **74**: 990–991.
12 Ruan J, Peyruchaud O, Nurden P *et al*: Family screening for a Glanzmann's thrombasthenia mutation using PCR-SSCP. *Platelets* 1998; **9**: 129–136.
13 Vinciguerra C, Bordet JC, Beaune G, Grenier C, Dechavanne M, Negrier C: Description of 10 new mutations in platelet glycoprotein IIb (alphaIIb) and glycoprotein IIIa (beta3) genes. *Platelets* 2001; **12**: 486–495.
14 Jacquelin B, Tuleja E, Kunicki TJ, Nurden P, Nurden AT: Analysis of platelet membrane glycoprotein polymorphisms in Glanzmann thrombasthenia showed the French gypsy mutation in the alphaIIb gene to be strongly linked to the HPA-1b polymorphism in beta3. *J Thromb Haemost* 2003; **1**: 573–575.
15 Nurden AT, Nurden P: Rebuttal: the French Gypsy mutation does not give rise to a particularly mild form of Glazmann's thrombasthenia. *J Thromb Haemost* 2003; **1**: 2459.
16 Kunicki TJ, Pidard D, Cazenave JP, Nurden AT, Caen JP: Inheritance of the human platelet alloantigen, PlA1, in type I Glanzmann's thrombasthenia. *J Clin Invest* 1981; **67**: 717–724.
17 Fiore M, Nurden AT, Vinciguerra C, Nurden P, Pillois X: Rapid diagnosis of the French gypsy mutation in Glanzmann thrombasthenia using high resolution melting analysis. *Thromb Haemost* 2010; **104**: 1076–1077.

18 French DL, Coller BS, Usher S *et al*: Prenatal diagnosis of Glanzmann thrombasthenia using the polymorphic markers BRCA1 and THRA1 on chromosome 17. *Br J Haematol* 1998; **102**: 582–587.

19 Austerlitz F, Kalaydjieva L, Heyer E: Detecting population growth, selection and inherited fertility from haplotypic data in humans. *Genetics* 2003; **165**: 1579–1586.

20 PHASEv2.1. www.stat.washington.edu/stephens/software.html.

21 International HapMap Consortium: The International HapMap Project. *Nature* 2003; **426**: 789–796.

22 French Senate report. http://www.senat.fr/rap/l96-283/l96-2831.html.

23 Mor-Cohen R, Rosenberg N, Peretz H *et al*: Disulfide bond disruption by a beta3-Cys549Arg mutation in six Jordanian families with Glanzmann thrombasthenia causes diminished production of constitutively active alphaIIbbeta 3. *Thromb Haemost* 2007; **98**: 1257–1265.

24 Rosenberg N, Hauschner H, Peretz H *et al*: A 13-bp deletion in alpha(IIb) gene is a founder mutation that predominates in Palestinian-Arab patients with Glanzmann thrombasthenia. *J Thromb Haemost* 2005; **3**: 2764–2772.

25 Fraser A: *The Gypsies*. Blackwell Publishers: Oxford, 1992.

26 Gresham D, Morar B, Underhill PA *et al*: Origins and divergence of the Roma (gypsies). *Am J Hum Genet* 2001; **69**: 1314–1331.

27 Kalaydjieva L, Gresham D, Calafell F: Genetic studies of the Roma (Gypsies): a review. *BMC Med Genet* 2001; **2**: 5.

28 Liegeois J-P: *Roma, Gypsies, Travellers*. Council of Europe Press: Strasbourg, 1994.

29 Morar B, Gresham D, Angelicheva D *et al*: Mutation history of the Roma/Gypsies. *Am J Hum Genet* 2004; **75**: 596–609.

30 Reyniers A: Manouches. *Tzigane studies* 2006; **26**: 1–144.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)