Am. J. Hum. Genet. 67:311–319, 2000

A Single Genetic Origin for the G101W CDKN2A Mutation in 20 Melanoma-Prone Families

Paola Ciotti,[1] Jeffery P. Struewing,[3] Michela Mantelli,[1] Agnès Chompret,[4] Marie-Françoise Avril,[4] Pier Luigi Santi,[2] Margaret A. Tucker,[3] Giovanna Bianchi-Scarrà,[1,*] Brigitte Bressac-de Paillerets,[5,*] and Alisa M. Goldstein[3,*]

[1]Department of Oncology, Biology and Genetics, University of Genova, and [2] Istituto Nazionale per la Ricerca sul Cancro, Geno; [3]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda; and [4]Service de Dermatologie and [5]Unité des Marqueurs Génétiques des Cancers, Institut Gustave Roussy, Villejuif Cedex, France

Germline mutations within the coding region of CDKN2A have been observed in affected members of melanoma-prone families. G101W is the most common CDKN2A missense mutation identified to date. It has been reported in several families from around the world, with a particularly high occurrence in France and Italy. Given the frequency of this mutation, we were interested in determining whether the mutation resulted from a single origin or represented a mutational hotspot in the CDKN2A gene. In addition, given the geographical distribution of the mutation, we examined the date of origination of the mutation and its migratory spread. We examined 10 families from Italy, 4 families from the United States, and 6 families from France with the G101W mutation. The following eight markers were employed for the haplotype analysis: IFNA, D9S736, D9S1749, D9S942, D9S1748, D9S1604, D9S171, and D9S126. Our findings showed no significant evidence for mutational heterogeneity, suggesting that all studied families derived from a single ancestral haplotype on which the mutation arose. Using maximum-likelihood methods, we estimated the mutation to have arisen 97 generations ago (1-LOD-unit support interval 70–133 generations) providing some explanation for the wide geographical spread of this common mutation, particularly in southwestern Europe. The presence of a founder mutation in a defined geographic area can facilitate carrier detection and genetic counseling and can provide an opportunity to study disease penetrance and the effect of environmental factors on the background of a common genetic susceptibility.

Introduction

Germline mutations within the coding region of CDKN2A (MIM 600160) have been observed in affected members of 9p21 linked families with inherited melanoma (Dracopoli and Fountain 1996; Hayward 1996). The population frequency of CDKN2A mutations, however, remains unknown. Overall, CDKN2A mutations have been observed in ~40% of 9p21-linked families (Hussussian et al. 1994; Soufir et al. 1998) but the frequency varies considerably across different studies (Hussussian et al. 1994; Kamb et al. 1994; Goldstein et al. 1995; Gruis et al. 1995; Holland et al. 1995; Walker et al. 1995; Borg et al. 1996; Flores et al. 1997; Harland et al. 1997; Platz et al. 1997; Soufir et al. 1998).

The CDKN2A gene encodes a protein, p16INK4a, which binds the cyclin-dependent kinases 4 and 6 (CDK4 and CDK6) inhibiting phosphorylation of the retinoblastoma protein (pRB). Therefore p16, the first melanoma-susceptibility gene to have been identified, negatively regulates progression through G1 into S phase of the cell cycle (Serrano et al. 1993, 1995). As such, it is an attractive candidate as a tumor-suppressor gene.

Many different mutations have been identified in melanoma-prone families from North America, Europe, and Australasia. Most mutations described to date are missense mutations scattered throughout the CDKN2A coding region. Some mutations have been observed only once (e.g., R58ter, L62P, and L97R), whereas others have repeatedly been found in different families (e.g., G101W, 113insArg, and 225del19). Haplotype analyses of recurrent mutations (113insArg, 225del19, and G-34T) in families from the same geographic areas have shown evidence for common founders (Gruis et al. 1995; Borg et al. 1996; Liu et al. 1999) rather than mutation hotspots in the CDKN2A gene. In addition, recent examination of two recurrent mutations from geographically diverse populations revealed evidence for a common founder for the M53I mutation and mul-

Received March 16, 2000; accepted for publication May 26, 2000; electronically published June 22, 2000.

Address for correspondence and reprints: Dr. Alisa M. Goldstein, Genetic Epidemiology Branch, National Cancer Institute, 6120 Executive Boulevard, MSC 7236, Bethesda, MD 20892-7236. E-mail: goldstea@exchange.nih.gov

* These authors contributed equally to this work.

0002-9297/2000/6702-0010$02.00

Table 1

Ancestry, Mutation, and Genotyping Information on G101W Families

			No. of					
Family	Location	Ancestry Origin	Genotyped Individuals	G101W Individuals	Affecteds	Genotyped Affecteds	Typed Unaffecteds/ Spouses	Reference
6101294	Italy (Liguria)	Celtic	9	4	3	1	5	Ghiorzo et al. (1999)
6091195	Italy (Liguria)	Celtic	1	1	2	1	0	Ghiorzo et al. (1999)
5080994	Italy (Liguria)	Celtic	2	1	3	1	1	Ghiorzo et al. (1999)
7070694	Italy (Liguria)	Celtic	3	1	2	0	2	Ghiorzo et al. (1999)
6120894	Italy (Liguria)	Celtic	1	1	2	1	0	Ghiorzo et al. (1999)
6090595	Italy (Liguria)	Celtic	8[a]	1	4	2[a]	6	Ghiorzo et al. (1999)
6110894	Italy (Liguria)	Celtic	5	3	3	1	2	Ghiorzo et al. (1999)
7091294	Italy (Liguria)	Celtic	10	2	4	3[b]	7	…
FMM 30	Italy (Rome)	…	2	2	2	2	0	…
6091098	Italy (Liguria)	Celtic	3	3	3	3	0	…
G	USA (KS)	S. Germany	6	5	5	3	3	Hussussian et al. (1994)
H	USA (PA)	Germany	4	4	3	3	1	Hussussian et al. (1994)
I	USA (PA)	S. Germany	5	5	4	2	2	Hussussian et al. (1994)
CC	USA (MD)	German/Alsatian	3	2	1	1	2	…
1648	France (Alsatian)	Celtic	2	4	3	2	0	Soufir et al. (1998)
2548	France (Alsatian)	Celtic	2	2	2	1	1	Soufir et al. (1998)
2000	France (SW)	Celtic	2	5	3	2	0	Soufir et al. (1998)
1220	France (SW)	Celtic	1	3	2	1	0	Soufir et al. (1998)
1379	France (Corsica)	Celtic	2	5	6	1	1	Soufir et al. (1998)
1893	Italy (Liguria)	Celtic	1	2	?[c]	1	0	…

[a] Includes one inferred genotype.
[b] One affected individual was typed for 9p markers but was not screened for the G101W mutation.
[c] Proband, who had multiple primary melanomas, was adopted. Family history is unknown.

tiple origins for the 23ins24 duplication (Pollock et al. 1998).

G101W is the most common CDKN2A missense mutation identified to date. It has been reported in numerous families from around the world (Hussussian et al. 1994; Kamb et al. 1994; Goldstein et al. 1995; Whelan et al. 1995; Soufir et al. 1998; Ghiorzo et al. 1999; Ruiz et al. 1999) with a particularly high occurrence in France and Italy (Soufir et al. 1998; Ghiorzo et al. 1999). The G101W mutation has been shown to result in a functionally defective protein yielding an impaired ability to inhibit the catalytic activity of complexes of cyclin D1-CDK4 and cyclin D1- CDK6 in vitro (Ranade et al. 1995), and to inhibit cyclin D1-CDK4 in a reconstituted pRb kinase assay (Parry and Peters 1996).

Given the frequency of this mutation, we were interested in determining whether the mutation resulted from a single origin or represented a mutational hotspot in the CDKN2A gene. In addition, given the geographical distribution of the mutation, we examined the date of origination of the mutation and its migratory spread. In this study, we present 10 families from Italy, 4 families from the United States, and 6 families from France with the previously reported (Soufir et al. 1998; Ghiorzo et al. 1999; Goldstein et al. 2000) G101W mutation in the CDKN2A gene. We have investigated whether the G101W mutation in families from different countries represents an independent mutational event or indicates the presence of a common founder. Analysis of a large number of genetic markers in the 9p21 region allowed us to evaluate this possibility in these families. Our findings showed no significant evidence for mutational heterogeneity, suggesting that all studied families derived from a single ancestral haplotype on which the mutation arose. In addition, the mutation was estimated to have arisen 97 generations ago (1-LOD-unit support interval 70–133 generations) providing some explanation for the geographical spread of this common mutation. The presence of a founder effect can facilitate carrier detection and genetic counseling in southwestern Europe and can provide an opportunity to study disease penetrance and the effect of environmental (and modifying genetic) factors in a more homogeneous setting.

Subjects and Methods

Subjects/Families

Table 1 shows details of the Italian, French, and American families. Written informed consent was obtained prior to participation, under IRB-approved protocols from the three study sites (the National Cancer Institute, the Institut Gustave Roussy, and the University of Genova). Further information on the geographic location of

the current and ancestral settlements of the families are described under Geographic Evidence.

Italian Families

The clinical data for 7 of 10 Italian families (families 6120894, 5080994, 7070694, 6110894, 6090595, 6101294, and 6091195) and isolation of DNA have been previously described (Ghiorzo et al. 1999). Three families (7091294, FMM 30, and 6091098) have not been previously described. Two affected individuals, four spouses, and three unaffected individuals were genotyped in family 7091294. An additional affected individual was haplotyped for 9p markers but was not screened for the G101W mutation. Two affected brothers were genotyped in family FMM 30. Three melanoma patients (sister, brother, and his son) were genotyped in family 6091098. In total, 14 patients with cutaneous malignant melanoma, 20 unaffected relatives, and 9 spouses from the 10 Italian families with a G101W mutation were haplotyped. Genotypes of two additional affected members were inferred.

French Families

Five French families (1648, 2548, 2000, 1220, and 1379) which harbored the G101W mutation have been reported elsewhere (Soufir et al. 1998). Three of the families contained at least three patients with melanoma; the other two families had two melanoma subjects each. The sixth family (1893) includes a single proband with multiple primary melanomas. This proband was adopted and has an unknown family history. Eight melanoma patients and two unaffected relatives from these six families were haplotyped.

American Families

Four American families from the National Cancer Institute carried the G101W mutation. All diagnoses of melanoma were confirmed using histologic review of pathological material or review of pathology reports. Data for three of the American families (families G, H, and I) have been previously presented (Hussussian et al. 1994; Goldstein et al. 2000). Briefly, these three families contained an average of four melanoma patients, with a median age at diagnosis of 28 years. Eight melanoma patients, four spouses, and two unaffected family members from these three families were haplotyped. The other family (CC) has not been previously reported. Although this family reported multiple family members with melanoma, diagnostic confirmation has only been possible in the proband, who has had multiple melanoma tumors. One melanoma patient, one spouse, and one unaffected mutation carrier were genotyped in this family.

Genotyping

The following eight markers were employed for the haplotype analysis: IFNA, D9S736, D9S1749, D9S942, D9S1748, D9S1604, D9S171, and D9S126. Primer sequences were obtained from the Genome Database. Markers D9S1749 and D9S942 flanked the CDKN2A gene. D9S1749 is located ~0.0105 M distal to exon 2 of the CDKN2A gene, the exon containing the G101W mutation; D9S942 is located ~0.0015 M proximal to CDKN2A exon 2.

Genotyping was performed under slightly different conditions at each of the three study sites. PCR assays were performed using either ^{32}P-labeled primers and were electrophoresed through 7 M urea/6%–8% polyacrylamide gels and were run alongside a size standard generated from pBSMB DNA using the forward M13 (lacZ) primer or with fluorescently-labeled primers and run on ABI 377 or ABI 310 semi-automated sequencers to determine allele sizes. Marker D9S1749 was amplified using 1.3 M betaine in the PCR reaction. Complete details of PCR conditions are available on request to the authors. The molecular results were combined using six common DNAs typed in each laboratory as allele size controls for all markers. Allele sizes for all markers except D9S736 and D9S126 are comparable with those from the haplotype study of Pollock et al. (1998).

Control Subjects

A pool of 43 control subjects was used for allele- and haplotype-frequency estimation. The controls comprised 20 unaffected spouses collected from the Italian families, 2 spouses/unaffected non–mutation carriers from the French families, and 21 spouses/non–mutation carriers from the American families. Of the 20 Italian spouses, 9 were from the G101W families and probably come from the same region (Liguria); the other 11 Italian spouses belong to p16 wild-type melanoma families and are from other regions of Italy. Of the American control subjects, 5 were from the G101W families, and the other 16 were from other American melanoma-prone families.

Dating the Mutation

We used two approaches to calculate when the G101W mutation originated. First, we used a maximum-likelihood (MLE) method developed and kindly provided by D. Goldgar (Neuhausen et al. 1996, 1998). Briefly, the joint likelihood of the G101W haplotypes was written as a function of the recombination fraction between the disease and each marker, the number of generations (G) since the mutation arose, and the mutation rate and allele frequencies at each marker locus. All markers except D9S1749 and D9S942 were assumed to have a mutation rate of 0.0006. D9S942 was assumed

Table 2

Haplotype Analysis of G101W Mutation Carriers for 9p Markers

	Haplotype for Family[a]													
	Italian							French			American			
Marker	6101294	5080994	7070694	7091294	6090595	6110894	6091098	2000	1648	1379	G	H	I	CC
IFNA	4	3	4	7	3	3	4	3,6	3	2	6	2	6	6
D9S736	3	3	3	3	3	3	3	5	4	4	4	4	4	6
D9S1749	**18**	**20**	**20**	**17**	**19**	**20**	**19**	**21**	**17**	**18**	**18**	**21**	**18**	15
D9S942	**9**	**9**	**9**	**9**	**9**	**9**	**9**	11	**9**	**9**	**9**	**9**	**9**	**9**
D9S1748	**9**	**9**	**9**	**9**	**9**	10	**9**	**9**	**9**	**9**	10	**9**	10	**9**
D9S1604	2	2	2	2	2	2,1	2,1	**2**	**2**	2,1	**2**	**2**	2,1	**2**
D9S171	1	5	5	9	5	5	5	5	1	1	1	1	1	8
D9S126	5	5	7	3	4	3	5,7	4,6	6	6	1	4,6	3,4	6

[a] Alleles that are part of the common disease-related haplotype are shown in boldface.

to have a mutation rate of 0.002, on the basis of the large number of alleles (n = 23 alleles). D9S1749 was assumed to have a mutation rate of 0.01, on the basis of both the large number of alleles (n = 32 alleles) and the observation of mutations within families. Another parameter, μ_D, the proportion of families with an independent mutation identical to that of the presumed ancestral haplotype, was incorporated into the approach to examine for mutational heterogeneity, in a manner similar to standard linkage analysis. The MLE method was used to find the value of G that best fitted the pattern of haplotype sharing at the 8 marker loci. When haplotypes could not be determined with certainty, all possible haplotypes consistent with the observed multilocus genotypes were considered in the analysis. Approximate support intervals were calculated by finding the value of G on either side of the most likely value that had a $\geq$10-fold decrease in likelihood. A heterogeneity test of mutation origin was conducted by comparing the likelihood at the MLE of G and μ_D with the analogous likelihood, assuming $\mu_D = 0$. Each generation was estimated to be 20 years.

Second, we used the approach suggested by Risch et al. (1995) (designated as "Q method"). Briefly, Q is the observed frequency of disease chromosomes not carrying the progenitor marker allele; p_N is the frequency of the progenitor marker allele on normal chromosomes; θ is the recombination fraction between the disease and marker locus; and G is the number of generations since the appearance of the mutation. As such, an estimate of the number of generations since the appearance of the G101W mutation is $G = \log[(1 - p_N - Q)/(1 - p_N)]/\log(1 - \theta)$. The two markers flanking the CDKN2A gene, D9S1749, and D9S942 were used to estimate the age of the origin of the G101W mutation in the second method. In a manner similar to the MLE method above, we calculated 90% confidence intervals for G. We used two sets of data to estimate Q. First, we used the proband from each family with a known disease chromosome for D9S1749 or D9S942 (n = 17 probands). Second, we used all melanoma cases and mutation carriers, who shared the same core haplotype (n = 37 for D9S1749 and n = 38 for D9S942). This latter approach is likely to be more accurate due to inclusion of a larger number of subjects and the allowance for known recombination events within the families (Risch et al. 1995).

Geographic Evidence

All Italian families (except family FMM 30; see table 1) come from a very small area (<40 km^2) on the eastern coast of the region of Liguria. The probands from these families reported that their families have always lived in this area and that they know of no ancestors who came from other regions of Italy. All of the American families had knowledge about their ancestries. Two families (G, I) migrated from southern Germany in the mid-to-late 1800s. One family (H) migrated from Germany before the mid-1800s, but their region of origin and date of migration are unknown. The fourth family (CC) had a German/Alsatian origin and migrated to America in the mid-1800s. For the French families, ancestries were obtained by asking for the birthplace of the proband's grandparents. Two families were of Alsatian origin; one family originated in Corsica. Two families came from the southwest of France (Ardeche, Lozere, and Aveyron, which are located within close proximity of each other). The proband of the sixth family was born in Liguria and was adopted there. No recent migration of the families has occurred.

Results

Haplotype analysis by means of eight polymorphic markers spanning the CDKN2A locus was performed on index cases and additional family members (when available) to determine whether carriers from different families harbored the same mutation identically by descent. Table 2 shows the disease haplotypes for families where phase could be determined. Both alleles are in-

dicated for markers for which segregating alleles could not be unequivocally determined. Table 3 presents genotypes for families that could not be haplotyped. There were six such families: three Italian (6091195, 6120894, and FMM 30) and three French (1220, 1893, and 2548). We determined whether the genotypes for these families were consistent with the haplotypes of families for whom two or three generations were available.

Table 4 shows the ages of origin of the G101W mutation estimated by use of the MLE and Q methods (Risch et al. 1995; Neuhausen et al. 1996, 1998). In addition to the families shown in tables 2 and 3, seven (for the MLE method) and three (for the Q method) additional French "sporadic" melanoma subjects were used to estimate the ages of origin of the G101W mutation. These melanoma patients were the only ones in their families affected by melanoma. In a separate study testing for CDKN2A and CDK4 mutations in patients affected by CMM with multiple primary melanoma tumors and no family history of CMM (B. Bressac-de Paillerets, A. Chompret, and M.-F. Avril, unpublished data), seven French melanoma patients were shown to carry the G101W mutation. Genotypes/haplotypes for these patients were added to the examination of the age of origin for the G101W mutation. From the MLE method, the mutation was estimated to have arisen 97 generations ago (1-LOD-unit support interval 70–133 generations) or ~1940 years ago (1-LOD-unit support interval 1,400–2,660 years). The results from the Q method were consistent with that of the MLE method but with wider confidence intervals in the Q method because age-of-origin estimates derive from a single marker. There was no significant evidence for mutational heterogeneity, suggesting that all families studied represent derivations from a single ancestral haplotype on which the mutation arose.

The D9S942-D9S1604 haplotype 9-9-2 appears to be common across all families (table 2), after recombination over time is allowed for. In addition, the six families from table 3 all have genotypes consistent with this disease haplotype. One French family from southwestern France (2000) had the 11 allele at D9S942, rather than the 9 allele seen in all other families. Similarly, three haplotyped families carried the 10 allele, rather than the 9 allele, at D9S1748. We observed a previously described variation in the allele size of the D9S1749 marker (Pollock et al. 1998), caused by replication slippage resulting in the loss or gain of ≥ 1 repeat units during meiosis. This slippage occurred in members of the same family, as well as across different families. After allowing for a high mutation rate in D9S1749, a common allele centering on allele 19 was observed. Among the haplotyped families, two carried the 17 allele, four carried the 18 allele, two carried the 19 allele, three carried the 20 allele, and two carried the 21 allele (table 2). All of the genotyped families carried at least one of these five alleles (table 3).

A common disease-related haplotype extended further in several families and was related to geographic origin. For example, the two American families from southern Germany (G and I) carried the same haplotype from IFNA to D9S171. The Italian families, who—except for FMM 30—came from a small area on the eastern coast of Liguria, all carried the same D9S736 allele (allele 3). In addition, two Italian families (5080994 and 7070694) also carried identical haplotypes from IFNA to D9S171. Ignoring 2-bp differences at D9S1749 increased this number to four haplotyped Italian families plus two genotyped Italian families consistent with this haplotype. Several French families had the same haplotype from D9S736 to D9S126, when replication slippage at marker D9S1749 was allowed for.

Forty-three controls were typed for the same markers used in the haplotype analysis for allele- and disease haplotype–frequency estimation. A core haplotype for D9S1749-D9S942-D9S1748-D9S1604 could be determined in 31 of 43 individuals. None of the haplotyped controls carried the core disease haplotype (0 of 62 haplotypes). In the entire pool (including 12 subjects for whom phase information was not available), no subjects carried the disease haplotype. The frequencies of the G101W-related alleles from the shared core disease haplotype—D9S1749 (alleles 17-21), D9S942, D9S1748, and D9S1604—were 0.264, 0.088, 0.122, and 0.56, respectively. Thus, the probability of finding the G101W haplotype in the general population is very low (i.e., 1.6×10^{-3}).

Table 3

Genotype Analysis of 9p Markers Where Phase Is Not Determined

	Genotype for Family[a]					
	Italian			French		
Marker	FMM 30	6091195	6120894	1220	1893	2548
IFNA	3,4	3,4	2,3	2,3	4,6	3,3
D9S736	3,2	3,4	3,4	4,5	4,5	4,4
D9S1749	**20**,23	**20**,22	**19**,**20**	**19**,**21**	**21**,24	**18**,32
D9S942	**9**,23	**9**,9	**9**,19	**9**,11	**9**,4	**9**,11
D9S1748	**9**,4	**9**,4	**9**,6	**9**,7	**9**,7	**9**,5
D9S1604	**2**,2	**2**,1	**2**,1	**2**,1	**2**,1	**2**,2
D9S171	5,8	5,8	5,2	5,7	5,5	5,9
D9S126	4,6	2,7	3,5	4,6	6,6	4,4

[a] Alleles consistent with the haplotype from table 2 are shown in boldface.

Discussion

In this paper, we have analyzed genotypic data from 20 geographically diverse melanoma-prone families from Italy, France, and the United States with the most frequent CDKN2A mutation so far identified, the G101W

Table 4

Estimates of the Age of the G101W Mutation

A. MLE Method			
Loci	Q	Generations (1-LOD Support Interval)	Years (1-LOD Support Interval)
IFNA, D9S736, D9S1749, CDKN2A, D9S942, D9S1748, D9S1604, D9S171, D9S126	...	97 (70–133)	1,940 (1,400–2,660)
B. Q Method			
Locus and Group	Q	Generations (90% CI)	Years (90% CI)
D9S1749:			
Probands	.588 (.392, .784)	115 (60–261)	2,297 (1,200–5,220)
Affecteds and carriers	.514 (.379, .649)	90 (57–141)	1,804 (1,140–2,820)
D9S942:			
Probands	.118 (0.0, .247)	92 (0–210)	1,846 (0–4,200)
Affecteds and carriers	.105 (.023, .187)	81 (17–153)	1,630 (340–3,060)

mutation. This mutation has been observed in melanoma-prone families from Australia, France, Italy, Spain, and North America; it is the most common CDKN2A mutation in France and Italy reported to date. Our findings showed no significant evidence for mutational heterogeneity, suggesting that all studied families derived from a single ancestral haplotype on which the mutation arose. When maximum-likelihood methods were used, the mutation was estimated to have arisen 97 generations ago (1-LOD-unit support interval 70–133 generations). A second method for estimating the origination of the mutation (Q method) showed results consistent with the MLE approach.

The age of origin of the G101W mutation depends on the methods used for estimation purposes (Risch et al. 1995; Neuhausen et al. 1996, 1998). The MLE method is particularly sensitive to the marker mutation rates. Similarly, the estimated distance between the disease and marker loci, as well as the precise determination of subjects with the known disease chromosome, is critical for the Q method. However, both of the methods, with their different assumptions, produced similar results, thus permitting an initial examination of the genetic history of this common mutation. Historical records suggest a number of hypotheses to explain the possible geographical spread of the G101W mutation and its place of origin.

All the Italian families, with the exception of FMM 30, came from a small area on the eastern coast of Liguria, and no proband ever recalled having ancestors who came from other regions of Italy. We, therefore, reviewed the documented Celtic ancestry and genetic structure of Liguria to try to identify an historical explanation for the geographically dispersed common founder mutation. Historical evidence supporting the hypothesis for a remote mutation origin come from archaeological findings of the late Hallstatt period (700–500 B.C.) from the Vara Valley, the geographic center of the area from which the Italian families derive. The archeological data support the existence of colonies or outposts of Celts who likely emigrated from the Po Valley (Hubert 1974), ~50 miles away.

Archeological evidence from as early as the 16th and 15th centuries B.C. show Celtic peoples in the upper course of the river Danube and in eastern France. The Central and Eastern Hallstatt populations (7th–6th centuries B.C.), who spread throughout most of Western Europe, including the northernmost part of Italy, were certainly Celtic (Dietler 1995). During the 6th century B.C., the Celts conquered all of the Rhineland, Belgium, and north-central France, along with part of the British Isles and a large portion of the Iberian peninsula. A new wave of migrations began around 400 B.C. (Kruta 1987; Vitali 1987; Violante 1993), when the Celts took control of northern Italy and settled the Po valley, as recorded by Livy(Historiae V). Although Liguria was inhabited from prehistorical times by the Ligurians, probably the most ancient Italian population, it appears that a Celtic tribe moved from the Po Valley, through the Appennine Mountains, and settled the nearby hills along the Ligurian coast. This region of Liguria is the area where the Italian G101W families have been living for generations.

Further, molecular results confirm the presence of Celtic populations in a large portion of the Po Valley, specifically in the provinces of Vercelli, Pavia, and Piacenza. Examination of HLA allele frequencies using population-based samples from Liguria, Piacenza, Pavia, and North Lombardy, revealed the presence of both ancient Ligurian and Celtic populations (Guglielmino et al. 1998). Moreover, earlier studies (Piazza et al. 1988) had shown that the genetic structure of present-day Italy likely reflected ancient patterns dating back to pre-Roman times. Specifically, the northern regions of

Italy, under Celtic influence, are well differentiated from the central-southern regions, and, within the former group, the western and eastern regions are remarkably distinct (Pallottino 1984; Rendine et al. 1997).

Alternatively, if the mutation dates to the beginning of the support interval, 70 generations or 1,400 years ago (i.e., 600 A.D.), other historical data provide another hypothesis for the spread of the mutation. During that period, a number of Irish monks began moving from their homeland towards the continent. In particular, Saint Columban, an Irish monk born near Leinster in 540 A.D., began a voyage towards Gaul in 590 A.D. with a group of companions and established himself in Burgundy, where he founded three monasteries (Annegray, Luxeuil, and Fontaine). In 610, for religious reasons, he was exiled. As an itinerant preacher he traveled through Alsace, the Rhineland, and Switzerland. He then crossed the Alps and reached Bobbio (Piacenza), in the upper Po Valley, in 612. There he founded a monastery and died a few years later (Mabillon 1733; Metlake 1914). Interestingly, Bobbio lies in the tract of the valley that starts behind the Appenines, $\leqslant$40 miles from the villages where the Italian families come from.

A recent study has questioned the intergenerational time intervals frequently used to date origins of mutations (Tremblay and Vezina 2000). Examination of a large population register from Quebec, Canada suggested that 30 years was a better estimate of intergenerational intervals than 20 or 25 years. However, the study was only able to examine relatively recent times (e.g., from the 17th century to the present). The applicability of the 30-year intergenerational interval to mutations of remote origin, such as the G101W mutation, remains unknown. Use of the 30-year generation interval produced a date of origin for the G101W mutation of 2,910 years (1-LOD-unit support interval 2,100–3,990 years), which is still consistent with the historical evidence for the spread of the mutation in southwestern Europe.

Haplotype analyses of recurrent mutations (e.g., 113insArg, 225del19, and G-34T) from geographically isolated areas have consistently shown evidence for common founders, rather than mutation hotspots in the CDKN2A gene (Gruis et al. 1995; Borg et al. 1996; Liu et al. 1999). However, for few of these mutations have there been enough families available to estimate the mutation's age of origination. The results from the present study are consistent with previous findings that show common founders for most of the recurrent mutations in the CDKN2A gene (Gruis et al. 1995; Borg et al. 1996; Pollock et al. 1998; Liu et al. 1999), suggesting that the CDKN2A gene is relatively stable. The only recurrent mutation, to date, that does not appear to derive from a single founder is the 23ins24 mutation (Pollock et al. 1998). This finding would be expected on the basis of the inherent instability of the tandem repeat region that produced the 24-bp insertion. Whether all founder CDKN2A mutations also have relatively remote origins remains to be tested.

Although it was not possible to unequivocally determine the precise geographic location where the mutation originated and how it spread around the world, it is likely that the mutation originated in Southwestern Europe. In addition, the historical and archeological records are consistent with the estimated date of the origin of the mutation, using both the MLE and Q methods, and the higher frequency of the mutation in Southwestern Europe compared to Northern Europe. Additional families from other geographic areas may help better discriminate between the various hypotheses for the origin and spread of this founder mutation.

Acknowledgments

We wish to thank the participating families, whose generosity and cooperation have made this study possible. We thank Danièle Pham, Johny Bombled, and Josyane Le Calvez for technical assistance; Mary Fraser, Laura Fontaine, and Deborah Zametkin for nursing support; and Jean-Pierre Cesarini, Rosine Guimbaud, Jean-Loïc Michel, and Didier Couillet for clinical/dermatological support. We acknowledge the Biomedical Supercomputing Center of the National Cancer Institute–Frederick Cancer Research and Development Center for allocation of computing time and staff support. This work was partly supported by grants from the AP-HP (CRC96011), the Fondation Sanofi Vaincre le Mélanome, and the Association pour la Recherche contre le Cancer, Subvention libre 1998; and the MURST National Project 1998 (to G.B.S) and the Ministero della Sanità 1997 (to P.L.S.).

Electronic-Database Information

The accession number and URL for data in this article are as follows:

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nim.nih.gov/Omim (for CDKN2A [MIM 600160])

References

Borg A, Johansson U, Johansson O, Hakansson S, Westerdahl J, Masback A, Olsson H, et al (1996) Novel germline p16 mutation in familial melanoma in southern Sweden. Cancer Res 56:2497–2500

Dietler M (1995) "Early Celtic" socio-political relations: ideological representation and social competition in dynamic comparative perspective. In: Arnold B, Gibson DB (eds) Celtic chiefdom, Celtic state. Cambridge University Press, Cambridge

Dracopoli NC, Fountain JW (1996) CDKN2A mutations in melanoma. Cancer Surv 26:115–131

Flores JF, Pollock P, Walker GJ, Glendening JM, Lin A, Palmer JM, Walters MK, et al. (1997) Analysis of the CDKN2A,

CDKN2B and CDK4 genes in 48 Australian melanoma kindreds. Oncogene 15:2999–3005

Ghiorzo P, Ciotti P, Mantelli M, Heouaine A, Queirolo P, Rainero ML, Ferrari C, et al (1999) Characterization of Ligurian melanoma families and risk of occurrence of other neoplasia. Int J Cancer 83:441–448

Goldstein AM, Fraser MC, Struewing JP, Hussussian CJ, Ranade K, Zametkin D, Fontaine LS, et al (1995) Increased risk of pancreatic cancer in melanoma-prone kindreds with p16INK4 mutations. N Engl J Med 333:970–974

Goldstein AM, Struewing JP, Chidambaram A, Fraser MC, Tucker MA (2000) Genotype-phenotype relationships in American melanoma-prone families with CDKN2A and CDK4 mutations. J Natl Cancer Inst (in press)

Gruis NA, van der Velden PA, Sandkuijl LA, Prins DE, Weaver-Feldhaus J, Kamb A, Bergman W, et al (1995) Homozygotes for CDKN2A (p16) germline mutation in Dutch familial melanoma kindreds. Nat Genet 10:351–353

Guglielmino CR, De Silvestri A, Rossi A, De Micheli V (1998) Surnames, HLA genes and ancient migrations in the Po Valley (Italy). Ann Hum Genet 62:261–269

Harland M, Meloni R, Gruis N, Pinney E, Brookes S, Spurr NK, Frischauf AM, et al (1997) Germline mutations of the CDKN2 gene in UK melanoma families. Hum Mol Genet 6:2061–2067

Hayward NK (1996) The current situation with regard to human melanoma and genetic inferences. Curr Opin Oncol 8: 136–142

Holland EA, Beaton SC, Becker TM, Grulet OMC, Peters BA, Rizos H, Kefford RF, et al (1995) Analysis of the p16 gene, CDKN2, in 17 Australian melanoma kindreds. Oncogene 11:2289–2294

Hubert H (1974) Les Celtes et l'expansion celtique jusquè l'époque de la Tène. Vol 1. Albin Michel, Paris (originally published 1932)

Hussussian CJ, Struewing JP, Goldstein AM, Higgins PAT, Ally DS, Sheahan MD, Clark WH Jr, et al (1994) Germline p16 mutations in familial melanoma. Nat Genet 8:15–21

Kamb A, Shattuck-Eidens D, Eeles R, Liu Q, Gruis NA, Ding W, Hussey C, et al (1994) Analysis of the p16 gene (CDKN2A) as a candidate for the chromosome 9p melanoma susceptibility locus. Nat Genet 8:22–26

Kruta V (1987) Il corallo, il vino e l'albero della vita: apporti peninsulari allo sviluppo della civiltà di La Tène. In: Vitali D (ed) Celti ed Etruschi nell'Italia centro-settentrionale dal V secolo a.C. alla romanizzazione. University Press Bologna, Bologna, pp 23–33

Liu L, Dilworth D, Gao L, Monzon J, Summers A, Lassam N, Hogg D (1999) Mutation of the CDKN2A 5′ UTR creates an aberrant initiation codon and predisposes to melanoma. Nat Genet 21:128–132

Livy. Historiae V.

Mabillon (1733) Acta sanctorum ordinis S. Benedicti. Vol 1. Venice, pp 3–26

Metlake G (1914) The life and writings of Saint Columban. Dolphin Press, Philadelphia

Neuhausen SL, Godwin AK, Gershoni-Baruch R, Schubert E, Garber J, Stoppa-Lyonnet D, Olah E, et al (1998) Haplotype and phenotype analysis of nine recurrent BRCA2 mutations in 111 families: results of an international study. Am J Hum Genet 62:1381–1388

Neuhausen SL, Mazoyer S, Friedman L, Stratton M, Offit K, Caligo A, Tomlinson G, et al (1996) Haplotype and phenotype analysis of six recurrent BRCA1 mutations in 61 families: results of an international study. Am J Hum Genet 58:271–280

Pallottino M (1984) Storia della prima Italia. Rusconi, Milano

Parry D, Peters G (1996) Temperature-sensitive mutants of p16CDKN2 associated with familial melanoma. Mol Cell Biol 16:3844–3852

Piazza A, Cappello N, Olivetti E, Rendine S (1988) A genetic history of Italy. Ann Hum Genet 52203–52213

Platz A, Hansson J, Mansson-Brahme E, Lagerlof B, Linder S, Lundqvist E, Sevigny P, et al (1997) Screening of germline mutations in the CDKN2A and CDKN2B genes in Swedish families with hereditary cutaneous melanoma. J Natl Cancer Inst 89:697–702

Pollock PM, Spurr N, Bishop T, Newton-Bishop J, Gruis N, van der Velden PA, Goldstein AM, et al (1998) Haplotype analysis of two recurrent CDKN2A mutations in 10 melanoma families: evidence for common founders and independent mutations. Hum Mutat 11:424–431

Ranade K, Hussussian CJ, Sikorski RS, Varmus HE, Goldstein AM, Tucker MA, Serrano M, et al (1995) Mutations associated with familial melanoma impair p16INK4 function. Nat Genet 10:114–116

Rendine S, Calafell F, Cappello N, Gagliardini R, Caramia G, Rigillo N, Silvetti M, et al (1997) Genetic history of cystic fibrosis mutations in Italy. I. Regional distribution. Ann Hum Genet 61:411–424

Risch N, de Leon D, Ozelius L, Kramer P, Almasy L, Singer B, Fahn S, et al (1995) Genetic analysis of idiopathic torsion dystonia in Ashkenazi Jews and their recent descent from a small founder population. Nat Genet 9:152–159

Ruiz A, Puig S, Malvehy J, Lazaro C, Lynch M, Gimenez-Arnau AM, Puig L, et al (1999) CDKN2A mutations in Spanish cutaneous malignant melanoma families and patients with multiple melanoma and other neoplasia. J Med Genet 36:490–493

Serrano M, Gomez-Lahoz E, DePinho RA, Beach D, Bar-Sagi D (1995) Inhibition of ras-induced proliferation and cellular transformation by p16INK4. Science 267:249–252

Serrano M, Hannon GJ, Beach D (1993) A new regulatory motif in cell cycle control causing specific inhibition of cyclin D/CDK4. Nature 366:704–707

Soufir N, Avril MF, Chompret A, Demenais F, Bombled J, Spatz A, Stoppa-Lyonnet D, et al (1998) Prevalence of p16 and CDK4 germline mutations in 48 melanoma-prone families in France. Hum Mol Genet 7:209–216

Tremblay M, Vezina H (2000) New estimates of intergenerational time intervals for the calculation of age and origins of mutations. Am J Hum Genet 66:651–658

Violante A (1993) La Tène period. I Celti a sud delle Alpi. Silvana Editoriale, Milano

Vitali D (1987) Celti ed Etruschi nell'Italia centro-settentrionale dal V secolo a.C. alla romanizzazione. Atti del Colloquio Internazionale, Bologna, 12–14 April 1985. University Press Bologna, Bologna

Walker GJ, Hussussian CJ, Flores JF, Glendening JM, Haluska FG, Dracopoli NC, Hayward NK, et al (1995) Mutations of the CDKN2/p16INK4 gene in Australian melanoma kindreds. Hum Mol Genet 10:1845–1852

Whelan AJ, Bartsch D, Goodfellow PJ (1995) A familial syndrome of pancreatic cancer and melanoma with a mutation in the CDKN2A tumor-suppressor gene. N Engl J Med 333: 975–977