

Analisi di 5 microsatelliti nelle tre maggiori isole del Mediterraneo Occidentale

C.M. Calò¹, L. Varesi², M. Memmi², G.E. Mameli¹, P. Moral³, G. Vona¹

¹ Dipartimento di Biologia Sperimentale, Sezione di Scienze Antropologiche, Università di Cagliari, Italy

² Faculté de Sciences et Technique (Civaren), Université de Corse, 20250 Corte, France

³ Facultat de Biologia, Secció d'Antropologia, Universitat de Barcelona, 0828 Barcelona, Spain.

Parole Chiave: microsatelliti, dinucleotide, Sardegna, Sicilia, Corsica

Riassunto

Sono stati esaminati 360 individui provenienti dalle tre maggiori isole del Mediterraneo Occidentale: Sardegna (Italia), Sicilia (Italia) e Corsica (Francia). Il DNA è stato amplificato tramite PCR e i frammenti sono stati separati usando un gel di poliaccrilammide. I microsatelliti analizzati nel presente lavoro sono: GABRB3, D15S108, LIPC, D13S115, D13S270. La distribuzione delle frequenze alleliche hanno mostrato differenze significative tra le isole. I 5 microsatelliti presentano un elevato potere di discriminazione (PD) e un elevato contenuto di informazione polimorfica (PIC). Il potere di discriminazione più alto è stato trovato per il GABRB3 e secondariamente per D13S115, che ha mostrato anche il più alto valore del PIC. I confronti con le altre popolazioni hanno evidenziato un netto raggruppamento delle popolazioni per continenti, in particolare gli europei mostrano una più ridotta eterogeneità genetica rispetto agli altri continenti. I risultati ottenuti dimostrano che tali marcatori sono utili nello studio della variabilità genetica delle popolazioni umane e nell'analisi forense.

Abstract

About 360 unrelated individuals from the three largest islands of the Western Mediterranean: Sardinia (Italy), Sicily (Italy) and Corsica (France) were examined. The DNA was amplified by means of PCR and the fragments were separated using a polyacrylamide gel. The microsatellites analysed in the present work are: GABRB3, D15S108, LIPC, D13S115, D13S270. The allele frequencies distribution showed significant differences between the islands. The 5 microsatellites have a high discrimination power (PD) and a high Polymorphic Information Content (PIC). The highest discrimination power was found for GABRB3 and secondarily for D13S115, which also showed the highest values of PIC. The comparison with other populations underlined the populations from the same continent cluster together. Particularly Europeans shows a less genetic heterogeneity than other continents. The results demonstrate that these microsatellites are useful in the study of genetic variability of the human populations and in forensic analysis.

Introduzione

Tra i marcatori che vengono usati nelle analisi genetiche, nelle diagnosi mediche e in medicina forense, gli STRs (Short Tandem Repeat polymorphisms) hanno recentemente assunto una grande importanza. Questi possono essere utilizzati con differenti scopi nel campo della genetica, e si sono dimostrati di grande aiuto sia nello studio della struttura genetica delle popolazioni sia in medicina forense. La variabilità mostrata da molti degli STRs li rende estremamente interessanti come marcatori antropologici (Harding, 1992).

Tecnicamente, gli STRs, in confronto con altri marcatori, presentano il vantaggio di poter essere analizzati tramite PCR e i loro alleli possono essere individuati senza ambiguità con un gel di poliacrilamide e non utilizzando il radioattivo. Questa relativa facilità di studio ha portato ad un rapido incremento nel numero degli STRs esaminati e delle popolazioni studiate.

Al fine di dare un contributo alla conoscenza della distribuzione delle frequenze alleliche degli STRs nelle popolazioni della Corsica (Francia), Sardegna e Sicilia (Italia), le tre maggiori isole del Mediterraneo Occidentale (Figura 1), sono stati studiati tre campioni provenienti dalle tre isole.



Figura 1. Localizzazione geografica delle popolazioni esaminate.

Figure 1. Geographical localization of the examined populations.

I cinque STRs utilizzati in questo studio sono microsatelliti dinucleotidi (CA)_n (GABRB3, D15S108, D13S115, D13S270) e (CT)_n (LIPC). La scelta di questi marcatori è stata suggerita dall'alto valore di eterozigosità e dall'alto contenuto di informazione polimorfica (PIC) riscontrato da vari autori (Bowcock, 1993; Bowcock *et al.*, 1993; Gyapay *et al.*, 1994) che li rendono adatti per l'utilizzo in genetica di popolazioni e in medicina forense (Bowcock *et al.*, 1994).

La Sicilia (Italia), la più grande isola del Mediterraneo Occidentale (25,707 km²), ha una densità di 193 abitanti per km². Per la sua particolare posizione geografica, la Sicilia è stata per generazioni teatro di numerose invasioni e dominazioni da parte di diverse popolazioni. Dall'arrivo nell'isola dei primi gruppi (Sicani, Elimi e Siculi) la Sicilia ha visto un'alternanza di Fenici, Greci, Romani, Vandali, Goti, Arabi, Normanni e Spagnoli. L'importanza del contributo genetico di questi gruppi sulla popolazione siciliana deve essere ancora chiarito (Vona *et al.*, 2000).

La popolazione della Sardegna (Italia) è un esempio di isolato genetico, come dimostrato dalla ben documentata unicità delle sue frequenze geniche (Piazza *et al.*, 1985; Modiano *et al.*,

1986; Walter et al., 1991; Vona et al., 1992). La Sardegna è rimasta isolata per lungo tempo in due modi: dal resto d'Italia e dalle regioni mediterranee ed all'interno del suo stesso territorio. La ragione di questo isolamento può essere spiegata dall'insularità e dalla posizione geografica, che non favoriscono vaste migrazioni, dalle condizioni geomorfologiche interne, che creano barriere tra le comunità frammentando la popolazione, e dalla scarsa densità di popolazione, che attualmente è 70 abitanti per km², in un territorio ampio (24.000 km²). Diversi studi sulla struttura genetica e matrimoniale, sulla consanguineità e sulla kinship dei cognomi hanno mostrato un alto grado di isolamento interno ed esterno (Cappello et al., 1996; Floris e Vona, 1984; Moroni et al., 1972, Vona et al., 1996; Vona, 1997).

La Corsica (Francia), la terza isola del Bacino del Mediterraneo (8,682 km²; 28.8 abitanti per km²), è caratterizzata da una storia estremamente complessa. L'isola fu sottoposta a svariati insediamenti da parte di differenti popolazioni: Liguri, Fenici, Greci, Romani, Bizantini, Pisani, e Francesi. Un alto grado di isolamento è stato suggerito dalle sue caratteristiche geomorfologiche, i suoi piccoli villaggi (spesso con meno di 400 abitanti), l'unicità delle sue frequenze geniche per alcuni polimorfismi e una forte eterogeneità se confrontata con le altre popolazioni mediterranee (Vona et al., 1995; Moral et al., 1996; Memmi et al., 1998; Varesi et al., 2000)

Materiali e Metodi

Sono stati esaminati 360 campioni provenienti dai centri trasfusionali e dagli ospedali della Corsica (Francia), Sardegna e Sicilia (Italia), precisamente 130 erano corsi del distretto di Corte e Ajaccio (centro e sud-ovest dell'isola), 130 siciliani provenienti dalle province di Palermo e Trapani (Sicilia occidentale) e 100 sardi delle province di Cagliari (sud Sardegna) e Nuoro (centro Sardegna)

Il campione è costituito da individui di entrambi i sessi, apparentemente sani, originari delle tre isole da almeno tre generazioni.

Il DNA è stato isolato da sangue intero, periferico, con il metodo tradizionale di digestione con proteinasi K ed estrazione con fenolo-cloroformio, nel Dipartimento di Biologia Sperimentale dell'Università di Cagliari (Italia).

L'amplificazione è stata condotta utilizzando i primers riportati in tabella 1.

GABRB3	L: 5'-CTCTTGTTCTCTGTTGCTTTCAATACAC-3' R: 5'-CACTGTGCTAGTAGATTCAGCTC-3'	(Beckmann et al., 1993)
D15S108	MFD102L: 5'-ATTCTTAACAGGAAGTGAGGG-3' MFD102R: 5'-AACATGAGTTTCAGAGGGG-3'	(Beckmann et al., 1993)
LIPC	HLIP1: 5'-ATGTGATGTCAGTGCTGCCAGTCCA-3' HLIP2: 5'-ACTGACATTTGAAAGATACGACCAC-3'	(Bhattacharrya et al., 1991)
D132115	MS34L: 5'-TGTAAGGAGAGAGAGATTTGACA-3' MS34R: 5'-TCTTAGCTGCTGGTGGTGG-3'	(Bowcock et al., 1993)
D13S270	MS34L: 5'-TGTAAGGAGAGAGAGATTTGACA-3' MS34R: 5'-TCTTAGCTGCTGGTGGTGG-3'	(Gyapay et al., 1994)

Tabella 1. Primers utilizzati per l'amplificazione e fonti bibliografiche

Table 1. Primers used for the amplification and references.

Tutti i microsatelliti sono stati amplificati in un volume totale di 12.5 µl contenente Buffer, MgCl₂ 1,5 mM, dNTP 2,5 mM ognuno, 10 µM di ogni primer e 0.6 unità di Taq polimerasi.

La PCR è stata portata avanti con i seguenti cicli: denaturazione iniziale a 94°C per 7', seguita da 30 cicli costituiti da denaturazione (94°C per 30"), ibridazione (55°C per 30") ed estensione (72°C per 30"). L'amplificazione è stata completata con un'estensione finale a 72°C per 7'. I prodotti amplificati sono stati sottoposti ad elettroforesi verticale in un gel di poliacrilamide all'8%. Dopo l'elettroforesi, le bande sono state visualizzate con una colorazione di nitrato d'argento. I genotipi sono stati determinati usando marcatori di peso molecolare disponibili in commercio e alcuni marcatori prodotti nel nostro laboratorio dalla combinazione di più campioni precedentemente sequenziati.

Analisi statistiche

Le frequenze alleliche sono state calcolate mediante la conta degli alleli. La possibile divergenza dall'equilibrio di Hardy-Weinberg è stata determinata con tre differenti metodi statistici: il likelihood ratio test, il test esatto della catena di Markov, in accordo con Guo e Thompson (1992) e l'Fis che dà una stima del coefficiente di inbreeding e rappresenta l'ampiezza della deviazione dall'equilibrio di Hardy-Weinberg

La stima dell'eterozigosità attesa è stata calcolata mediante il metodo descritto da Nei e Roychoudhury (1974). Il linkage disequilibrium per i loci situati sullo stesso cromosoma è stato verificato attraverso il test esatto della catena di Markov. E' stata anche studiata l'eterogeneità genetica tra le popolazioni esaminate seguendo il metodo di Fisher. Tutti i test sono stati eseguiti con il programma GENEPOP, versione 1.2 (Raymond e Rousset, 1995).

I valori del contenuto di informazione polimorfica (PIC) sono stati ottenuti in accordo con Weber (1990) e il potere di discriminazione (PD) è stato calcolato secondo la formula di Fisher (Fisher, 1951). I livelli di diversità genetica intra- e inter-popolazione sono stati stimati tramite l'eterozigosità e i coefficienti di Nei (1973) e la significatività dei valori del Gst è stata testata utilizzando il metodo suggerito da Paoli e Franceschi (1990).

Le distanze genetiche tra le popolazioni sono state calcolate con il metodo di Reynolds *et al.* (Reynolds *et al.*, 1983) attraverso il programma Phylip 3.5 (Felsenstein, 1989), utilizzando come confronti i dati gentilmente concessi dal Prof. L.L. Cavalli-Sforza e dal Prof. P. Moral per la tesi di dottorato di uno di noi (Calò, 1997). Per la costruzione del dendrogramma dalla matrice delle distanze genetiche è stato utilizzato il metodo neighbor-joining (Saitou e Nei 1987). Le popolazioni prese in esame per i confronti sono: EU1=Nord Italia; EU2=Nord Europa; BAS= Baschi; AF1=Pigmei Biaka dell'Africa centrale; AF2= Lisongo del Gabon; AF3= Pigmei Mbuti dello Zaire; AS1=Cina; AS2= Giappone; AS3= Cambogia; AM1= Maya dello Yucatan, Messico; AM2= Karitiana dell'Amazzonia, Brasile; AM3= Surui dell'Amazzonia, Brasile; OC1= Melanesiani di Bouganville, Isole Salomone; OC2= Nuova Guinea; OC3= Australia.

Risultati

Distribuzione delle frequenze alleliche

Le frequenze alleliche dei 5 STRs esaminati nelle popolazioni delle tre isole mediterranee con i loro errori standard sono mostrate nella tabella 2.

Locus GABRB3

Per il microsatellite GABRB3 sono stati trovati complessivamente 13 alleli, con alcune particolarità, e la taglia degli alleli varia da 175 a 210. La distribuzione delle frequenze appare estremamente variabile, ma in tutte le popolazioni l'allele più frequente è risultato il 179, seguito dal 181 nei sardi e dal 183 nei Siciliani e nei Corsi. La distribuzione delle frequenze appare bimodale in Sardegna e in Sicilia e multimodale in Corsica. Nella popolazione sarda è stata riscontrata la presenza di un allele raro, il 175.

Locus D15S108

La taglia degli alleli per il marcatore D15S108 ha un range di 141-163 e sono stati trovati 12 alleli, sebbene alcuni di essi (141, 147, 149) presentano frequenze molto basse e sono presenti solo in alcune popolazioni. L'allele 157 si è dimostrato essere il più frequente in tutte le popolazioni, con una frequenza massima (0.583) nel campione siciliano. La distribuzione delle frequenze appare bimodale con i picchi più alti in corrispondenza degli alleli 143 e 157.

Locus LIPC

Il marcatore LIPC ha un range di 165-177 bp e mostra 6 differenti alleli, con una distribuzione bimodale e con i picchi maggiori sugli alleli 167 e 171 in tutte le popolazioni. L'allele con frequenze maggiori in tutte le popolazioni esaminate è risultato il 167, con un range di frequenza tra 0.567 nel campione siciliano e 0.649 nel campione sardo.

Locus D13S115

Per il locus D13S115 sono stati trovati complessivamente 8 alleli (161-175 bp). L'allele 165 è risultato essere il più frequente in tutte le popolazioni esaminate. La distribuzione delle frequenze appare unimodale.

Locus D13S270

Sono stati riscontrati 9 alleli per il locus D13S270, con una taglia di 79-97. La distribuzione delle frequenze, sebbene variabile nelle diverse popolazioni, risulta bimodale, con i picchi sugli alleli 79 e 89. Alcuni alleli mostrano frequenze estremamente basse e sono presenti solo in alcune popolazioni: gli alleli 81, 85 e 97 solo nel campione sardo e l'allele 93 è stato trovato solo nei corsi. Non è stato trovato l'allele 83 nei nostri campioni.

GABRB3	Sardegna	Sicilia	Corsica	LIPC	Sardegna	Sicilia	Corsica
175	0.005 ± 0.005	0.000 ± 0	0.000 ± 0	165	0.102 ± 0.023	0.133 ± 0.024	0.093 ± 0.019
179	0.437 ± 0.048	0.442 ± 0.041	0.434 ± 0.042	167	0.649 ± 0.058	0.567 ± 0.049	0.613 ± 0.050
181	0.116 ± 0.025	0.085 ± 0.018	0.053 ± 0.015	169	0.015 ± 0.009	0.033 ± 0.012	0.061 ± 0.016
183	0.074 ± 0.020	0.100 ± 0.020	0.094 ± 0.020	171	0.224 ± 0.034	0.212 ± 0.030	0.202 ± 0.029
185	0.037 ± 0.014	0.085 ± 0.018	0.041 ± 0.013	173	0.010 ± 0.007	0.050 ± 0.014	0.032 ± 0.011
187	0.032 ± 0.013	0.031 ± 0.011	0.049 ± 0.014	177	0.000 ± 0	0.004 ± 0.004	0.000 ± 0
189	0.021 ± 0.011	0.023 ± 0.009	0.012 ± 0.007	D13S115	Sardegna	Sicilia	Corsica
191	0.079 ± 0.020	0.062 ± 0.015	0.074 ± 0.017	161	0.021 ± 0.010	0.030 ± 0.011	0.025 ± 0.010
193	0.089 ± 0.022	0.062 ± 0.015	0.090 ± 0.019	163	0.021 ± 0.010	0.004 ± 0.004	0.025 ± 0.010
195	0.063 ± 0.018	0.054 ± 0.014	0.037 ± 0.012	165	0.546 ± 0.053	0.515 ± 0.044	0.425 ± 0.042
197	0.026 ± 0.012	0.058 ± 0.015	0.078 ± 0.018	167	0.201 ± 0.032	0.263 ± 0.031	0.271 ± 0.034
199	0.005 ± 0.005	0.019 ± 0.009	0.012 ± 0.007	169	0.129 ± 0.026	0.078 ± 0.017	0.175 ± 0.027
201	0.016 ± 0.009	0.008 ± 0.006	0.025 ± 0.010	171	0.026 ± 0.012	0.026 ± 0.010	0.058 ± 0.016
D15S108	Sardegna	Sicilia	Corsica	173	0.046 ± 0.015	0.063 ± 0.015	0.017 ± 0.008
141	0.000 ± 0	0.004 ± 0.004	0.000 ± 0	175	0.010 ± 0.007	0.022 ± 0.009	0.017 ± 0.008
143	0.085 ± 0.021	0.150 ± 0.025	0.141 ± 0.023	D13S270	Sardegna	Sicilia	Corsica
145	0.030 ± 0.012	0.013 ± 0.007	0.050 ± 0.014	79	0.364 ± 0.044	0.468 ± 0.043	0.446 ± 0.043
147	0.000 ± 0	0.000 ± 0	0.011 ± 0.006	81	0.005 ± 0.005	0.000 ± 0	0.000 ± 0
149	0.005 ± 0.005	0.000 ± 0	0.000 ± 0	85	0.038 ± 0.014	0.000 ± 0	0.000 ± 0
151	0.015 ± 0.009	0.008 ± 0.006	0.000 ± 0	87	0.190 ± 0.032	0.039 ± 0.012	0.004 ± 0.004
153	0.050 ± 0.016	0.046 ± 0.014	0.031 ± 0.011	89	0.386 ± 0.046	0.205 ± 0.028	0.244 ± 0.032
155	0.115 ± 0.024	0.083 ± 0.019	0.107 ± 0.020	91	0.011 ± 0.008	0.276 ± 0.033	0.285 ± 0.034
157	0.580 ± 0.054	0.583 ± 0.049	0.542 ± 0.045	93	0.000 ± 0	0.000 ± 0	0.004 ± 0.004
159	0.090 ± 0.021	0.088 ± 0.019	0.080 ± 0.017	95	0.000 ± 0	0.008 ± 0.006	0.017 ± 0.008
161	0.020 ± 0.010	0.025 ± 0.010	0.027 ± 0.010	97	0.005 ± 0.005	0.000 ± 0	0.000 ± 0
163	0.010 ± 0.007	0.000 ± 0	0.012 ± 0.007				

Tabella 2. Frequenze alleliche dei 5 marcatori esaminati con relativi errori standard.

Table 2. Allele frequencies for the 5 markers examined with standard errors.

Tests per l'equilibrio di Hardy-Weinberg

Al fine di analizzare la struttura genetica delle popolazioni studiate per i loci esaminati, sono stati eseguiti tre differenti tests statistici. I risultati sono mostrati nella tabella 3.

GABRB3	Sardegna	Sicilia	Corsica
Eterozigosità osservata	64.2	70.0	79.5
Eterozigosità attesa	77.2	77.1	77.7
Likelihood ratio test	p<0.01	n.s.	n.s.
Catena di Markov (p)	0.017	0.056	0.650
Fis	+0.169	+0.093	-0.024
PIC (%)	75.2	75.3	75.8
PDo (%)	90.3	92.2	92.5
PDa (%)	93.0	93.1	93.3
<hr/>			
D15S108			
Eterozigosità osservata	57.0	60.8	66.4
Eterozigosità attesa	63.4	62.2	66.7
Likelihood ratio test	n.s.	n.s.	n.s.
Catena di Markov (p)	0.061	0.864	0.093
Fis	+0.102	+0.022	+0.004
PIC (%)	60.8	59.2	63.8
PDo (%)	71.6	82.7	85.2
PDa (%)	84.1	82.7	86.1
<hr/>			
LIPC			
Eterozigosità osservata	53.1	50.8	57.3
Eterozigosità attesa	52.2	61.5	54.8
Likelihood ratio test	n.s.	p<0.05	n.s.
Catena di Markov (p)	0.435	0.001	0.001
Fis	-0.017	+0.174	+0.043
PIC (%)	46.6	57.1	53.5
PDo (%)	70.6	79.3	75.8
PDa (%)	71.6	80.5	77.4
<hr/>			
D13S115			
Eterozigosità osservata	59.8	68.2	65.8
Eterozigosità attesa	64.4	65.6	71.3
Likelihood ratio test	n.s.	n.s.	n.s.
Catena di Markov (p)	0.690	0.722	0.426
Fis	+0.072	-0.039	+0.078
PIC (%)	60.3	61.0	66.5
PDo (%)	83.1	83.2	81.5
PDa (%)	83.3	83.5	81.8
<hr/>			
D13S270			
Eterozigosità osservata	67.4	63.0	66.1
Eterozigosità attesa	68.4	66.4	66.3
Likelihood ratio test	n.s.	n.s.	n.s.
Catena di Markov (p)	0.525	0.136	0.495
Fis	+0.015	+0.051	+0.002
PIC (%)	62.1	60.2	59.4
PDo (%)	85.1	83.0	87.7
PDa (%)	83.7	82.6	87.1

Tabella 3. Test di Hardy-Weinberg, PIC e PD nelle popolazioni esaminate (o=osservate, a=attese)

Table 3. Hardy-Weinberg tests, PIC and PD on the examined populations

(o=observed, a=expected)

E' stata riscontrata una deviazione significativa dall'equilibrio tramite il test della catena di Markov e il likelihood ratio test nel campione sardo per il solo locus GABRB3 e nel campione siciliano per il locus LIPC. Inoltre, una deviazione dell'equilibrio di HW è stata suggerita dal solo test della catena di Markov per il locus LIPC nel campione corso.

I valori di Fis sono in accordo con i risultati ottenuti dagli altri test per l'equilibrio di HW e confermano le deviazioni significative nei Sardi per il locus GABRB3 e nei Siciliani per il LIPC presentando le due popolazioni rispettivamente il 7% e il 10,7% di omozigoti osservati superiore a quelli attesi. Questo dato potrebbe trovare una spiegazione nel campionamento effettuato, infatti, tra le aree sarde e siciliane analizzate vi sono alcuni centri (es. Alia in Sicilia, Nuorese in Sardegna) per i quali sono stati evidenziati fenomeni di isolamento e consanguineità, che potrebbero aver favorito l'incremento dell'omozigotità.

Poiché due dei cinque marcatori studiati sono localizzati sul cromosoma 13 (D13S115 e D13S270), mentre gli altri tre (GABRB3, D15S108 e LIPC) sono localizzati sul cromosoma 15, è stata verificata l'esistenza di linkage disequilibrium per i loci situati sullo stesso cromosoma. I risultati non hanno dato valori significativi per nessuna popolazione, perciò i loci non appaiono associati in linkage.

Anche i valori del PIC e del PD sono mostrati nella tabella 3. I valori ottenuti per il PIC variano da un locus all'altro e da una popolazione all'altra. I valori medi per locus hanno un range compreso tra 52.4% del LIPC e il 75.4% del GABRB3. Si può notare che tutti i loci presentano un alto valore di PD, che varia tra il 75% e l'85% rispettivamente per il LIPC e per il GABRB3.

Differenziazione tra le popolazioni

E' stata studiata la differenziazione per coppie di popolazioni e per l'insieme di esse per ciascun locus e poi combinando i diversi loci, usando il test esatto di Fisher. Le popolazioni studiate sono risultate tutte fortemente differenziate, in particolare la Sardegna e la Sicilia (Tabella 4).

	GABRB3	D15S108	LIPC	D13S115	D13S270
Sardegna-Sicilia	0.539	0.761	0.046	0.018	0.000
Sardegna-Corsica	0.149	0.411	0.108	0.329	0.000
Sicilia-Corsica	0.727	0.374	0.174	0.008	0.027
Tutte le popolazioni	0.411	0.520	0.063	0.008	0.000

Test di combinazione: altamente significativo

Tabella 4. Valori di probabilità del test esatto di Fisher per la differenziazione tra coppie di popolazioni e considerando tutte le popolazioni nel loro insieme.

Table 4. Probability of Fisher exact test for differentiation between populations pairs and considering all populations together.

La differenziazione è dovuta prevalentemente alla distribuzione allelica dei loci D13S270 e D13S115. La differenza che emerge tra Sardegna e Corsica è data esclusivamente dal D13S270. Infatti quando si riapplica il test esatto di Fisher escludendo il D13S270, si nota un'assenza di eterogeneità tra le due popolazioni ($p > 0.151$).

Le tre popolazioni del Mediterraneo studiate in questo lavoro sono state successivamente messe a confronto con altre popolazioni appartenenti sia all'Europa sia ad altri continenti.

Per stimare il grado di variabilità tra le popolazioni, sono stati analizzati i dati di 18 campioni dell'Europa, dell'America, dell'Africa, dell'Asia e dell'Oceania per verificare l'influenza relativa delle suddivisioni etniche intra e inter-popolazioni attraverso i coefficienti di Nei (1973). La diversità genica per locus nelle tre popolazioni esaminate varia da 0.57 del LIPC a 0.77 del GABRB3. Gli stessi loci segnano i limiti della variabilità di Hs anche negli altri due campioni di Europei e nel complesso delle 18 popolazioni considerate. La media della diversità genica (Hs) interna alle popolazioni della Corsica, Sardegna e Sicilia (Tabella 5) è simile sia a quella degli Europei sia a quella delle popolazioni americane, mentre appare maggiore rispetto all'Hs delle popolazioni oceaniche e inferiore a quello dell'Africa.

I valori di Hs osservati per le popolazioni da noi studiate e per gli altri due campioni di Europei sono inferiori al range (0.71-0.74) che viene riportato da Da Silva *et al.* (1999) e ricavato dalla letteratura per i loci ipervariabili. Ma i nostri valori coincidono con quelli riportati da Jorde *et al.* (2000) per gli Europei.

E' interessante notare come le popolazioni africane abbiano un maggiore livello di diversità genica, al contrario delle popolazioni dell'Oceania che appaiono caratterizzate dal valore più basso.

	n. pop.	Hs	Ht	Gst	Gst'	
Tutte le popolaz.	17	0.635	0.761	0.164	0.176	p<0.001
America	3	0.618	0.668	0.075	0.108	n.s.
Africa	3	0.701	0.775	0.063	0.137	p<0.01
Asia	3	0.634	0.669	0.052	0.076	n.s.
Oceania	3	0.569	0.630	0.096	0.139	n.s.
Europa	2	0.619	0.636	0.025	0.052	n.s.
Euro-Med.	5	0.645	0.668	0.035	0.043	p<0.01
Sar.-Sic.-Cor.	3	0.662	0.671	0.013	0.020	p<0.005

Tabella 5. Valori degli indici dei Nei.

Table 5. Nei's index values.

L'ampiezza relativa della differenziazione genica (Gst') assume un valore globale nelle 14 popolazioni pari a 0.18, molto simile a quello riportato da Livshits e Nei (1990) nello studio di vari continenti utilizzando i marcatori genetici classici.

Il Gst' ottenuto per le nostre tre popolazioni, relativo al complesso dei loci analizzati, è lievemente inferiore a quello ottenuto negli altri Europei, ma è notevolmente più basso di quello calcolato nei gruppi di popolazioni degli altri continenti che raggiungono con l'Africa e l'Oceania valori di 0.137 e 0.139 rispettivamente.

I valori di Gst (tabella 5) appaiono significativi secondo il test proposto da Paoli e Franceschi (1990) per le popolazioni delle tre isole da noi studiate, per le popolazioni europee nel loro complesso, per l'Africa e per tutte le popolazioni introdotte nel confronto considerate nel loro insieme.

Successivamente, le relazioni tra le 18 popolazioni sono state analizzate anche tramite le distanze genetiche secondo il metodo suggerito da Reynolds et al. (1983) che considera la deriva genetica casuale come fattore principale di differenziazione.

Dalle distanze genetiche nella tabella 6, appare, come ci si attendeva, che i campioni provenienti dalle tre isole del Mediterraneo sono molto vicini alle altre popolazioni europee. Le distanze minori risultano essere quelle dei confronti tra coppie di popolazioni europee e mediterranee, mentre molto più elevate appaiono le distanze all'interno degli altri continenti.

L'albero filogenetico ottenuto con le distanze di Reynolds e costruito con il metodo Neighbor-joining (Saitou e Nei, 1987) è mostrato nella figura 2. La robustezza dell'albero è stata testata ricostruendo 100 alberi attraverso il metodo del bootstrap (Efron, 1982). L'albero è composto da tre clusters principali: uno costituito dalle popolazioni europee, un secondo formato da due subclusters nei quali sono disposte le popolazioni dell'Africa e dell'Asia ed infine il terzo costituito da due subclusters che discriminano le popolazioni americane ed oceaniche. I nodi principali presentano valori di bootstrap superiori al 50%.

	Sar.	Sic.	Cor.	Bas.	Eul1	Eul2	Am1	Am2	Am3	Af1	Af2	Af3	Oc1	Oc2	Oc3	As1	As2	As3
Sar	-																	
Sic	0.025	-																
Cor	0.030	0.009	-															
Bas	0.043	0.017	0.027	-														
Eul1	0.060	0.070	0.093	0.076	-													
Eul2	0.024	0.035	0.053	0.058	0.050	-												
Am1	0.149	0.128	0.132	0.135	0.139	0.175	-											
Am2	0.237	0.203	0.309	0.226	0.275	0.259	0.248	-										
Am3	0.168	0.141	0.152	0.161	0.165	0.186	0.168	0.171	-									
Af1	0.152	0.149	0.143	0.154	0.175	0.186	0.423	0.524	0.346	-								
Af2	0.182	0.171	0.173	0.173	0.185	0.203	0.430	0.428	0.358	0.247	-							
Af3	0.238	0.215	0.208	0.225	0.227	0.258	0.389	0.544	0.327	0.253	0.276	-						
Oc1	0.203	0.206	0.208	0.212	0.196	0.218	0.236	0.364	0.341	0.492	0.548	0.540	-					
Oc2	0.189	0.166	0.161	0.196	0.202	0.212	0.269	0.301	0.282	0.452	0.520	0.433	0.224	-				
Oc3	0.172	0.156	0.166	0.145	0.180	0.188	0.248	0.184	0.212	0.379	0.413	0.439	0.153	0.164	-			
As1	0.201	0.191	0.199	0.212	0.215	0.219	0.179	0.243	0.157	0.203	0.171	0.270	0.263	0.260	0.230	-		
As2	0.203	0.201	0.197	0.215	0.203	0.242	0.167	0.284	0.182	0.178	0.163	0.201	0.241	0.234	0.232	0.096	-	
As3	0.175	0.169	0.172	0.194	0.192	0.195	0.142	0.220	0.136	0.189	0.172	0.229	0.152	0.176	0.170	0.047	0.083	-

Tabella 6. Matrice delle distanze genetiche secondo Reynolds.
Table 6. Matrix of genetic distances in according to Reynolds.

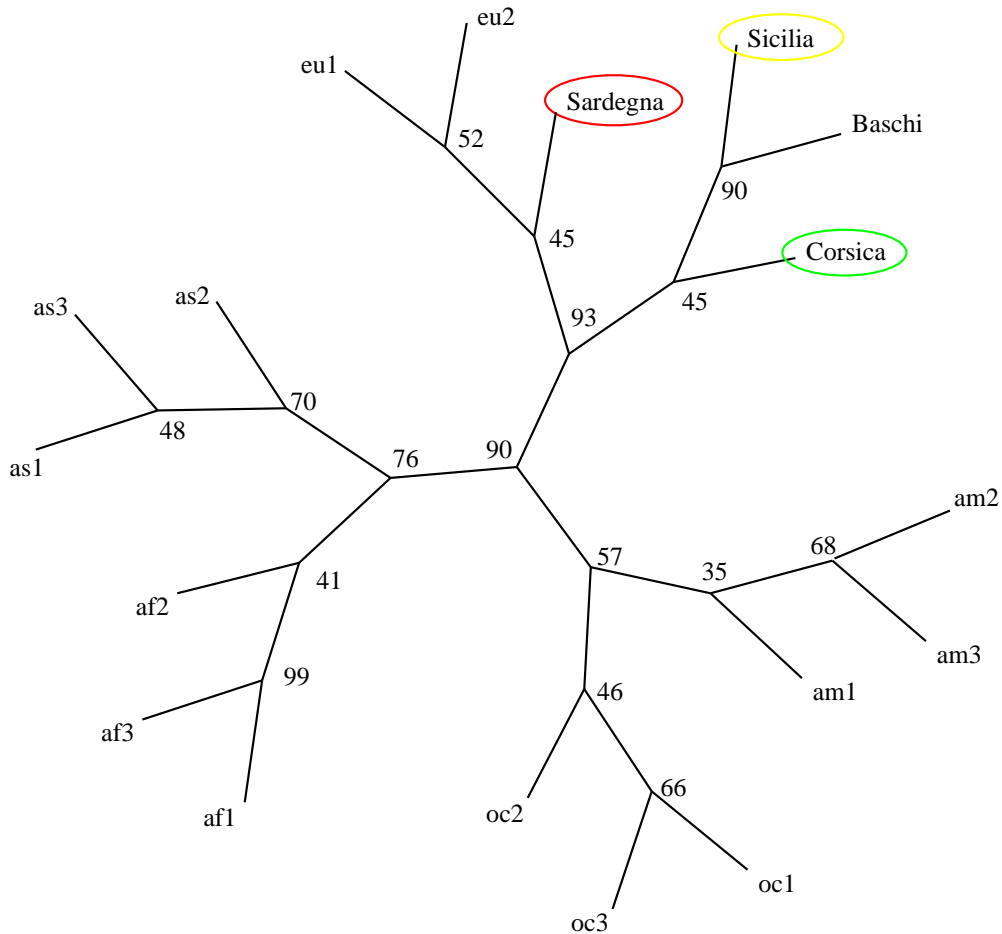


Figura 2. Albero genetico costruito con il metodo Neighbor-joining, ricavato dalla matrice di tabella 6.
Figure 2. Neighbor-joining tree, obtained from matrix reported on table 6.

Discussione

Lo spettro delle frequenze alleliche delle nostre popolazioni mostra per quasi tutti i loci delle distribuzioni bimodali. I campioni appaiono in generale in un buon adattamento all'equilibrio di Hardy-Weinberg. L'alto grado di eterozigosità (dal 50.8% al 79.5%) potrebbe essere ascritto all'elevato tasso di mutazione caratteristico degli STR, che incrementa la variazione intragruppo rispetto a quella tra gruppi portando ad un decremento dei valori di G_{st} (Jin e Chakraborty, 1995).

Il rapporto tra H_s , l'eterozigosità all'interno delle sottopopolazioni, e H_t , l'eterozigosità per il campione complessivo, senza tenere conto delle suddivisioni, dimostra che il 98% della diversità totale tra le popolazioni da noi studiate può essere attribuita alla variabilità intrapopolazione, mentre risulta dell'83% la variabilità interna, se si considerano anche le popolazioni del confronto.

Le relazioni tra diversità genetica intra- e inter-popolazione studiata mediante il coefficiente di differenziazione genetica G_{st} sembrano indicare l'esistenza di una eterogeneità genetica tra le popolazioni delle tre isole. Anche il confronto operato tra le popolazioni mediate il test di eterogeneità, come si è detto, appare molto significativo. Le popolazioni Corsa, Sarda e Siciliana hanno in comune molti degli alleli trovati nei cinque loci esaminati. Ma non sempre le frequenze maggiori spettano agli stessi alleli. Inoltre, vi sono alleli che vengono esibiti esclusivamente da una sola delle popolazioni. La Sardegna presenta 5 alleli dei loci D15S108, GABRB3 e D13S270 che non compaiono tra i Corsi e i Siciliani. Queste due ultime popolazioni, a loro volta, hanno due

alleli esclusivi. La maggior parte di questi alleli compaiono, comunque, in altre popolazioni, solo il 175 del GABRB3, ritrovato in un solo individuo sardo, non è mai stato descritto in bibliografia. La presenza di alleli rari non è un evento straordinario ed è spiegato dall'elevato tasso di mutazione che caratterizza gli STR.

La presenza di numerosi alleli comuni nelle tre popolazioni testimoniano la loro origine comune. Le differenze nelle frequenze alleliche possono essere il risultato dell'effetto fondatore e della deriva genetica, fattori in grado di condizionare la struttura genetica di popolazioni relativamente isolate, come, ad esempio, la Sardegna e la Corsica (Memmi *et al.*, 1998; Vona *et al.* 1992; Vona *et al.*, 1994). Ma queste differenze possono essere dovute anche alle mutazioni che creano nuovi alleli o reintrodurre in una popolazione isolata alleli che si erano persi.

Gli Europei sono caratterizzati all'interno del loro cluster da branche più brevi di quelle che si riscontrano negli altri clusters, indicando una loro più ridotta eterogeneità genetica rispetto a quella presente all'interno delle popolazioni degli altri continenti. Inoltre, i valori di bootstrap all'interno di questo cluster non appaiono così significativi da permettere delle considerazioni sulle associazioni che vengono mostrate.

Sulle basi del modello di un tasso costante di evoluzione le distanze attese dovrebbero essere uguali tra tutte le popolazioni non africane e africane.

Le distanze genetiche medie indicano che le popolazioni europee sono quasi equidistanti da quelle degli altri continenti, pur risultando lievemente più prossimi alle popolazioni africane di quanto non lo siano le asiatiche. Molto più elevate appaiono le distanze genetiche dell'America e dell'Oceania rispetto all'Africa. I risultati potrebbero essere interpretati come una derivazione degli Europei da un mescolamento tra popolazioni africane e asiatiche ancestrali, sostenuto da numerosi altri autori (Cavalli Sforza *et al.* 1994; Matullo *et al.*, 1997; Fernandez-Santander *et al.*, 2001).

Lo studio dei 5 STRs ha permesso di apportare nuovi contributi alla conoscenza della struttura genetica della Corsica, Sardegna e Sicilia che sinora erano state studiate prevalentemente mediante i marcatori genetici classici.

Le frequenze dei 5 loci differenziano in modo netto le popolazioni che appartengono ai diversi continenti. Questi marcatori si sono rivelati, in generale, efficaci anche per lo studio delle popolazioni umane soprattutto a livello macrogeografico essendo in grado di differenziare in modo netto popolazioni appartenenti a continenti diversi. Inoltre, il buon adattamento all'equilibrio genico riscontrato rende questi marcatori utili non solo nelle applicazioni comuni alla genetica delle popolazioni, ma anche in medicina forense, dato l'elevato numero di alleli che presentano e il loro considerevole contenuto informativo (PIC) e discriminante (PD).

Bibliografia

- Beckmann, J.S., Tomfohrde, J., Barnes, R.I., Williams, M., Broux, O., Richards, I., Weissenbach, J., e Bowcock, A.M., 1993, A linkage map of human chromosome 15 with an average resolution of 2 cM and containing 55 polymorphic microsatellites. *Hum. Mol. Genet.*, 2, 2019-2030.
- Bhattacharya, S., Ameis, D., Cullen, P., Narcisi, T.M., Bayliss, J., Greten, H., Schotz, M.C., e Scott, J., 1991, VNTR polymorphism in the hepatic lipase gene (LIPC). *Nucl. Acids Res.*, 19, 5088.
- Bowcock, A., Osborne-Lawrence, S., Barnes, R., Chakravarti, A., Washington, S., e Dunn, C., 1993, Microsatellite Polymorphism Linkage Map of human Chromosome 13q. *Genomics* 15, 376-386.
- Bowcock, A.M., 1993, A linkage map of human chromosome 15 with an average resolution of 2 cM and containing 55 polymorphic microsatellites. *Hum. Mol. Genet.*, 2, 2019-2030.
- Bowcock, A.M., Ruiz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J.R., e Cavalli-Sforza, L.L., 1994, High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* 368, 455-457.
- Calò, C.M., 1997, Variabilità genetica mediante VNTR nelle tre maggiori isole del Mediterraneo Occidentale. Ph.D. thesis, University of Cagliari, Italy.

- Cappello, N., Rendine, S., Griffo, R., Mameli, G.E., Succa, V., Vona, G., e Piazza, A., 1996, Genetic analysis of Sardinia: I. Data on 21 linguistic domains. *Ann. Hum. Genet.*, 60, 125-141.
- Cavalli-Sforza, L.L., Menozzi, P., e Piazza, A., 1994, *History and Geography of Human genes.* (Princeton: Princeton University Press).
- Da Silva, W.A., Bortolini, M.C., Meyer, D., Salzano, F.M., Elion, J., Krishnamoorthy, R., Cruz Schneider, M.P., Castro De Guerra, D., Layrisse, Z., Mendez Castellano, H., De Azevedo Weimer, T., e Zago, M.A., 1999, Genetic Diversity of two African and sixteen south American populations determined on the basis of six hypervariable loci. *Am. J. Phys Anthr.*, 109, 425-437.
- Efron, B., 1982, *The jackknife, the bootstrap, and other resampling plans.* CBMS-NSF Regional Conference Series in Applied Mathematics. Monograph 38. Society of Industrial and Applied Mathematics, Philadelphia.
- Felsenstein, J., 1989, PHYLIP - Phylogeny inference package (Version 3.2). *Clastics*, 5, 164-166.
- Fernández-Santander, A., Luna, F., Villegas, G., e Moral, P., 2001, Seventy-three nuclear DNA polymorphisms in a Spanish population and a comparison with ten other populations. *Am. J. Hum. Biol.*, 13, 212-219.
- Fisher, R., 1951, Standard calculation for evaluating a blood group system. *Heredity*, 5, 95-102.
- Floris, G., e Vona, G., 1984, Calcolo dell'isonimia e del coefficiente di relazione in nove comuni sardi. *Antrop. Contemp.*, 7, 17-22.
- Guo, S.W., e Thompson, E.A., 1992, Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, 48, 361-372.
- Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lanthrop, M., e Weissenbach, J., 1994, The 1993-94 Génethon human genetic linkage map. *Nature Genetics*, 7, 304-307.
- Harding, R.M., 1992, VNTR in Review. *Evol. Anthropol.*, 1, 62-71.
- Jin, L., e Chakraborty, R., 1995, Population structure, stepwise mutations, heterozygote deficiency and their implications in DNA forensics. *Heredity*, 74, 247-285.
- Jorde, L.B., Watkins, W.S., Banskad, M.J., Dixon, M.E., Ricker, C.E., Seielstad, M.T., e Batzer, M.A., 2000, The distribution of human genetic Diversity: a comparison of mitochondrial, autosomal and Y chromosome data. *Am. J. Hum. Genet.*, 66, 979-988.
- Livshits, G., e Nei, M., 1990, Relationships between intrapopulation and interpopulation genetic diversity in man. *Ann Hum. Biol.*, 17, 501-513.
- Matullo, G., Griffo, R.M., Mountain, J.L., Calafell, F., Guarrera, S., Piazza, A., e Cavalli-Sforza, L.L., 1997, Seventy-five nuclear DNA polymorphisms in an Italian sample: A comparative worldwide study. *Gene Geography*, 11, 15-35.
- Memmi, M., Moral, P., Calò, C.M., Autuori, L., Mameli, G.E., Succa, V., Varesi, L., e Vona, G., 1998, The genetic structure of south-western Corsica (France). *Am.J.Hum.Biol.*, 10, 567-577.
- Modiano, G., Terrenato, L., Scozzari, R., Santachiara Benerecetti, S.A., Ulizzi, L., Santolamazza, C., Petrucci, R., e Santolamazza, P., 1986, Population Genetics in Sardinia. *Atti Acc. Lincei Memorie*, 18, 257-330.
- Moral, P., Memmi, M., Varesi, L., Mameli, G.E., Succa, V., Gutierrez, B., Lutken, N., e Vona, G., 1996, Study on the variability of seven serum protein markers in Corsica (France). *Anthrop. Anz.*, 54, 97-107.
- Moroni, A., Anelli, A., Anghinetti, W., Lucchetti, E., Rossi, O., e Siri, E., 1972, La consanguineità umana nell'isola di Sardegna dal secolo XVII al secolo XX. *Ateneo Parmense*, 8, 69-92.
- Nei, M., 1973, Analysis of Gene Diversity in Subdivided Populations. *Proc. Nat. Acad. Sci. USA*, 70, 3321-3323.
- Nei, M., e Roychoudhury, A.K., 1974, Sampling variance of heterozygosity and genetic distance. *Genetics*, 76, 379-390.
- Paoli, G., e Franceschi, M.G., 1990, Genetic studies in the Garfagnana population. *Anthrop. Anz.*, 48, 333-345.

- Piazza, A., Mayr, W.R., Contu, L., Amoroso, A., Borelli, I., Curtoni, E.S., Marcello, C., Moroni, A., Olivetti, E., Richiardi, P., e Ceppellini, R., 1985, Genetic and population structure of four Sardinian Villages. *Ann. Hum. Genet.*, 45, 47-63.
- Raymond, M., e Rousset, F., 1995, GENEPOP (Version 1.2): Population genetics software for exact tests and ecumenicism. *J. Hered.*, 86, 248-249.
- Reynolds, J., Weir, B.S., e Cockerham, C.C., 1983, Estimation of the coancestry coefficient. Basis for a short-term genetic distance. *Genetics*, 105, 767-779.
- Saitou, N., e Nei, M., 1987, The neighbor-joining method; a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, 4, 406-425.
- Varesi, L., Memmi, M., Cristofari, M.C., Mameli, G.E., Calò, C.M., e Vona, G., 2000, Mitochondrial control-region sequence variation in the Corsican population, France. *Am. J. Hum. Biol.*, 12, 339-351.
- Vona, G., 1997, The peopling of Sardinia (Italy): history and effects. *Int. J. Anthrop.*, 12, 71-87.
- Vona, G., Chiarelli, B., Ghiani, M.E., e Sineo, L., 2000, Genetic structure of Sicily: A review. *Biennial Book of EAA. C. Susanne e E.B. Bodzsar Ed.*, 1, 63-78 (Budapest: Eötvös University Press).
- Vona, G., Floris-Masala, R., Mameli, G.E., e Succa, V., 1992, Red cell and serum protein polymorphisms in Sardinia. *Int. J. Anthrop.*, 7, 25-33.
- Vona, G., Francalacci, P., Paoli, G., Latini, V., e Salis, M., 1996, Study of the matrimonial structure of the population of Central Sardinia (Italy). *Anthrop. Anz.*, 54, 317-329.
- Vona, G., Memmi, M., Varesi, L., Mameli, G.E., e Succa, V., 1995, A study of several genetic markers in the Corsican population (France). *Anthrop. Anz.*, 53, 125-132.
- Walter, H., Matsumoto, H., e De Stefano, G.F., 1991, Gm and Km allotypes in four Sardinian population samples. *Am. J. Phys. Anthrop.*, 86, 45-50.
- Weber, J.L., 1990, Informativeness of human (dC-dA)n (dG-dT)n polymorphisms. *Genomics*, 7, 524-530.