



Università degli Studi di Cagliari
Dipartimento di Biologia Sperimentale
Sezione di Scienze Antropologiche

**DOTTORATO DI RICERCA IN BIOLOGIA E BIOCHIMICA
DELL'UOMO E DELL'AMBIENTE
CICLO XXIII**

L'isolamento degli isolati sardi: Carloforte e Benetutti

Settore scientifico disciplinare di afferenza: BIO/08 Antropologia

Presentata da:	Dott.ssa Laura Corrias
Coordinatore Dottorato	Prof. Emanuele Sanna
Relatore	Dott.ssa Carla Maria Calò

Esame finale anno accademico 2009 - 2010

Indice

1. Introduzione	- 1 -
1.1 Le comunità isolate	- 1 -
1.2 Gli isolati in Sardegna	- 3 -
1.3 Indicatori biodemografici dell'isolamento	- 4 -
1.4 Descrizione delle popolazioni analizzate.....	- 9 -
1.4.1 Benetutti.....	- 9 -
1.4.2 Carloforte	- 11 -
1.5 Utilizzo dei Marcatori Uniparentali	- 15 -
1.6 Cromosoma Y	- 15 -
1.6.1 Marcatori del cromosoma Y.....	- 17 -
1.6.2 Aplogruppi del cromosoma Y.....	- 18 -
1.7 DNA mitocondriale	- 24 -
1.7.1 Aplogruppi del DNA mitocondriale	- 28 -
1.8 Scopo del lavoro	- 32 -
2. Materiali e metodi.....	- 33 -
2.1 Campionamento.....	- 33 -
2.2 Estrazione del DNA da saliva	- 34 -
2.3 Estrazione del DNA da sangue intero.....	- 35 -
2.4 Determinazione quantitativa del DNA	- 36 -
2.5 Amplificazione	- 37 -
2.6 Analisi del DNA mitocondriale	- 38 -
2.7 Analisi del cromosoma Y	- 41 -
2.8 Digestione.....	- 43 -
2.9 Elettroforesi su gel di agarosio.....	- 44 -
2.10 Elettroforesi su gel di poliacrilammide	- 45 -
2.11. Analisi statistica.....	- 46 -
2.11.1 Frequenze alleliche e aplotipiche.....	- 46 -
2.11.2 Analisi degli aplotipi	- 47 -
2.11.3 Network filogenetici.....	- 48 -
2.11.4 Distanze genetiche	- 48 -

2.11.5 Albero filogenetico.....	- 49 -
2.11.6 Struttura genetica	- 50 -
3. Risultati.....	- 51 -
3.1 Risultati Cromosoma Y	- 51 -
3.1.1. Frequenze alleliche.....	- 51 -
3.1.2 Analisi degli aplogruppi	- 55 -
3.1.3 Network filogenetici.....	- 58 -
3.1.4 Distanze ed albero filogenetico.....	- 61 -
3.2 Risultati DNA mitocondriale	- 64 -
3.2.1. Analisi delle sequenze della regione ipervariabile I e II.	- 64 -
3.2.2 Analisi degli aplogruppi	- 76 -
3.2.3 Network filogenetici.....	- 81 -
3.2.4 Distanze ed alberi genetici	- 84 -
4. Confronto fra due diversi metodi di campionamento	91
4.1 Analisi degli Aplogruppi.....	92
4.2 Network filogenetici	94
4.3 Alberi filogenetici	95
5. Discussioni	99
Bibliografia	112
<i>Ringraziamenti</i>	126

1. Introduzione

1.1 Le comunità isolate

Una comunità si definisce isolata quando deriva da un piccolo numero di individui fondatori e persiste per più generazioni in una condizione di isolamento geografico o culturale durante il quale non riceve alcun contributo genetico dall'esterno.

L'isolamento può essere dovuto a barriere di tipo geografico quali la presenza di montagne e corsi d'acqua, all'assenza di vie di comunicazione stradale adeguate, al verificarsi di calamità quali ad esempio terremoti, guerre, carestie, a fattori culturali quali la religione e la lingua.

Solitamente le popolazioni isolate sono caratterizzate da un limitato numero di fondatori, da un alto tasso di endogamia e consanguineità, da un'espansione della popolazione bassa o nulla e da un'immigrazione trascurabile.

L'isolamento porta alla formazione di un pool genico differente rispetto a quello della popolazione d'origine e delle popolazioni circostanti. Tale differenziamento del pool genico avviene ad opera di vari meccanismi che solitamente agiscono in maniera combinata. I principali sono la deriva genetica, l'adattamento all'ambiente e l'incrocio non casuale.

La **deriva genetica**, ovvero la variazione casuale delle frequenze alleliche, può avere importanti implicazioni riguardo la sopravvivenza e l'evoluzione di popolazioni piccole. Infatti, più piccola è la popolazione maggiore è l'effetto della deriva genica, mentre nelle popolazioni grandi o in rapida espansione è parzialmente compensata dall'elevato numero di prole per generazione (Cavalli-Sforza et al. 1994). La deriva genetica conduce alla perdita di variabilità all'interno delle popolazioni ed a fluttuazioni casuali col tempo delle frequenze alleliche di una popolazione.

Una modalità mediante la quale si può verificare questo fenomeno è il cosiddetto “**collo di bottiglia**”. Questo evento si ha quando una popolazione si trova in condizioni sfavorevoli (come, ad esempio, cambiamenti climatici, catastrofi naturali, malattie infettive) che riducono notevolmente il numero di individui. In seguito a tale “collo di bottiglia” le frequenze alleliche possono variare drasticamente, soprattutto in piccole popolazioni.

Un altro modo in cui si può manifestare deriva genetica è per il **principio del fondatore**. Questo si verifica quando una specie si espande in un’altra regione e un piccolo numero di individui fonda una nuova popolazione. È improbabile che tali fondatori possiedano tutti gli alleli presenti nella popolazione parentale e anche nell’ipotesi in cui questo avvenga è probabile che la frequenza di ciascun allele sia diversa da quella della popolazione originaria. Più piccolo è il numero di fondatori, maggiore sarà la probabilità che essi non siano rappresentativi della popolazione più ampia da cui provengono. Quindi, sebbene la popolazione possa in seguito crescere di numero, il pool genico sarà sempre quello derivato dai geni presenti negli individui fondatori. Ciò - ovviamente - a patto che non vi siano migrazioni, dal momento che il flusso genico tende a ridurre le divergenze genetiche esistenti tra le popolazioni.

Su tale pool genico agirà poi la selezione naturale, a cui gli individui risponderanno con una reazione adattativa che dipenderà dall’ambiente. Infatti, due popolazioni isolate aventi lo stesso pool genico possono sviluppare, col passare delle generazioni, pool genici diversi se esposte a condizioni ambientali differenti e quindi a pressioni selettive diverse. Può anche accadere però che, pur subendo la stessa pressione selettiva, esse sviluppino meccanismi adattativi diversi a causa della deriva genetica.

Non va infine dimenticato l’effetto dell’accumulo pressoché costante delle **mutazioni**, i cui risultati su ciascuna popolazione possono essere completamente differenti.

1.2 Gli isolati in Sardegna

Le popolazioni isolate sono campioni ideali per lo studio di fenomeni evolutivi quali il popolamento e le migrazioni e per la ricostruzione di genealogie (Colonna et al., 2007). Studiare gli effetti dell'isolamento sulla struttura genetica di una popolazione consente di valutare il grado di differenziazione rispetto alla popolazione di appartenenza, il grado di coinvolgimento nei flussi migratori che hanno investito l'area di interesse in tempi preistorici e storici, l'effettiva età del pool genico degli isolati, la possibilità di riconoscere influssi di altre popolazioni, l'evoluzione autoctona con geni e aplotipi propri. Sotto questi aspetti la Sardegna è un'area di studio eccellente poiché sono ben note la sua archeologia, storia, cultura, nonché le sue peculiarità linguistiche e genetiche.

La popolazione Sarda è stata ampiamente studiata per le sue particolari caratteristiche genetiche che la differenziano dalle altre popolazioni italiane e mediterranee. Tale eterogeneità genetica si riscontra anche all'interno della stessa popolazione sarda ed è dovuta alla frammentazione geografica e/o culturale in numerosi isolati (Gatti et al., 1990; Walter et al., 1991; Vona et al., 1992; Morelli et al., 1993; Moral et al., 1994; Zavattari et al., 2000; Angius et al., 2001; Calò et al., 2008; Pistis et al., 2009). Questa differenziazione può essere il risultato di vari fattori che hanno agito per lungo tempo sulla popolazione sarda, quali ad esempio le barriere linguistiche e geografiche, l'estensione del territorio sovradimensionato rispetto al numero degli abitanti, l'alto tasso di endogamia e consanguineità e la presenza della malaria che ha esercitato una forte pressione selettiva in alcune zone dell'isola modificando direttamente la frequenza di alcuni geni (Cappello et al., 1996; Gatti, 1990; Vona et al., 1996).

La maggior parte degli autori ricollega l'eterogeneità genetica tra le popolazioni isolate all'effetto della deriva genetica (Fraumene et al., 2006); infatti, le barriere naturali e la complessa orografia hanno spesso mantenuto alcune aree estremamente isolate (Workman et al., 1975). Il suddetto isolamento interno è evidenziato anche dalla parcellizzazione dialettale della lingua, chiaro indice della frammentazione della popolazione sarda in sottopopolazioni. Un recente lavoro sul DNA antico che ha riguardato il DNA mitocondriale ha evidenziato un basso livello di eterogeneità genetica nel periodo nuragico (tra 3430 e 2700), il che suggerisce una recente differenziazione del pool genico sardo (Caramelli et al., 2007).

Alcuni piccoli paesi della Sardegna sono particolarmente indicati per il mappaggio di geni coinvolti nelle malattie complesse (Zavattari et al, 2000; Angius et al, 2001), dal momento che le comunità isolate di ridotte dimensioni eliminano tutte quelle problematiche che si devono affrontare quando si utilizzano popolazioni aperte (Heutink et al., 2002; Bourgain et al., 2001) per questa tipologia di studi. Per questo motivo, la Sardegna risponde perfettamente alle condizioni necessarie per questo tipo di indagini, poiché la sua popolazione è distribuita in pochi grandi aggregati urbani di recente formazione e in molti piccoli villaggi isolati.

1.3 Indicatori biodemografici dell'isolamento

Il grado di isolamento di una popolazione può essere studiato attraverso analisi di tipo biodemografico che, esaminando i comportamenti matrimoniali, traducono un fenomeno sociale nel fenomeno biologico della riproduzione. Infatti, tramite la scelta del coniuge, il matrimonio rappresenta l'evento in cui le componenti biologiche e quelle socio-culturali interagiscono tra loro, influenzando l'evoluzione del patrimonio genetico della popolazione. Lo studio dell'evoluzione della struttura matrimoniale si avvale di diversi indicatori bio-

demografici: i tassi di endogamia ed esogamia, le distanze matrimoniali, i coefficienti di consanguineità, e altri coefficienti il cui calcolo si basa sulla distribuzione dei cognomi (coefficienti isonimici).

Questi studi bio-demografici possono essere effettuati utilizzando i dati riportati nei *Quinque Libri*: registri dei battesimi, delle morti, delle cresime, dei matrimoni e dello stato delle anime, conservati presso gli archivi parrocchiali o diocesani, e nei registri anagrafici comunali.

Nelle popolazioni umane la scelta del coniuge nell'ambito di una determinata area viene condizionata da fattori di tipo etnico, sociale, religioso, psicologico, socio-economico, culturale, geografico. In popolazioni isolate si ha un aumento delle unioni matrimoniali endogamiche, cioè fra individui appartenenti alla stessa cerchia matrimoniale (comune, etnia, casta etc.). I matrimoni esogamici, invece, sono quelli nei quali almeno uno dei coniugi non appartiene alla cerchia matrimoniale in questione. I tassi di endogamia ed esogamia non danno solo informazioni sul grado di isolamento riproduttivo, ma determinano anche il livello di affinità genetica all'interno e tra gruppi, dato che più elevata è l'endogamia, maggiore è l'omogeneità interna del gruppo e minore è l'affinità con i gruppi circostanti (Pettener, 1995).

L'estremizzazione dell'endogamia porta al matrimonio fra consanguinei, in cui i due individui hanno almeno un progenitore relativamente prossimo in comune, la cui progenie viene detta inincrociata (inbred). Il livello di inincrocio (inbreeding) è correlato direttamente con il grado di parentela. Molti dei metodi adottati per misurare il livello di inincrocio furono inventati negli anni 20 dal genetista americano Sewall Wright. L'effetto dell'inincrocio è un continuo decremento della frazione degli eterozigoti ed il corrispondente incremento di quella degli omozigoti, rispetto alle proporzioni attese secondo l'equilibrio di Hardy-Weinberg; ma il tasso di incremento dell'omozigosi dipende dal livello di inincrocio. (Mange et Mange, 1993).
Sebbene la consanguineità non modifichi di per sé le frequenze alleliche, incrementa la

proporzione di omozigoti alla generazione successiva, esponendo quindi alla selezione fenotipi recessivi (Gelehrter et al., 2002).

Un individuo inincrociato può essere omozigote o perchè possiede alleli aventi una storia indipendente o perchè possiede alleli aventi una storia comune, ossia che derivano esattamente dallo stesso allele che era presente nel progenitore comune. La probabilità che i due alleli presenti in una stessa persona siano identici per discendenza rappresenta il *coefficiente di inincrocio* (F) della persona stessa che si può così calcolare:

$$F = \sum_i^k (1/2)^{n_i + p_i + 1}$$

in cui n_i e p_i sono il numero di passi generazionali che uniscono i due genitori dell'individuo in esame passando per l'antenato comune e k sono le linee indipendenti di parentele possibili.

Dunque, ogni generazione in più dimezza la probabilità che l'allele ancestrale sia presente in entrambi i coniugi e che da questi sia trasmesso all'individuo in esame.

Se a sua volta l'antenato comune è consanguineo con un coefficiente di consanguineità F_z , la formula diventa:

$$F = \sum_i^k (1/2)^{n_i + p_i + 1} (1 + F_z)$$

F	PARENTELA
1/8	Zio/a-nipote
1/16	Cugini primi
1/32	Cugini primi con una generazione di distanza (Cugini primi e ½)
1/64	Cugini secondi
1/128	Cugini secondi con una generazione di distanza (Cugini secondi e ½)
1/256	Cugini terzi

Tabella 1. Alcuni esempi di coefficiente di inincrocio fra parenti. (Mange. et Mange, 1993.).

Nell'uomo ed in tutte le specie a sessi separati, la forma di inincrocio più stretta risulta dall'accoppiamento fra genitore e figlio e fra fratello e sorella. Tali accoppiamenti sono generalmente proibiti dalla legge e condannati quasi universalmente da tutte le religioni. I

matrimoni fra zio (o zia) e nipote sono molto rari nei gruppi umani eccetto quelli ove i costumi sociali permettono questa forma di promiscuità. Il matrimonio fra cugini di primo grado rappresenta l'accoppiamento tra consanguinei più frequente nell'uomo (Mange e Mange, 1993.).

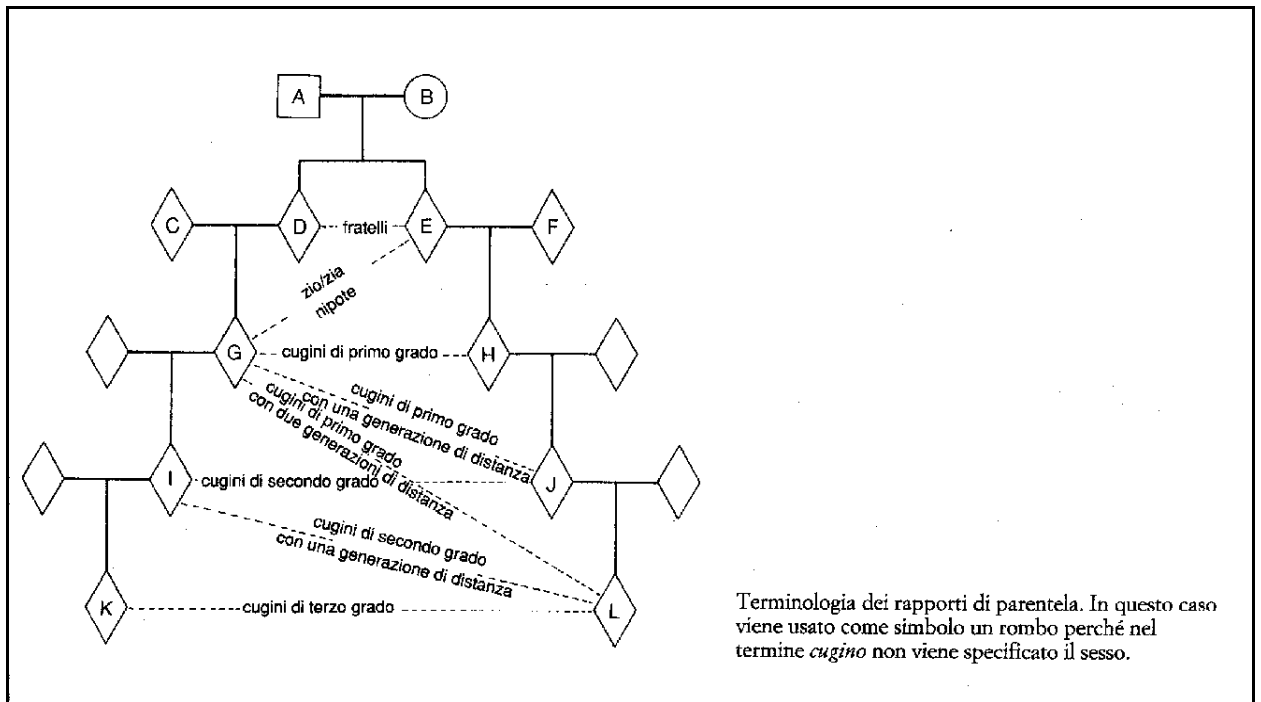


Figura 1. Albero genealogico che spiega la terminologia e il grado dei vari cugini (Mange e Mange , 1993)

In una popolazione, invece, la consanguineità può essere stimata dalla incidenza percentuale di matrimoni consanguinei rispetto al numero totale di matrimoni verificatesi in un dato arco di tempo, calcolata come proposto da Bernstein (1930) con la formula:

$$\alpha = \sum p_i F_i$$

in cui p_i indica la frequenza relativa dei matrimoni tra consanguinei di un determinato grado , mentre F_i è il coefficiente di consanguineità relativo.

La stima di α indica la probabilità che due alleli omologhi di un individuo scelto a caso nella popolazione siano identici.

Il livello medio di consanguineità è considerato un buon indicatore dell'isolamento riproduttivo; infatti, è risultato correlato in modo statisticamente significativo con il livello di endogamia e con l'altitudine (Floris et al., 2000).

Generalmente i coefficienti di consanguineità delle popolazioni umane presentano valori inferiori a 10^{-3} . Questo perché i matrimoni tra parenti di grado più stretto tendono ad essere evitati nelle società umane.

In Sardegna il livello medio di consanguineità (α), dal 1765 al 1965, è risultato essere pari a $1,54 \times 10^{-3}$ (Moroni et al., 1972).

Un altro tipo di analisi sui matrimoni fra consanguinei si basa **sull'isonimia maritale**, ossia sul ricorrere di cognomi uguali nei due individui che contraggono il matrimonio (Crow e Mange, 1965; Lasker, 1985 cit da Mange e Mange, 1993). Con il metodo dell'isonimia si può stimare il grado di omogeneità in una popolazione, o tra popolazioni, in quanto la condivisione dello stesso cognome potrebbe indicare un antenato comune.

I cognomi, che hanno un comportamento paragonabile a quello di alleli neutri (Zei et al., 1983), consentono di pervenire ad una stima dei valori di "inbreeding" in una data comunità, di valutare il grado di suddivisione in sottopopolazioni, di analizzare i rapporti di affinità genetica tra gruppi umani, di stimare la consistenza delle migrazioni, di calcolare l'isolamento da distanza.

Nel corso degli ultimi secoli si è assistito ad una tendenza alla diminuzione dell'endogamia ed al conseguente aumento dell'esogamia. Questo fenomeno particolarmente accentuato dagli inizi del Novecento in poi, è da mettere in relazione con l'aumento della mobilità individuale e con lo spopolamento delle aree montane e delle zone geograficamente più isolate in seguito ai cambiamenti socioculturali legati all'industrializzazione, ed è noto come **rottura degli isolati**.

1.4 Descrizione delle popolazioni analizzate

1.4.1 Benetutti

Benetutti è un piccolo centro in provincia di Sassari facente parte della Comunità Montana del Goceano. Il *Goceano* è una regione ubicata nella parte più interna dell'Isola orlata dalla Catena omonima e circondata a sud-est dal Marghine e, procedendo in senso orario, dal Logudoro, le ultime propaggini del Monte Acuto, gli altopiani di Alà, Buddusò e Bitti e dal Nuorese. Il suo territorio, esteso per 481 kmq, comprende nove comuni, di cui sette aggrappati alla montagna nel versante sud occidentale del Marghine (*Illorai, Burgos, Esporlatu, Bottidda, Bono, Anela, Bultei*), e due poco più a nord nella vallata attraversata dal fiume *Tirso* (*Benetutti e Nule*).

Benetutti, quindi, sorge alle pendici dell'altopiano del Goceano a 406 metri s.l.m. e il suo territorio è attraversato dal fiume *Tirso* e dai suoi affluenti *Rio Mannu* e *Rio Minore*. Conta una popolazione di 2181 abitanti (secondo il censimento ISTAT del 2001) e ha una superficie di 94,5 Kmq per una densità abitativa di 23,1 abitanti per Kmq (Angioni et al, 1997). Con gli altri centri del Goceano condivide la crisi delle campagne e il forte calo demografico, dovuto alla fuga degli abitanti verso luoghi più sicuri ed economicamente più affidabili. Dal censimento del 1971 a quello del 2001 Benetutti ha perso 262 abitanti (il Goceano nello stesso arco di tempo ha perso 2637 abitanti) (Angioni et al, 1997). L'economia tarda a decollare per la mancata realizzazione da parte delle amministrazioni competenti di vie di collegamento adeguate, nonché di altri interventi infrastrutturali che possano favorire l'agricoltura come la costruzione dell'invaso sul Riu Mannu, atteso da quasi cinquant'anni. Benetutti vive principalmente dell'allevamento del bestiame e della trasformazione del latte, ma è conosciuto anche per la sua rinomata produzione enologica, i suoi due centri termali e

non mancano iniziative industriali significative nel campo tessile e in quello sughericolo. Negli ultimi anni ha avuto un notevole impulso anche l'attività edilizia come confermano le numerose case costruite alla periferia del paese.

È un paese antichissimo che conserva nel suo territorio testimonianze del periodo prenuragico e di quello nuragico, che confermano come l'area fosse abitata fin dal tremila avanti Cristo. Si trovano, infatti, varie domus de janas, dolmen, tombe dei giganti e numerosi nuraghi disseminati nel vasto territorio che si estende per quasi 9500 ettari. Anche i Romani hanno lasciato una traccia della loro permanenza sul territorio, come dimostra per esempio, la vasca termale che si trova all'interno delle terme di San Saturnino. I Romani avevano scoperto per primi le qualità terapeutiche di quell'acqua sulfurea e ferruginosa che sgorga alla temperatura di 34 gradi, e vicino alla sorgente avevano costruito la città di Lesa, più volte ricordata da Tolomeo. Secondo la leggenda il nome del paese deriva proprio dalla presenza di queste sorgenti termali che si trovano sparse qua e là per il suo territorio. Dai millenni della preistoria fino al Medioevo, se si esclude la parentesi della dominazione romana, un lungo periodo di silenzio avvolge le vicende della zona di Benetutti. Le prime notizie del centro abitato risalgono infatti all'età giudiciale.

Nel Medioevo il paese ospitò una comunità di monaci Camaldolesi. Questi ultimi edificarono, nel XIII secolo, la chiesetta di San Saturnino su un'area nuragica. Erano anni di grande fervore religioso, tanto che il paese vantava un singolare primato fra i comuni della zona per il numero delle chiese possedute: ve ne erano ben dieci, sette si trovavano all'interno dell'abitato e tre nella vicina campagna. In una di queste, la parrocchiale di Sant'Elena sono custoditi gli scomparti del retablo - databile attorno al 1549 - dipinto da Giovanni del Giglio, meglio noto come il Maestro di Ozieri (Oppes, 1990).

Da un recente studio che ha analizzato la struttura matrimoniale di 20 paesi sardi ubicati nel territorio a diverse altitudini è emerso un alto livello di isolamento di Benetutti (Sanna et al,

2004). In tale lavoro son stati analizzati i matrimoni celebrati dal 1800 al 1974 ed è stato riscontrato un alto tasso di endogamia (84,4%) e di consanguineità con un α di Bernstein pari a $3,85 \times 10^{-3}$. In generale si è visto che i matrimoni consanguinei più frequenti in Sardegna sono quelli tra cugini di secondo grado ($F = 1 / 64$), mentre un contributo importante per il livello medio di consanguineità deriva anche da matrimoni fra cugini di primo grado ($F = 1 / 16$) e cugini di terzo grado ($F = 1 / 256$).

1.4.2 Carloforte

Carloforte è una cittadina dell'Isola di San Pietro situata nell'arcipelago del Sulcis presso le coste sud-occidentali della Sardegna. La popolazione di Carloforte conta 6444 abitanti secondo il censimento ISTAT del 2001.

L'isola era conosciuta dai Fenici col nome di *Enosim*, dai navigatori Greci come *Hieracòn Nésos* e dai Romani come *Accipitrum Insula* (o isola degli sparvieri) Il suo nome attuale è dalla tradizione attribuito al passaggio del santo Apostolo, quando, per una tempesta, durante un suo viaggio dall'Africa verso Roma con tappa a Cagliari, fu costretto a fermarsi in questa isola.

A testimonianza di una antichissima frequentazione in epoca nuragica, sono ancora visibili i resti di qualche nuraghe. In seguito, vista la sua posizione strategica rispetto alle antiche rotte di navigazione che costeggiavano la Sardegna, attirò sulle sue coste fenici e romani.

La storia dell'attuale comunità di abitanti di Carloforte ha origine da un piccolissimo gruppo di abitanti Genovesi che nel 1542 andarono ad abitare una piccola isola disabitata della Tunisia (Tabarka), situata tra Bona e Biserta. L'Isola di Tabarka fu ceduta ai Lomellini, signori di Pegli, dalla Spagna, come punto d'appoggio per la ricca industria del corallo.

Questa piccola colonia, costituita quasi interamente da Pegliesi, crebbe floridamente, ma visse sempre isolata rispetto alle popolazioni del Nord Africa. Fece accumulare ricchezze ai Lomellini per circa un secolo e mezzo ma, quando ai primi anni del 1700 le risorse di corallo si impoverirono, la popolazione, che ormai contava 2000 abitanti, conobbe periodi durissimi. A rendere la situazione ancora più difficile si aggiunsero le continue molestie dei corsari e i taglieggiamenti del Bey di Tunisi e di Algeri, ai quali i Lomellini già pagavano un tributo annuo di 9000 scudi. Quando i tabarkini vennero a conoscenza del fatto che il re di Sardegna Carlo Emanuele III di Savoia concedeva in feudo territori disabitati del suo Regno, si adoperarono ad individuare un'area adatta alle loro attitudini lavorative ed intavolarono le trattative per la colonizzazione dell'Isola di San Pietro. Fu così che il 22 Febbraio 1738 un primo gruppo di 86 persone piene di entusiasmo e speranze si recarono verso l'Isola di San Pietro per iniziare i lavori di costruzione delle prime case di quella cittadina che chiameranno **Carloforte** in onore del re sabauda. Due mesi dopo vennero raggiunti da altri 302 tabarchini. La giovane cittadina attirò anche alcune famiglie liguri. Secondo il primo censimento il gruppo di fondatori era costituito da 100 famiglie di Tabarka (381 persone complessive), 26 famiglie Liguri (83 persone complessive) e 2 curati.

Questa comunità continuò a praticare la raccolta del corallo, ma intraprese anche altre attività, quali la pesca del tonno, la produzione di sale, l'agricoltura. Inoltre, non mancarono quanti si dedicarono al commercio, spingendosi per tutto il Mediterraneo.

Il più drammatico evento nella storia della comunità carlofortina si verificò la notte del 2 settembre 1798, quando, alcune centinaia di corsari tunisini, penetrarono nella città e la devastarono, compiendo angherie, soprusi e massacri nei confronti degli abitanti. Quando ripartirono condussero con sé, come schiavi, ben 830 abitanti. Dopo cinque anni, il Re di Sardegna, Vittorio Emanuele I, riuscì a riscattare dalla loro prigionia i carlofortini che finalmente poterono tornare nella loro isola. Fortunatamente, dopo quell'episodio, non se ne

ripeterono altri e la comunità poté godere di un periodo di tranquillità (Vallebona, 1974 ; Ferraro, 1989; Luxoro, 1977).

La comunità di Carloforte ha suscitato notevole interesse da parte della comunità scientifica che l'ha studiata e continua ancora oggi ad analizzarla sotto molteplici punti di vista: linguistico, antropometrico, demografico e genetico, che evidenziano tutti il forte grado di isolamento e l'eterogeneità rispetto alla popolazione sarda. L'isolamento - che sembra avere avuto un effetto conservativo sui tratti genetici probabilmente già presenti nel nucleo fondatore - ha molteplici componenti tra cui quella geografica rafforzata da quella linguistica. Infatti, la comunità carlofortina parla ancora oggi il dialetto dei suoi antenati pegliesi. La barriera linguistica, può aver ostacolato il libero scambio di geni con le popolazioni sarde limitrofe, almeno per un buon periodo di tempo. Secondo Cavalli-Sforza (1994) le lingue si evolvono in modo simile ai caratteri genetici e quindi è possibile ricostruire una filogenesi linguistica che può essere utilmente impiegata per interpretare la storia, anche genetica, delle popolazioni moderne.

Uno studio sulla struttura matrimoniale dal 1738 al 1993 (Calò et al, 1994) ha evidenziato un elevato tasso di endogamia, la ristretta area matrimoniale e la marcata consanguineità. Secondo tale lavoro a Carloforte, dal 1738 al 1993, sono stati celebrati 8104 matrimoni di cui il 75,42% endogamici e solo il 24,58% esogamici. I matrimoni fra coniugi entrambi non nativi di Carloforte sono il 5,02% del totale. I matrimoni fra carlofortini e sardi sono però andati aumentando nel tempo (dal 7,14% del periodo 1738-49 al 79,30% del 1950-93), motivo per cui si è portati a pensare che in un futuro vi potrà essere una omogeneità genetica fra le due popolazioni. La distanza matrimoniale media totale è risultata di 42,36 Km. I matrimoni fra consanguinei sono stati 563 su 8104 (6,95%), ma a partire dagli anni '60 tale pratica è andata in disuso.

Il coefficiente medio di consanguineità è risultato pari a $1,63 \times 10^{-3}$, un valore leggermente superiore a quello riportato da Moroni per la Sardegna ($1,54 \times 10^{-3}$). (Moroni et al., 1972).

Anche dallo studio demografico inerente l'isonimia tra il 1876 e il 1976 è emerso un bassissimo grado di relazione tra carlofortini e sardi (Floris et al., 1980). Un lavoro del 2006 (Sanna et al., 2006) mostra come l'isonimia osservata aumenta fino al 1900-1949 della popolazione verso matrimoni con individui esterni o comunque non appartenenti allo stesso clan familiare, con la conseguente introduzione di nuovi cognomi per poi decrescere nel periodo successivo, mostrando una tendenza generale dell'apertura all'interno della comunità. Studi antropometrici condotti negli anni '70-'80, hanno riscontrato una maggiore affinità morfologica dei carlofortini con i liguri piuttosto che con i sardi (Cosseddu et al., 1979; Lucia e Vona, 1982).

Altri studi hanno, invece, analizzato la struttura genetica di tale comunità attraverso l'utilizzo di differenti marcatori.

Dall'analisi di 12 marcatori genetici classici (ABO, Rh, MN, PGD, AK, PGM1, DIA, ACP1, ESD, HP, GC, C3), i carlofortini hanno mostrato frequenze geniche fuori dal campo di variabilità di quelle sarde, mentre si caratterizzavano rispetto ai liguri per le frequenze più elevate di ABO*0 e dall'aplotipo CDe. (Vona et al., 1994). Analoghe conclusioni si trovano in un lavoro del 1996 (Vona et al., 1996) dove è riportato l'albero genetico e le distanze genetiche medie che collocano Carloforte in una posizione intermedia tra sardi e continentali, ma con un maggiore orientamento verso questi ultimi. Si è quindi trovato un basso grado di omogeneizzazione fra tra sardi e carlofortini e una maggiore affinità di questi ultimi con gli italiani continentali.

Una ricerca medica del 2001 (Heath et al., 2001) ha riscontrato un'elevata incidenza fra i carlofortini di una patologia genetica rara (la miopia elevata). Quest'incidenza potrebbe essere dovuta all'effetto del fondatore che ha subito la popolazione di Carloforte. Probabilmente fra

gli individui fondatori era presente il gene mutante che per effetto dell'isolamento, dell'elevato tasso di endogamia e consanguineità ha potuto diffondersi.

1.5 Utilizzo dei Marcatori Uniparentali

Il DNA uniparentale (cromosoma Y e DNA mitocondriale) ha la particolarità di non dar luogo a ricombinazione e quindi, di passare di generazione in generazione senza subire cambiamenti, ad eccezione delle mutazioni che si accumulano nel tempo. L'assenza di ricombinazione rende perciò più facile l'interpretazione filogenetica. È possibile quindi tracciare le origini geografiche di una certa popolazione e risalire alle componenti che la hanno generata. Per tale motivo vengono largamente impiegati in antropologia molecolare per studiare l'evoluzione dell'uomo, ricostruire i flussi migratori e il mescolamento delle popolazioni. Un'altra applicazione utilissima è quella di ricercare o scoprire legami di parentela (Caramelli, 2009).

1.6 Cromosoma Y

Il cromosoma Y, ereditato come entità aploide per via paterna, è un piccolo cromosoma acrocentrico caratterizzato da una bassissima densità di geni e da una bassa frequenza di mutazioni. È una molecola lineare costituita da circa 50 Mb. Da un punto di vista strutturale (Fig. 2) il cromosoma Y è costituito da due regioni localizzate alle estremità, chiamate pseudoautosomiche (PAR 1 e PAR 2) e da una zona non ricombinante (NRPY) localizzata fra queste. Le regioni PAR durante la meiosi nei maschi si appaiano e scambiano materiale genico con zone di omologia del cromosoma X. Quindi, i geni presenti all'interno di queste

sequenze presentano un'ereditarietà di tipo autosomico. Oltre a ciò le PAR (in particolare la PAR1) sono molto importanti, in quanto permettono una corretta segregazione cromosomica durante la meiosi. La regione PAR 1 è localizzata all'estremità del braccio corto ed ha una lunghezza di circa 2,5 Mb, mentre la regione PAR 2 si trova all'estremità del braccio lungo e occupa una porzione inferiore a 1 Mb. Fra queste zone pseudoautosomiche, come già detto, si trova una regione non omologa al cromosoma X che si chiama, per tale motivo, regione non ricombinante (NRY non-ricombinig region of the human Y chromosome). Alcuni ricercatori hanno attribuito a questa regione il nome di male specific region (MSY) per mettere in rilievo l'appartenenza di questa regione a individui di sesso maschile (Skaletsky et al., 2003). La regione NRPY comprende circa il 95% dell'intero cromosoma ed è costituita da una regione eucromatica e da una eterocromatica. Un gruppo di ricercatori è riuscito a sequenziare una parte eucromatica del cromosoma Y di circa 23 Mb, di cui 8 Mb situate sul braccio piccolo Yp e 14,5 Mb localizzate sul braccio lungo Yq (Skaletsky et al., 2003). La restante parte del cromosoma, corrispondente a una regione di eterocromatina di circa 30Mb localizzata sul braccio lungo, è composta da sequenze altamente ripetute che con l'attuale tecnologia non è possibile sequenziare. Ad oggi non sono stati ancora localizzati geni all'interno della regione eterocromatica. Le unità trascrizionali che sono state messe in evidenza sono contenute esclusivamente all'interno delle sequenze eucromatiche (sono presenti solo poche dozzine di geni).

Dal momento che a livello della NRY non avviene ricombinazione, si avrà all'interno di questa regione un accumulo di eventi mutazionali avvenuti nel corso delle generazioni lungo la linea esclusivamente maschile. Pertanto lo studio dei polimorfismi localizzati in questa regione presenta svariate applicazioni fra cui lo studio dell'evoluzione umana.

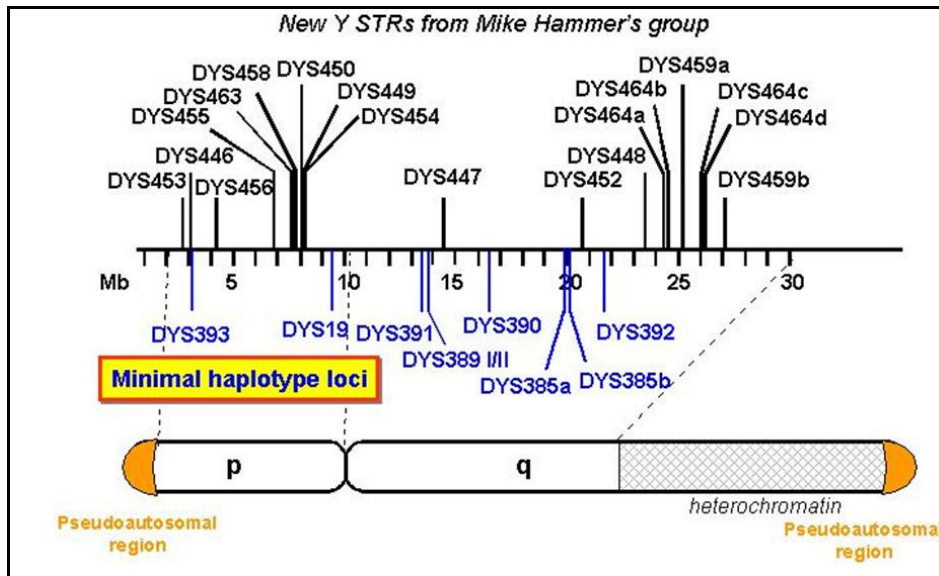


Figura 2. Struttura del cromosoma Y e posizione di alcuni STRs.
<http://www.cstl.nist.gov/strbase/ystrpos1.htm>

1.6.1 Marcatori del cromosoma Y

Le due categorie di marcatori molecolari utilizzate per lo studio della variabilità del cromosoma Y sono i loci bi-allelici e i loci multi-allelici. Fanno parte dei loci bi-allelici i polimorfismi a singolo nucleotide (SNPs) e l'inserzione o delezione degli elementi Alu che portano entrambi alla comparsa di due possibili forme alleliche. Questi marcatori hanno un basso tasso di mutazione e sono considerati quasi eventi unici (circa $<10^{-9}$ per generazione). I loci multi-allelici comprendono 2 minisatelliti o VNTRs (Variable Number of Tandem Repeat) e più di 200 microsatelliti o STRs (Short Tandem Repeat). Questi ultimi due polimorfismi sono dovuti a variazioni in lunghezza di una sequenza del DNA, quindi, ogni satellite è caratterizzato da una determinata sequenza nucleotidica di una certa lunghezza ripetuta diverse volte. Un allele si differenzia da un altro per il numero di ripetizioni. I VNTRs consistono in ripetizioni di 15-20 bp, con una lunghezza totale che varia da 300 bp ad

1kb. Gli STRs hanno un segmento ripetuto di non più di 6 bp per una lunghezza complessiva di 100-200 bp.

I loci multi-allelici hanno un alto potere discriminatorio di aplotipi diversi dovuto all'alto tasso di mutazione (circa 10^{-2} per i VNTRs e circa 10^{-3} per gli STRs). I polimorfismi più comunemente usati negli studi evolutivi sono i marcatori bi-allelici e i microsatelliti che risultano essere particolarmente funzionali nell'analisi di diversi periodi evolutivi in quanto presentano differenti tassi di mutazione. I marcatori biallelici per il loro basso tasso di mutazione permettono di suddividere un gruppo di cromosomi in aplogruppi, mentre, i marcatori multiallelici, come per esempio i microsatelliti, possono essere usati per definire gli aplotipi dentro gli aplogruppi, indicando una maggiore diversità (Scozzari et al., 2001).

1.6.2 Aplogruppi del cromosoma Y

Data la mancanza di ricombinazione e la frequenza estremamente bassa di retromutazione, è stato possibile dedurre un albero filogenetico estremamente preciso e con minore ambiguità di interpretazione rispetto a quello ottenuto tramite l'mtDNA. La regione non ricombinante del cromosoma Y (NRY) ha oggi la maggior risoluzione aplotipica ad ogni locus di qualsiasi altro sistema del genoma, e sono stati riconosciuti e raccolti in una dettagliata filogenesi 153 aplogruppi, definiti da marcatori biallelici (figura 3).

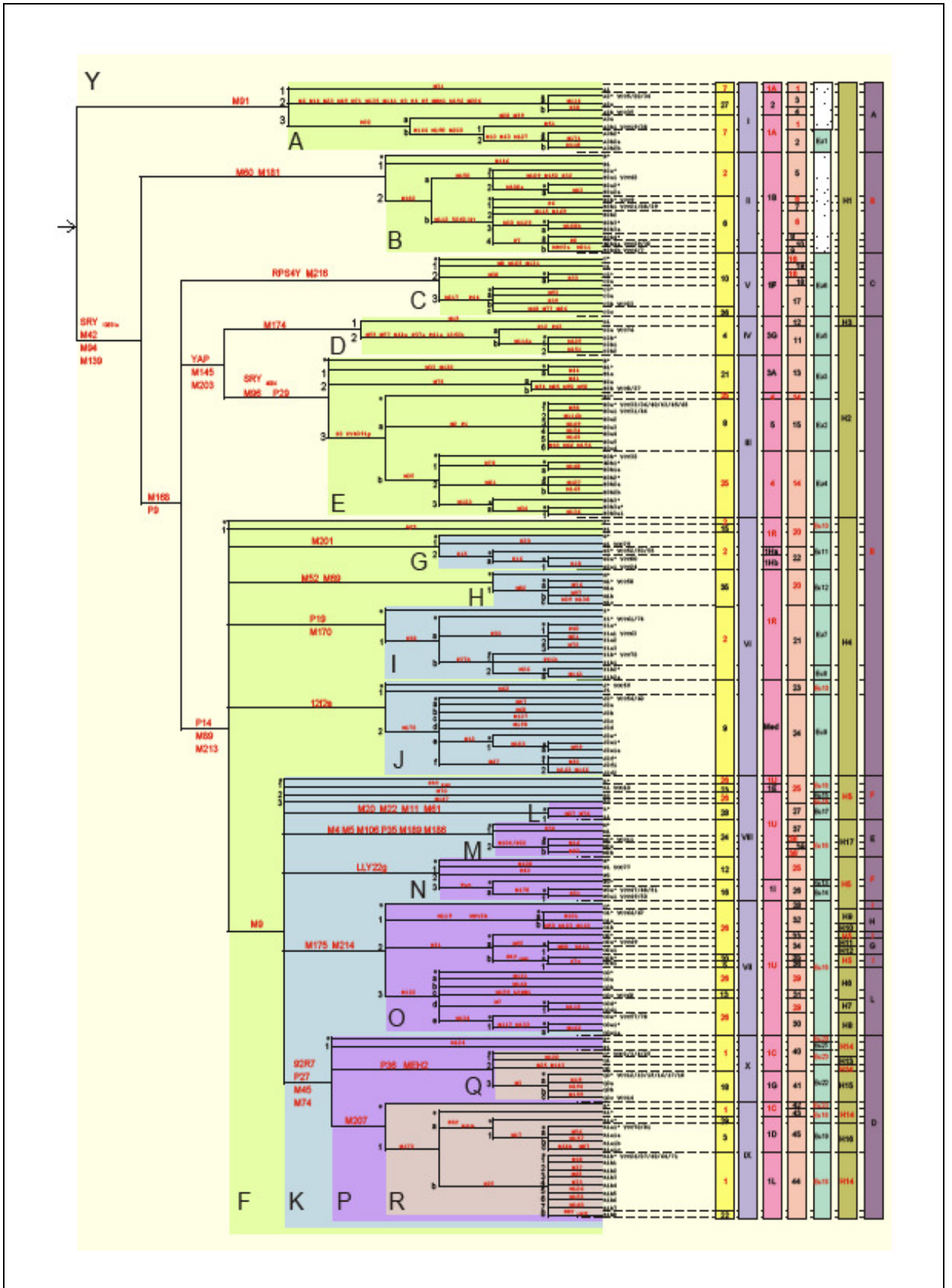


Figura 3. Albero filogenético del cromosoma Y (YCC, 2002)

Con il termine aplotipo si intende la combinazione di diversi stati allelici di un set di marcatori polimorfici che si trovano fisicamente associati sulla stessa molecola di DNA come per esempio un cromosoma od una regione cromosomica. L'aplogruppo rappresenta un insieme di aplotipi di cui si ipotizza un'origine comune, grazie alla condivisione di mutazioni caratteristiche (generalmente ad evoluzione lenta).

Per il cromosoma Y, l'aplotipo è costituito dalla somma della variabilità di polimorfismi microsatelliti (STRs) (Fig. 4). L'aplogruppo si definisce sulla base della condivisione di mutazioni specifiche per marcatori biallelici (SNPs) e non per i microsatelliti, poiché hanno un tasso di mutazione troppo alto (fig. 5).

Aplotipo	DYS19	DYS389-1	DYS389-2	DYS390	DYS391
1	15	8	25	21	10
2	15	10	28	21	10
3	17	10	27	21	10
4	16	10	27	21	10

Figura 4. Esempio di aplotipo per il cromosoma Y

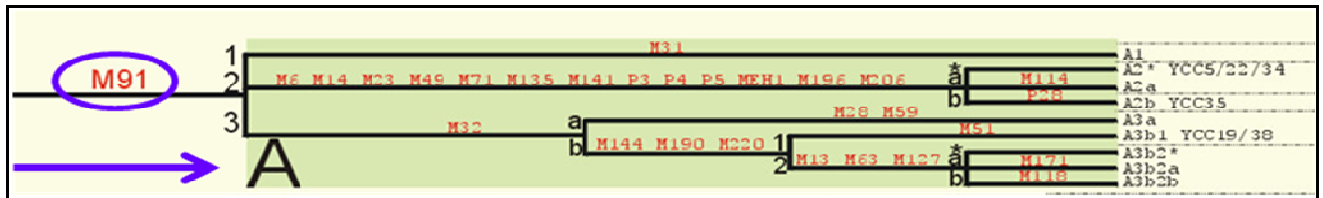


Figura 5. Esempio di aplogruppo per il cromosoma Y

Negli ultimi decenni è andato sempre più crescendo il numero dei marcatori biallelici scoperti che ha portato al nascere di almeno sette diverse nomenclature. Questo ha recato notevoli disagi nel confronto dei risultati delle varie pubblicazioni, tanto che, nel 2002, l'Y Chromosome Consortium (YCC) ha pubblicato un albero filogenetico del cromosoma Y costruito tramite l'analisi di marcatori (SNPs), genotipizzati in un set di campioni rappresentativo di tutte le popolazioni (Fig. 3). È stato messo a punto un semplice insieme di

norme per etichettare in maniera non ambigua i diversi clades localizzati all'interno di questo albero. Le Lettere maiuscole da A a R, sono state utilizzate per identificare i 18 principali clades (aplogruppi) e ulteriori divisioni sono state apportate usando numeri e lettere al pedice. Questo sistema di nomenclatura gerarchico, che è stato ampiamente accettato dalla comunità scientifica, unifica tutte le nomenclature utilizzate in precedenza. In seguito, in un lavoro di Jobling and Tyler-Smith del 2003 è stata proposta una versione modificata di tale albero (Fig.6) e, nel 2008 Karafet e collaboratori hanno pubblicato un ulteriore aggiornamento dell'albero e introdotto una modifica della nomenclatura basata sulle regole stabilite dall'YCC. Nel lavoro di Jobling e Tyler-Smith (2003) sono, inoltre, presenti informazioni sulla distribuzione geografica degli aplogruppi (Fig. 7).

Nel novembre del 2005 è stato costituito il gruppo ISOGG (International Society of Genetic Genealogy) che ha creato un documento web, che comprende una lista ordinata di marcatori e aplogruppi corrispondenti, che viene aggiornata per tenere il passo con i rapidi sviluppi nel campo (<http://www.isogg.org/tree/index10.html>). Il programma Haplogroup predictor impiegato per la determinazione degli aplogruppi nel presente lavoro (fruibile dal web al sito <http://www.hprg.com/hapest5/index.html#>) utilizza la nomenclatura proposta dal gruppo ISOGG.

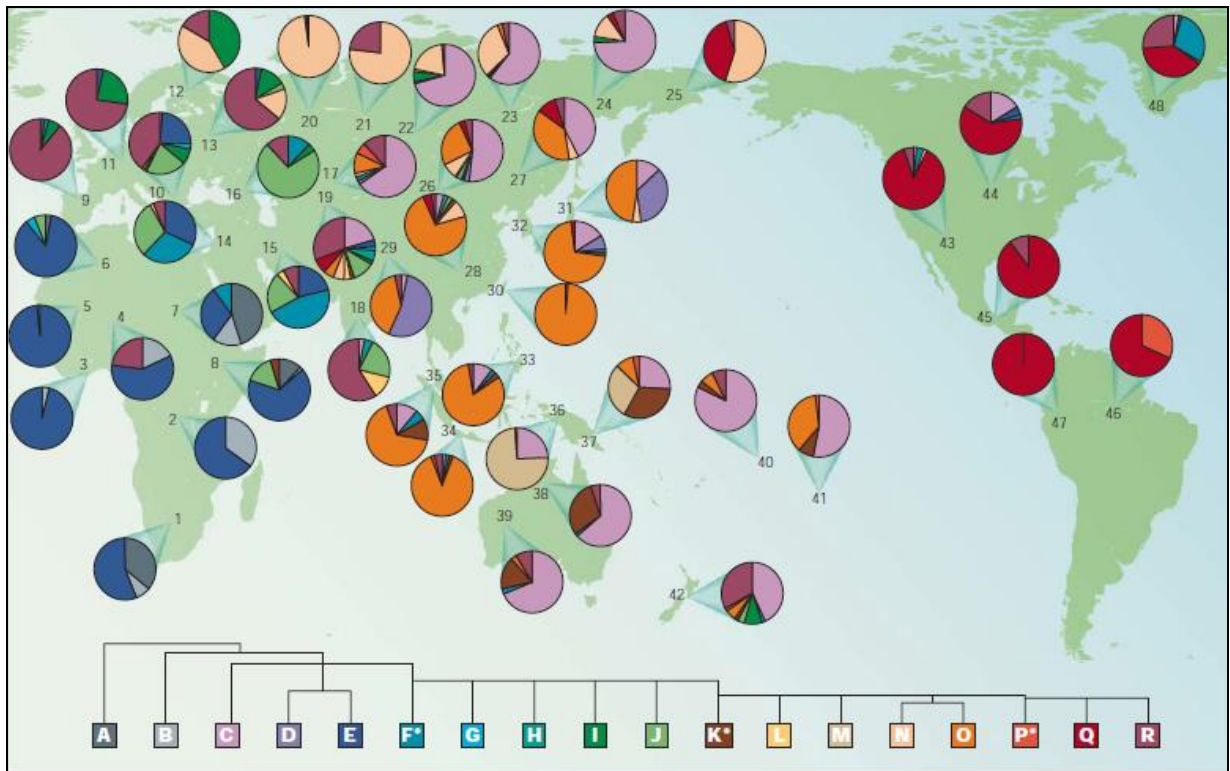


Figura 7. Distribuzione mondiale degli aplogruppi del cromosoma Y (Jobling and Tyler-Smith, 2003)

Si pensa che la linea evolutiva caratterizzata dalla mutazione M89 si sia originata in Africa orientale, a partire dalla linea originale M168 dell'Adamo Eurasiatico, circa 45.000 anni fa. Essa si sarebbe diffusa dall'Africa verso l'Asia occidentale, e da qui si sarebbe espansa dapprima in Medio Oriente e poi verso nord ed est circa 40.000 anni fa, fino a colonizzare gran parte del continente (Underhill et al., 2001). Probabilmente la penuria di selvaggina ha spinto i cacciatori all'esplorazione di nuove terre più ricche di prede. A partire da M89 si sono in seguito differenziate ulteriori linee caratterizzate ognuna da particolari polimorfismi. Due di tali linee, M170 e M173, sembrano essere state presenti in Europa sin dal paleolitico (Semino et al., 2000). Di queste, M170 sembra essere una linea tipicamente europea, in quanto virtualmente assente, se non teniamo conto dei recenti flussi genici, al di fuori del continente (Bosch et al., 2001). Lungo questa linea si trova la mutazione M26 che è caratteristica della Sardegna dove si trova con una frequenza di circa il 37%. (Ghiani et al., 2009; Contu et al., 2008; Francalacci et al., 2003; Semino et al., 2000; Passarino et al., 2001).

Un'altra linea molto rappresentata in Europa Occidentale (50-80%) è M173 che in Sardegna si trova con una frequenza del 22% (Semino et al., 2000).

1.7 DNA mitocondriale

Il mtDNA si trova all'interno dei mitocondri che sono degli organelli citoplasmatici deputati alla produzione di energia necessaria per tutte le funzioni cellulari. Il numero dei mitocondri varia da cellula a cellula (fino a 100), come pure varia il numero di copie di DNA presenti all'interno di ciascun mitocondrio (da 1 a 15).

Il grande numero delle molecole di mtDNA per cellula lo ha reso uno dei migliori strumenti d'indagine rispetto al DNA nucleare in quei campioni in cui il DNA può essere stato danneggiato dal tempo e dalle condizioni climatico-ambientali. La molecola di DNA mitocondriale è circolare ed è costituita da due filamenti superavvolti, entrambi codificanti, non associati a proteine istoniche, che differiscono per la composizione in basi (Fig. 4). Questi sono il filamento pesante (H), che è ricco di residui guaninici e codifica per 28 prodotti genici e il filamento leggero (L), ricco di residui citosinici, che trascrive gli 8 tRNA e l'enzima ND6 (NADH 6 deidrogenasi).

Il genoma mitocondriale è costituito da circa 16569 bp ed è caratterizzato dalla presenza di regioni codificanti contenenti polimorfismi dei frammenti di restrizione (RFLPs) ed una regione di controllo non codificante chiamata D-loop. Il nome D-loop (displacement loop) è dovuto alla presenza di un tratto di RNA quiescente che impedisce l'accoppiamento dei due filamenti.

È in quest'ultima regione che si osserva la maggiore variabilità fra gli individui (3%), a causa del fatto che non codifica per alcun prodotto necessario alle funzioni cellulari; quindi, le mutazioni si accumulano col tempo.

Invece, i tratti che possiedono il tasso di sostituzione più basso sono i geni per i tRNA e per gli rRNA. Il tasso di mutazione del DNA mitocondriale è da 5 a 10 volte maggiore rispetto a quello nucleare, poiché nel mitocondrio non esiste l'attività di correzione di bozze (proof-reading) da parte della DNA polimerasi mitocondriale durante la replicazione.

La D-loop (Fig.5), indicata anche come regione di controllo, lunga circa 1122 bp, contiene l'origine della replicazione per il filamento pesante del mtDNA (OH) e di essa fanno parte le regioni ipervariabili I, II e III (HVRI, HVRII e HVRIII). Le regioni ipervariabili I e II in totale si espandono per circa 900 bp.

La HVRI si trova tra le posizioni 16.024 e 16.365, mentre la HVRII si colloca tra le posizioni 73 e 340. Il segmento HVRIII, che si estende per circa 137 bp, viene più che altro utilizzato in campo forense e si trova tra la posizione 438 e la 574.

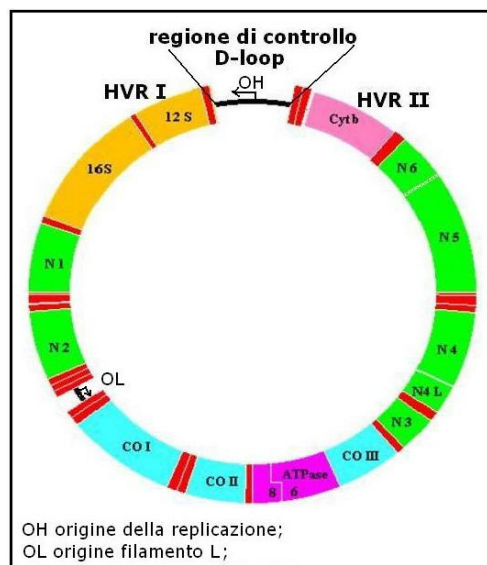


Figura 3. Molecola circolare di mtDNA.

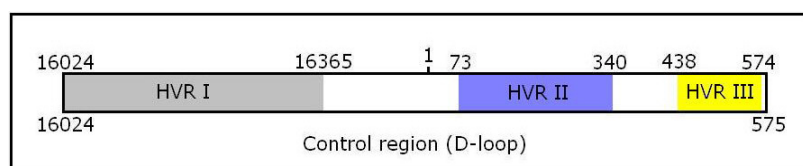


Figura 4. Segmenti HVRs del mtDNA

Il segmento ipervariabile I presenta una variabilità doppia rispetto al segmento ipervariabile II. Il rapporto tra transizioni e transversioni è piuttosto elevato (5:1), con alcune variazioni a seconda della regione interessata. Le transizioni sono sostituzioni nucleotidiche dello stesso tipo di base azotata, quindi, una purina con un'altra purina o una pirimidina con un'altra pirimidina, mentre le transversioni sono sostituzioni nucleotidiche tra tipi diversi di basi azotate come, per esempio, la sostituzione di una purina con una pirimidina e viceversa.

La regione HVRI, per il caratteristico grado di omologia con scimpanzé e con altri mammiferi, risulta essere il segmento più informativo per confronti tra popolazioni che si sono separate in tempi evolutivamente recenti (Caramelli, 2009).

I geni presenti sul mtDNA vengono trasmessi lungo la linea parentale materna dalla madre ai suoi figli siano essi maschi o femmine (Fig. 6). Il citoplasma, all'interno del quale sono contenuti i mitocondri, viene trasmesso allo zigote esclusivamente dalla cellula uovo. Ciò sarebbe dovuto al fatto che le molecole di mtDNA dello spermatozoo una volta entrate nella cellula uovo vengono degradate perché riconosciute estranee dai meccanismi di degradazione cellulari appena formati dell'embrione (Caramelli, 2009).

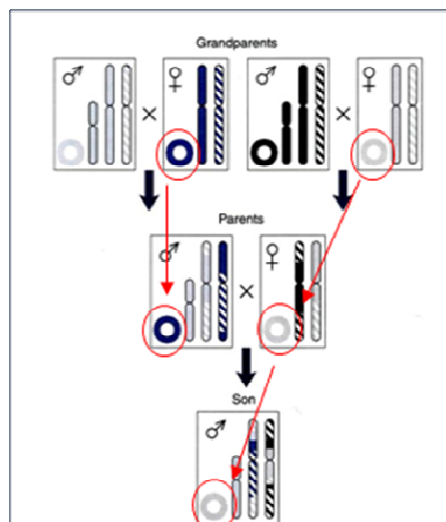


Figura 5. Ereditarietà materna del mtDNA

Lo studio delle mutazioni applicato all'analisi della struttura delle popolazioni può essere eseguito in due diversi modi: attraverso l'analisi delle regioni codificanti della molecola mediante gli enzimi di restrizione (che operano tagli in specifici siti) (Fig 7) e l'analisi delle sequenze delle regioni HVR I e HVR II della regione di controllo.

In quest'ultimo tipo di analisi la sequenza di interesse viene confrontata con una sequenza di riferimento chiamata CRS (Cambridge Reference Sequence) che è stata identificata da Anderson e collaboratori nel 1981, presso il MRC, Laboratory of Molecular Biology di Cambridge (Anderson et al., 1981) Si tratta della sequenza completa del mtDNA umano di un soggetto di nazionalità inglese (Fig. 8).

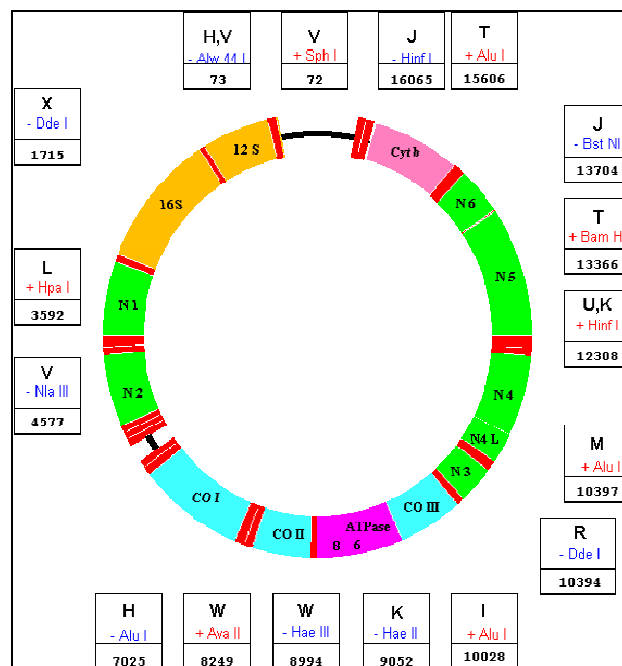


Figura 6. Siti di restrizione presenti nella regione codificante dell' mtDNA

```

1 gatcacaggt ctatcaccct attaaccact cacgggagct ctccatgcat ttggtatatt
61 cgtctggggg gtatgcacgc gatagcattg cgagacgctg gagccggagc accctatgtc
121 gcagtatctg tctttgattc ctgcctcadc ctattattta tcgcaacctac gtccaatatt
181 acaggcgaac atacttacta aagtgtgtta attaattaat gcttgtagga cataataata
241 acaattgaat gtctgcacag ccactttcca cacagacatc ataacaataa atttccacca
301 aacccccctt cccccgcttc tggccacagc acttaaacac atctctgcca aacccccaaa
[...]
16021 ctgttctttc atggggaagc agatttgggt accaccaagc tattgactca cccatcaaca
16081 accgctatgt atttcgtaca ttactgccag ccaccatgaa tattgtacgy taccataaat
16141 acttgaccac ctgtagtaca taaaaaccca atccacatca aaacccccct cccatgctta
16201 caagcaagta cagcaatcaa ccctcaacta tcacacatca actgcaactc caaagccacc
16261 cctcaccacc taggatacca acaaacctac ccacccttaa cagtacatag tacataaagc
16321 catttaccgt acatagcaca ttacagtcaa atcccccttc gtccccatgg atgaccccc
16381 tcagataggg gtcccctgac caccatcctc cgtgaaatca atatcccgca caagagtgct
16441 actctcctcg ctccggggcc ataacacttg ggggtagcta aagtgaactg tatccgacat
16501 ctggttccta ctccagggtc ataaagccta aatagcccac acgttccccct taaataagac
16561 atcacgatg

```

Figura 7. Sequenza di Anderson
(<http://www.mitomap.org/MITOMAP/HumanMitoSeq>)

1.7.1 Aplogruppi del DNA mitocondriale

Per quanto riguarda il DNA mitocondriale, l'aplotipo è costituito dalla somma dei siti delle variazioni della sequenza nucleotidica (Fig.9) e l'aplogruppo si definisce sulla base della condivisione di mutazioni specifiche in posizioni aventi un basso tasso di mutazione (mutazioni stabili). Quindi, aplotipi non identici possono appartenere allo stesso aplogruppo (Fig. 10).

Aplotipo rif.	16038	16126	16187	16189	16223	16264	16270	16278	16293	16300	16311
1	G	C	T	C	T	T	T	T	G	T	C
2		C	T	C	T	T	T	T	G		C
3		C	T	C	T	T	T	T			C
4		C	T	C	T	T	T	T			C

Figura 8. Esempio di aplotipo per il mtDNA

Aplotipo rif.	16038	16126	16187	16189	16223	16264	16270	16278	16293	16300	16311	Aplog.
1	G	C	T	C	T	T	T	T	G	T	C	L1b1
2		C	T	C	T	T	T	T	G		C	L1b1
3		C	T	C	T	T	T	T			C	L1b
4		C	T	C	T	T	T	T			C	L1b

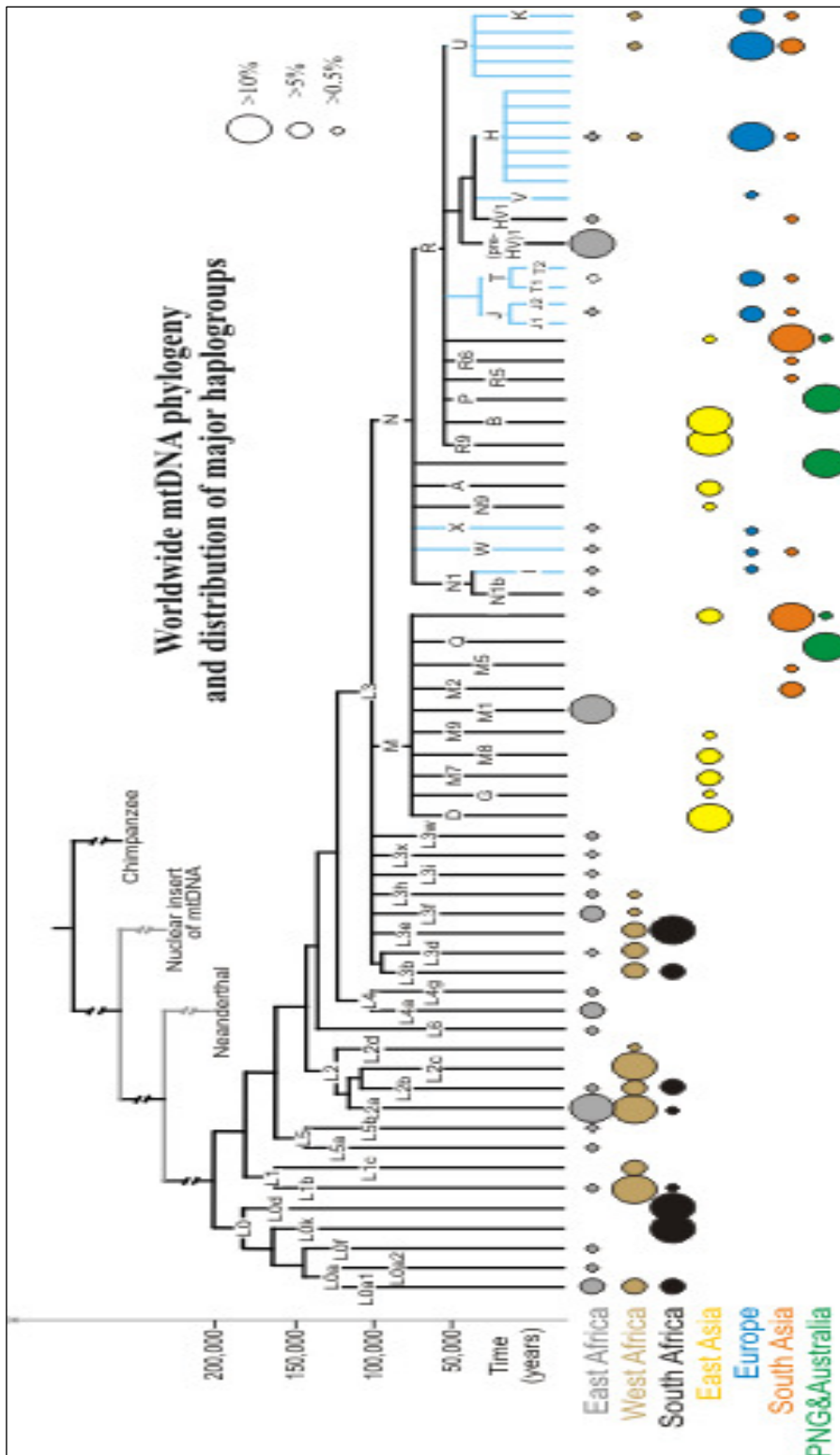
Figura 9. Esempio di aplogruppo per il mtDNA.

Nella figura 11 è riportato l'albero filogenetico del mtDNA con la distribuzione geografica degli aplogruppi. Come si può notare, vi è una forte connessione fra aree geografiche e aplogruppi, ovvero fra linee mitocondriali e migrazioni passate dell'uomo.

L'aplogruppo L rappresenta l'aplogruppo ancestrale, distribuito nel continente africano. L'aplogruppo M racchiude la prima migrazione umana fuori dall'Africa che raggiungerà le coste meridionali dell'Asia. La sua presenza si riscontra fra le popolazioni dell'Africa Orientale, in Asia, in America e Melanesia.

L'aplogruppo N è probabile che rappresenti un'altra grande migrazione dal continente africano che ha puntato verso nord. Dall'Aplogruppo N prende origine R che consiste di due sottogruppi: uno con distribuzione in Asia Sud-Orientale e in Oceania e l'altro che contiene quasi tutte le popolazioni europee.

Nella figura 12 è riportato l'albero filogenetico dell'aplogruppo H dove sono visibili i 15 sotto-aplogruppi (H1-H15) identificati da Achilli (Achilli et al. 2004). Questo è l'aplogruppo più rappresentato in Europa, con frequenze che si aggirano intorno al 40-50% (Quintans et al., 2004).



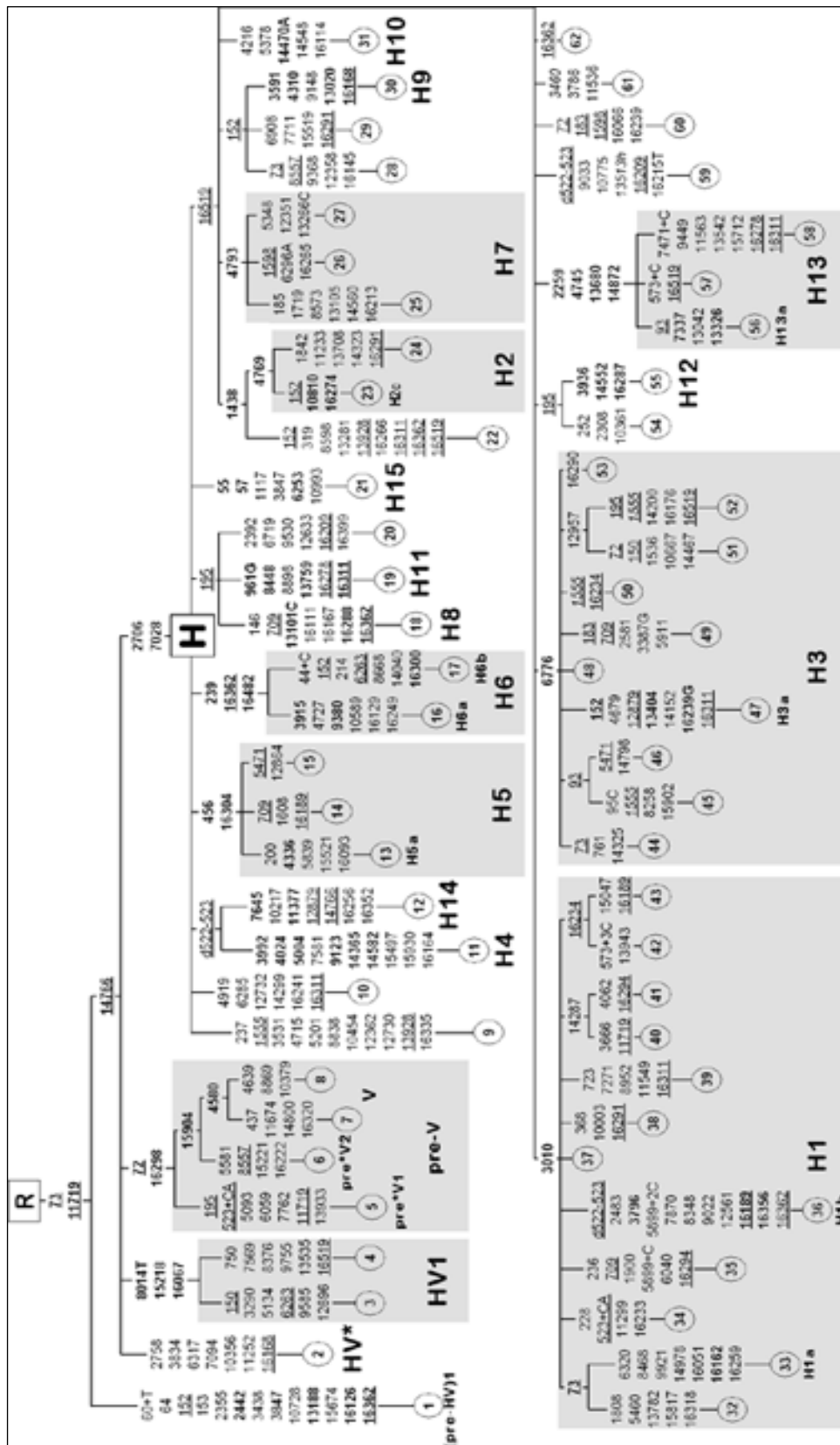


Figura 11. Albero filogenetico dell'aplogruppo H del mtDNA (Achilli et al.,2004)

1.8 Scopo del lavoro

L'intento della ricerca è quello di analizzare due isolati sardi, uno linguistico (Carloforte) e uno geografico (Benetutti), attraverso i marcatori del DNA uniparentali (mtDNA e cromosoma Y), con lo scopo di valutare il livello di isolamento e appurare se le barriere culturali e linguistiche abbiano la stessa influenza delle barriere geografiche sull'isolamento e, quindi, nella determinazione della struttura genetica di una popolazione.

Si procederà altresì all'analisi comparativa dei risultati ottenuti con i marcatori ad eredità materna e paterna per verificare se forniscano informazioni concordanti. Inoltre, per quanto riguarda Carloforte, si metteranno a confronto i risultati del DNA mitocondriale ottenuti con due differenti metodi di campionamento: quello in base ai cognomi fondatori e quello secondo le genealogie.

2. Materiali e metodi

2.1 Campionamento

Le due comunità sarde che si è deciso di analizzare sono state selezionate sulla base dei seguenti criteri di isolamento: esiguo numero di fondatori, alto tasso di endogamia e consanguineità, espansione della popolazione bassa o nulla ed immigrazione trascurabile. Il campionamento ha interessato individui maschi adulti apparentemente sani, non legati fra loro da vincoli di parentela sino al terzo grado, che sono stati scelti in base alla genealogia del cognome, ossia preferendo quei soggetti che presentavano il cognome dei fondatori. Per la precisione è stato campionato un solo individuo per ciascun cognome fondatore selezionato. Ogni soggetto che ha partecipato al campionamento ha firmato un regolare consenso informato per l'autorizzazione al trattamento dei dati.

Per quanto riguarda la comunità di Benetutti i cognomi fondatori sono stati individuati grazie all'analisi del *Liber Baptizatorum*, il più antico dei *Quinque Libri* presente negli archivi parrocchiali della Chiesa di Sant'Elena. Nello specifico sono stati analizzati i battesimi registrati dal 1600 al 1700. Da tale analisi sono stati individuati 217 cognomi diversi presenti nella popolazione di quegli anni, dei quali 70 sono ancora rintracciabili nella comunità attuale. I cognomi fondatori di Carloforte sono invece stati acquisiti dalla bibliografia storica (Vallebona, 1974 ; Ferraro, 1989). Secondo queste fonti i cognomi fondatori erano complessivamente 55, quasi tutti presenti ancora oggi.

Per il comune di Carloforte sono stati campionati 43 individui e per Benetutti 50. Inoltre, sono stati selezionati due campioni di confronto nelle zone più prossime alle aree di riferimento dei due villaggi analizzati. A tal fine, sono stati campionati 50 individui maschi della zona del Sulcis Iglesiente (Sud Sardegna) e 50 individui maschi di vari paesi del Nord Sardegna

appartenenti alle aree del sud-Logudoro-Planargia, del Monte Ferru e della Gallura. Il campionamento di Benetutti e del rispettivo campione di confronto è stato effettuato mediante prelievo salivare tramite tampone buccale, mentre per Carloforte e il rispettivo campione di controllo si è utilizzato il sangue.

2.2 Estrazione del DNA da saliva

Per estrarre il DNA dal tampone buccale si è utilizzato il QIAamp® DNA mini kit fornito dalla QIAGEN, che ha il vantaggio di essere un metodo molto rapido, di facile utilizzo che permette di ottenere un prodotto altamente purificato. In un primo passaggio si provoca la lisi delle membrane cellulari e la denaturazione delle proteine cellulari ed extracellulari. A tal fine nei tubicini da 2 ml contenenti lo spazzolino si versano 500µl di PBS (tampone fosfato salino, Ph 7.4), 20 µl di proteinasi K , 500 µl di Buffer AL (Buffer di lisi), si passano immediatamente al vortex per 15 secondi e si incubano a 56°C per 10 minuti.

Tolti i campioni dal termostato si centrifugano per far scendere le gocce formatesi nella superficie interna del tappo e si procede aggiungendo 500 µl di etanolo (96-100%) e nuovamente vortexando per 15" e centrifugando per pochi secondi. A questo punto si procede trasferendo il contenuto del tubo da 2ml (in due momenti successivi, ciascuno da 700 µl) nella colonnina fornita nel Kit. In tale colonnina è contenuta una membrana di gel di silice in cui viene adsorbito il DNA (in seguito a centrifugazione 8000 rpm). A questo punto il DNA legato alla membrana viene lavato con due buffer di lavaggio (AW1 a 8000 rpm per 1' e AW2 a 14000 rpm per 3') per eliminare tutte le componenti cellulari e successivamente eluito dalla membrana per mezzo di un buffer di eluizione. Per quest'ultimo passaggio si posiziona la colonnina contenente la membrana di silice in un tubicino da 1.5 ml, si versano 100 µl di

buffer di eluizione, la si lascia incubare a temperatura ambiente per 5' e poi si centrifuga a 8000 rpm per 1'.

2.3 Estrazione del DNA da sangue intero

Per estrarre il DNA dal sangue intero è stata utilizzata la metodica fenolo-cloroformio, che permette di rimuovere le proteine e i frammenti cellulari derivanti dalla rottura della membrana cellulare e degli organuli cellulari, ottenendo in tal modo un ottimo quantitativo di DNA a doppio filamento con un elevato grado di purezza, il tutto con un basso dispendio economico. Per contro, la metodica richiede tempi molto lunghi e i due reagenti principali (fenolo e cloroformio) sono sostanze chimiche pericolose che vanno maneggiate con cautela.

La prima fase consiste nell'incubare 100 µl di sangue intero, a 37°C per 2h, con 150 µl di un apposito tampone di lisi (31.25 µl SDS al 10%, 6.25 µl DTT 1M, 0.3 µl proteinasi K, 112.2 µl STE). Successivamente, si procede aggiungendo 250 µl di Fenolo, agitando lentamente ciascun campione fino a quando non assume una colorazione marrone, mettendo in ghiaccio per 10' e centrifugando a 13000 rpm. I campioni così trattati presentano una divisione bifasica, in cui, nella fase superiore acquosa si trova il DNA e in quella inferiore organica le proteine. Questa ripartizione è dovuta al fatto che il fenolo è un liquido pesante che interagisce con le regioni scarsamente polari delle proteine, non miscibile con l'acqua, mentre il DNA mostra una scarsa solubilità nel fenolo.

Una volta ottenuta questa soluzione bifasica, si recupera il DNA spillando il surnatante e lo si pone in un nuovo tubicino, lo si tratta con 250 µl di una soluzione di fenolo - cloroformio - alcool isoamilico (in rapporto 25:24:1), si agita lentamente a mano e si centrifuga a 13000 rpm per 10'. Il risultato di questa operazione consisterà nella formazione di una nuova

soluzione bifasica di cui si spilla il surnatante e lo si tratta con 250 µl di una miscela di cloroformio - alcool isoamilico, agitando a mano il campione, centrifugando a 13000rpm per 10' per eliminare ogni traccia di fenolo presente nella fase acquosa. Dopodiché, si fa precipitare il DNA con 500 µl di etanolo ghiacciato (centrifuga 13000 rpm per 10'), si elimina il surnatante e si lava il pellet aggiungendo 500 µl di etanolo al 70% (13000rpm per 10') al fine di rimuovere gran parte dei cationi monovalenti che possono interferire con gli enzimi usati nelle successive manipolazioni.

Infine, si elimina il surnatante, si fa evaporare tutto l'etanolo e si sospende il sedimento in 100 µl di TE a 37°C per tutta la notte.

2.4 Determinazione quantitativa del DNA

Una volta estratto, il DNA genomico è stato quantificato mediante l'utilizzo del fluorimetro fornito dalla Invitrogen (QubitTM). Questa metodica ha il vantaggio di essere molto rapida e semplice, richiede l'utilizzo di poco DNA, fornisce un risultato accurato ed ha costi ridotti.

La procedura richiede prima di tutto la calibrazione dello strumento mediante l'utilizzo di due appositi standard presenti nel Quant-iT ds DNA Hs Assay Kit (i reagenti di tale Kit sono stati progettati per eseguire l'esperimento a una temperatura ambiente di circa 22-28°C). Successivamente, si dispone 1 µl del DNA da testare in un tubicino da 0.5 ml e si prepara la mix che è costituita da 199 µl di Quant-it Buffer e da 1 µl di Quant-it Reagent (quantità per ciascun campione). La mix viene passata al vortex per 3 secondi e distribuita fra i vari campioni da testare (il volume finale di ciascun tubicino deve essere 200 µl). Si passano al vortex i campioni per 3 secondi, si incubano per 2 minuti a temperatura ambiente e si procede

con la lettura al fluorimetro. Il valore che si visualizza nello strumento è espresso in ng/ml. Il valore così ottenuto viene poi convertito in ng/μl

2.5 Amplificazione

La PCR, o reazione a catena della DNA polimerasi è una metodica di laboratorio, ideata verso la metà degli anni '80 da Kary Mullis, che consente di realizzare in vitro milioni di copie di una specifica sequenza di DNA da una miscela di DNA, per mezzo di un processo chiamato amplificazione. Per eseguire la PCR è necessario conoscere le estremità del frammento di DNA da amplificare, così da poter sintetizzare due primers oligonucleotidici a singolo filamento complementari uno all'estremità 3' e l'altro all'estremità 5' del frammento di DNA che si vuole amplificare. I primers si comportano da inneschi per l'attività della DNA polimerasi che riconosce l'estremità 3' libera e inizia le sintesi del DNA in direzione 5'→3'.

Ogni ciclo di PCR è diviso in tre fasi successive:

Denaturazione: il DNA a doppio filamento viene riscaldato a 94-95°C per far avvenire la denaturazione e cioè la separazione dei due filamenti.

Annealing: si ha l'appaiamento dei *primers*, la temperatura varia solitamente tra i 50 e 68°C a seconda della sequenza oligonucleotidica e del contenuto in guanine e citosine.

Estensione: avviene ad una temperatura fra i 68 e i 72°C e porta alla sintesi di una nuova elica complementare per opera di una speciale DNA polimerasi resistente al calore isolata dal batterio *Thermus aquaticus*.

Ad ogni ciclo il numero di copie della sequenza specifica compresa tra i due primers raddoppia e il DNA del campione e il DNA neo amplificato fungono entrambi da filamento stampo per i cicli successivi. In questo modo, la PCR procederà con andamento esponenziale.

Dopo 30 cicli si è formato quasi un miliardo di copie della sequenza di DNA situata fra i due inneschi. Tuttavia, superati i 35-40 si osserva un effetto *plateau* dovuto al depauperamento dei reagenti e all'inibizione causata dall'accumulo del pirofosfato, per tale motivo non si procede oltre tale limite.

2.6 Analisi del DNA mitocondriale

Per quanto riguarda il DNA mitocondriale, si è proceduto amplificando i due segmenti ipervariabili (HVRI e HVRII), seguendo le condizioni riportate nelle tabelle 3 e 4 e utilizzando i *primers* riportati nella tabella 2. La lunghezza di tali amplificati era rispettivamente di 750 e 478 bp. Successivamente, le regioni HVRI e HVRII amplificate sono state sequenziate da strutture esterne utilizzando il *primer forward* specifico per ciascuna regione ipervariabile. Le sequenze di Benetutti e del rispettivo campione di controllo sono state effettuate dalle ditte Genoscreen di Lille (Francia) e Shardna di Pula (Cagliari), mentre quelle di Carloforte e della zona limitrofa sono state eseguite presso l'Università "La Trobe" di Melbourne (Australia).

Le sequenze ottenute sono state esportate in formato *fasta* tramite il programma Chromas (Ver. 2.33). Le sequenze sono state poi allineate con la sequenza di Anderson attraverso il software ClustalX (Higgins et al., 1992). Una volta creato il file di allineamento, questo è salvato in formato **aln* ed esportato in DAMBE (<http://web.hku.hk/~xxia/software/software.htm> Xia, X. 2000). Questo programma permette l'editing delle sequenze nucleotidiche allineate.

Inoltre, si è proceduto alla determinazione degli aplogruppi: in prima analisi con l'ausilio di Haplogroup Predictor, che ha permesso di fare una prima attribuzione

(<http://nnhgtool.nationalgeographic.com/classify/index.html>) e in seconda analisi mediante Phylotree, l'albero filogenetico umano del DNA mitocondriale disponibile in rete al sito <http://www.phylotree.org/tree/main.htm>, che ha permesso una assegnazione più dettagliata degli aplogruppi.

Per i campioni che sono risultati appartenere all'aplogruppo H, vista la difficoltà nell'attribuzione con phylotree si è proceduto alla verifica analizzando il polimorfismo 7025 AluI localizzato nella regione codificante della molecola. Le sequenze dei primers sono riportate nella tabella 5 e le condizioni utilizzate per la PCR e per la digestione mediante enzima di restrizione sono riassunte nelle tabelle 6, 7 e 8 (Helgason et al., 2000).

mtDNA	PRIMER	Ampl.
HVRI	5' CTCCACCATTAGCACCCAAAGC 3' for 5'AATAGGATGAGGCAGGAATCAA 3' rev	750bp
HVRII	5' TAAAGCCTAAATAGCCCACA 3' for 5 5'CTGTAAAAGTGCATACCGCCA 3' rev	478bp

Tabella 2. Sequenza primer mtDNA HVRI e HVRII

REAGENTI	CONCENTRAZ. REAGENTI	QUANTITÀ UNITARIA (µl)
Buffer	1X *	2.5
MgCl₂	1.5 mM	0.75
dNTPs	0.2 mM per ognuno	4
Primer for	0.5 µM	0.75
Primer rev	0.5 µM	0.75
Taq	1 U	0.15
H₂O		14.6
Dna		1.5
		25 Vol.finale

Tabella 3. Protocollo amplificazione mtDNA (HVRI e HVRII). *15 mM Tris-HCl, pH 8.0, 50 mM KCl

	TEMPO	TEMPER.(°c)	
Denaturazione	5'	95	
Denaturazione	30''	95	25 cicli
Annealing	30''	55	
Estensione	10'	72	
Denaturazione	30''	95	35 cicli
Annealing	30''	55	
Estensione	90''	72	
Estensione	10'	72	
Stop		4	

Tabella 4. Condizioni PCR per il mtDNA (HVRI, HVRII)

mtDNA	PRIMER	AMPL	Bibliografia
6909F	5' AAGCAATATGAAATGATCTG 3' for	243bp	Helgason et al 2000
7115R	5' CGTAGGTTTGGTCTAGG 3' rev		

Tabella 5. Sequenze primers per il polimorfismo 7025 Alu I

REAGENTI	CONCENTRAZ. REAGENTI	QUANTITÀ UNITARIA (µl)
Buffer	1X *	2.5
MgCl ₂	1.5 mM	1
dNTPs	0.2 mM per ognuno	1
Primer for	0.5 µM	0.5
Primer rev	0.5 µM	0.5
Taq	1 U	0.1
H ₂ O		18.4
Dna		1
		25 Vol. finale

Tabella 6. Protocollo per per il polimorfismo 7025 Alu I. *15 mM Tris-HCl, pH 8.0, 50 mM KCl

	TEMPO	TEMPER.(°c)	
Denaturazione	1'	94	
Denaturazione	30''	94	30 cicli
Annealing	30''	50	
Estensione	5'	72	
Stop		4	

Tabella 7. Condizioni per per il polimorfismo 7025 Alu I

Mix di digestione		
Reagenti	Concentrazione	Quantità per campione (µl)
Ampl.		10
Buffer	10 u/ µl	2
Enzima	10 X (Alu I)	0.2
H ₂ O		7.8

Tabella 8. Mix di digestione per il polimorfismo 7025 Alu I

2.7 Analisi del cromosoma Y

Per lo studio del cromosoma Y si è iniziato con l'analisi di 17 *loci* STR attraverso una PCR-multiplex, utilizzando il *Kit AmpFISTR Yfiler* (Applied Biosystems). Tale Kit analizza:

- i *loci* che definiscono “l'European minimal haplotype”: DYS19, DYS385 a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393;
- i *loci* raccomandati dal Scientific Working Group-DNA Analysis Methods: DYS438, DYS439;
- *loci* addizionali altamente polimorfici: DYS437, DYS448, DYS456, DYS458, DYS635 (Y GATA C4) e Y GATA H4.

L'amplificazione è stata eseguita tramite il termociclatore Geneamp®PCR System 9700 (Applied Biosystems). Per la separazione dei frammenti amplificati è stato utilizzato il sequenziatore automatico ABI Prism 3100. I dati sono stati poi analizzati tramite il software GeneScan 3.1.

Tali analisi sono state condotte dal Reparto Investigazioni Scientifiche dei Carabinieri di Cagliari.

Mediante l'utilizzo del programma Haplogroup predictor (<http://www.hprg.com/hapest5/>) per mezzo degli aplotipi STRs è stato possibile determinare gli aplogruppi di appartenenza degli individui.

Gli aplogruppi che non si è riusciti a individuare con tale metodo sono stati determinati nel nostro laboratorio mediante l'analisi di alcuni marcatori biallelici (Y-SNPs ed elementi Alu). Si è prima di tutto scelto quali SNPs indagare e poi si è proceduto ad amplificare e digerire i campioni.

Per quanto riguarda le sequenze dei primers e le condizioni utilizzate per amplificare i vari polimorfismi del cromosoma Y, si possono consultare le tabelle 9, 10, 11, 12. Per l'amplificazione degli SNPs del cromosoma Y è stato utilizzato un programma della PCR chiamato touch-down, in cui, durante i vari cicli, la temperatura viene diminuita ciclo dopo ciclo per rendere instabili eventuali appaiamenti dei primer a sequenze non perfettamente complementari, rendendo quindi la PCR più specifica.

MARKERS	PRIMERS	AMPLIFICATO (bp)	Bibliografia
M9	5'-GCAGCATATAAACTTTTCAGG-3' 5'-AAAACCTAACTTTGCTCAAGC-3' rev	340	Underhill et al., 2001
M170	5'-CTATTTTATTTACTTAAAAATCATTGATC-3' for 5'-AGACCACACAAAAACAGGTC-3' rev	88	Flores et al.,2003
M172	5'-AAATTAGGAGCCAGATGACC-3' for 5'-AATAATAATTGAAGACCTTTTGAGT-3' rev	176	Flores et al.,2003
M173	5'-AAGTTGATGCCACTTTTCAG-3' for 5'-TTCTGAATATTAACAGATCACAAAG-3' rev	199	Flores et al.,2003
M201	5'-CTAATAATCCAGTACCAACTGAGG-3' for 5'-TGAAAGTTCAAACGTCAAACAG-3' rev	207	Flores et al.,2003
YAP	5'-CAGGGGAAGATAAAGAAATA-3' for 5'-ACTGCTAAAAGGGGATGGAT-3' rev	Yap + → 455 Yap - → 150	Hammer et al.,1995

Tabella 9. Polimorfismi cromosoma Y analizzati

Marker	Buffer	MgCl ₂	DNTPs	P1	P2	BSA	H ₂ O	Taq	DNA	Tot
M9	1.5	0.5	0.75	0.52	0.52	1	8.41	0.3	1.5	15
M170	2.5	0.75	1	0.7	0.7		17.15	0.2	2	25
M172	2.5	0.75	1	0.75	0.75		18.05	0.2	1	25
M173	2.5	0.75	1	1	1		16.55	0.2	2	25
M201	2.5	0.75	1	0.75	0.75		18.05	0.2	1	25
YAP	2.5	1	0.3	0.15	0.15		18.7	0.2	2	25

Tabella 10. Protocollo di amplificazione. Le quantità sono espresse in µl.

	TEMPO	TEMPER.(°c)	
Denaturazione	5'	94	14 cicli*(-0.5°C ogni ciclo)
Denaturazione	20"	94	
Annealing	1'	63 *	
Estensione	1'	72	
Denaturazione	20"	94	35 cicli
Annealing	45"	56	
Estensione	90"	72	
Estensione	10'	72	
Stop	∞	4	

Tabella 11. Touch-down per.

	TEMPO	TEMPER.(°c)	
Denaturazione	3'	94	30 cicli
Denaturazione	1'	94	
Annealing	1'	51	
Estensione	1'	72	
Stop	∞	15	

Tabella 12. Condizioni PCR YAP

2.8 Digestione

L'amplificazione degli SNPs del cromosoma Y è seguita dalla digestione che viene effettuata mediante degli enzimi di restrizione, i quali riconoscono, quando presente, un sito specifico sul frammento di DNA amplificato e tagliano in quella specifica posizione. L'attività dell'enzima si misura in Unità di Attività enzimatica: la quantità di enzima che è in grado di tagliare completamente 1 grammo di DNA in un'ora alla temperatura adatta (37°C). Gli enzimi di restrizione utilizzati per ciascun polimorfismo e i frammenti generati sono riportati

nella tabella 14. I reagenti adoperati per la mix di digestione e le rispettive quantità sono elencati nella tabella 13.

Mix di digestione		
Reagenti	Concentrazione	Quantità per campione (μ l)
Ampl.		10
Buffer	10 u/ μ l	2
Enzima	10 X	0.2
H₂O		7.8

Tabella 13. Mix di digestione SNPs cromosoma Y.

Marker	Amplificato	Enz.restriz.	°C	Frammenti	Bibliografia
M9	340	Hinf I	37	181/64/95 (C) → (G) 245/95	
M170	88	Bcl I	55	63/25 (A) → (C) 88	Flores et al.,2003
M172	176	Hinf I	37	176 (T) → (G) 151/25	Flores et al.,2003
M173	199	Dra III	37	178/21 (A) → (C)199	Flores et al.,2003
M201	207	BseL I	55	108/78/21 (G) → (T) 129/78	Flores et al.,2003

Tabella 14. Protocollo di digestione SNPs Y

2.9 Elettroforesi su gel di agarosio

Ultimata la PCR, i frammenti di DNA amplificati vengono testati mediante l'elettroforesi su gel di agarosio al 2%. L'elettroforesi è una tecnica di analisi chimica che permette la migrazione di sostanze in soluzione (come gli acidi nucleici) per effetto di un campo elettrico. La preparazione del gel viene descritta di seguito. Si procede pesando una certa quantità di agarosio che viene diluita in un tampone di TAE. Questo preparato viene riscaldato fino a diventare un liquido limpido, raffreddato a circa 60°C e infine, sotto la cappa, colorato con bromuro di etidio. Tale liquido viene versato in una vaschetta provvista di pettinini movibili e dopo circa 15 minuti solidifica (il gel polimerizza a 40°C). A questo punto, si possono rimuovere i pettinini che hanno permesso la formazione di una serie di pozzetti, all'interno dei

quali verranno seminati i campioni. La vaschetta contenente il gel solidificato viene immersa in una cella da elettroforesi colma di una soluzione tampone (TAE) che contiene gli ioni necessari a condurre l'elettricità.

Si procede seminando in ogni pozzetto 5 μ l di DNA precedentemente miscelato con 1 μ l di Bromo Fenolo Blu che funge da colorante e appesantente. Dopo un certo numero di campioni seminati in un determinato pozzetto si semina un marcatore di peso molecolare (0.8 μ l). A questo punto, viene fornita la corrente ai due elettrodi, così da creare una differenza di potenziale di 100-150 Volt, che permette al DNA di migrare dal catodo verso l'anodo grazie alla carica negativa conferitagli dai gruppi fosfato.

Dopo circa 20 minuti tale corsa termina e il gel viene collocato in un transilluminatore a raggi ultravioletti che permette di visualizzare i frammenti di DNA grazie al bromuro di etidio che è una sostanza fluorescente a 590nm. Questa sostanza si lega al DNA intercalandosi fra le basi e permettendo così la visualizzazione delle bande.

2.10 Elettroforesi su gel di poliacrilammide

Per la lettura degli SNPs del cromosoma Y è stato utilizzato un gel di poliacrilamide al 10%, dal momento che permette di discriminare bande che differiscono di pochissime paia di basi.

Per la preparazione di una piastra si adoperano:

- Acrilamide-BisAcrilamide: 1.65 ml
- TBE 1X: 5 ml
- APS (20%): 23.25 μ l
- TEMED: 4.74 μ l

Il gel, una volta preparato, viene iniettato tra due piastre di vetro distanziate da un separatore di plastica grazie al quale si crea un'intercapedine in cui si formerà un gel sottilissimo di circa 0.75 mm. Sulla parte alta, si dispone un pettinino per la formazione dei pozzetti. Dopo circa 50 minuti a temperatura ambiente il gel è polimerizzato.

A questo punto, si estrae il pettinino, si sistema la piastra nella cella elettroforetica verticale (riempita con TBE 1X) e si procede con la semina. In ogni pozzetto vengono seminati 8 µl di campione digerito, ai quali vengono aggiunti 2µl di Bromo Fenolo Blu. Per identificare l'altezza delle bande viene seminato anche un marcatore di peso molecolare noto. Successivamente, si applica una differenza di potenziale di 200 Volt e una volta terminata la corsa si dispone il gel in una soluzione di bromuro di etidio per 15 min. Poi, grazie al transilluminatore a luce UV si effettua la lettura.

2.11. Analisi statistica

2.11.1 Frequenze alleliche e aplotipiche.

Le frequenze alleliche sono state calcolate tramite il metodo della conta degli alleli. Considerati due o più loci polimorfici vicini, i loro possibili assortimenti allelici sono detti aplotipi. Se l'aplotipo t è osservato x_t volte in un campione costituito da n individui, allora la frequenza stimata (p_t) è data da:

$$p_t = x_t / n$$

2.11.2 Analisi degli aplotipi

I dati aplotipici sono stati analizzati con il programma Arlequin (ver.3.1) (<http://anthro.unige.ch/software/arlequin/>), il quale analizza una vasta quantità di marcatori genetici grazie a numerosi metodi e test statistici che vengono fatti in ordine per ricavare informazioni sulle caratteristiche geniche e demografiche delle popolazioni in esame. Di seguito vengono indicati alcuni parametri individuati.

Il numero di aplotipi è un parametro descrittivo che indica il numero dei diversi aplotipi osservati nella popolazione.

L'indice standard di diversità (*Gene diversity*), è la diversità genetica che si riferisce a qualsiasi variazione di nucleotidi, dei geni, dei cromosomi. Esso è rappresentato dalle differenze nella sequenza nucleotidica che formano il DNA. Maggiore è il valore assunto dalla diversità genetica, maggiore sarà la diversità tra le popolazioni.

Il Mean number of pairwise differences è il numero medio di differenze fra coppie di sequenze.

L'Average gene diversity over loci è la media della diversità genica tra loci.

Il grado di differenziamento delle due popolazioni è stato valutato tramite il parametro F_{ST} (Fixation Index-Statistics) che misura la strutturazione di una popolazione. Questo indice misura la deviazione delle frequenze osservate dei genotipi eterozigoti rispetto a quanto atteso secondo il principio di Hardy-Weinberg. Il valore di questo indice varia fra 0 e 1. Quando vi è poca differenziazione, dovuta a flusso genico tra le popolazioni il valore è vicino a 0, se al contrario sono molto differenziate il valore è prossimo a 1.

2.11.3 Network filogenetici

I network filogenetici sono stati realizzati con il programma Network (ver. 4.5.1.6). Sono ricostruzioni filogenetiche che risultano particolarmente adatte per rappresentare processi evolutivi in cui i vari taxa non si siano sempre e solo separati gli uni dagli altri ma si siano anche mescolati.(Caramelli, 2009)

Gli aplotipi (sequenze di DNA o profili di micro satelliti) vengono confrontati fra di loro e viene costruita una matrice in cui sono indicati i passi mutazionali che separano ogni aplotipo dall'altro. Si uniscono prima i taxa che sono separati da un unico passo mutazionale, e poi si procede unendo i taxa separati da passi mutazionali via via sempre più grandi, fino a che tutti i taxa non sono uniti in un'unica struttura. Nel grafico che si viene a formare gli aplotipi sono rappresentati mediante cerchi la cui area è direttamente proporzionale alla frequenza dell'aplotipo nella popolazione. Le mutazioni sono rappresentate con delle linee.

2.11.4 Distanze genetiche

Per le matrici delle distanze è stato usato il programma Arlequin, pairwise genetic distances. Come matrici sono stati utilizzati i valori di F_{st} . L' F_{st} a coppie può essere usato come uno short-term di distanze genetiche fra le popolazioni, con l'applicazione di una leggera trasformazione che linearizza le distanze genetiche con il tempo di divergenza delle popolazioni (Reynolds et al., 1983; Slatkin, 1995). Nel metodo linearizzato da Slatkin nel 1995 si considera un modello demografico dove due popolazioni aploidi di taglia N si sono separate da t generazioni fa da una popolazione di taglia identica. Inoltre, queste due popolazioni si sono mantenute separate, senza scambi determinati da flusso genico. Sotto tale condizione, F_{st} può essere espresso in termini di tempo di coalescenza t_1 , che è il tempo di

coalescenza di due geni presi da due differenti popolazioni, e t_0 che è il tempo di coalescenza di due geni provenienti dalla stessa popolazione.

2.11.5 Albero filogenetico

Dalla matrice delle distanze è stato possibile costruire un dendrogramma con il metodo Neighbor-joining (Saitou e Nei, 1987). Il neighbor-joining, assume la lunghezza delle branche come una associazione tra le popolazioni e costruisce un albero per successivi raggruppamenti di popolazioni. L'albero non assume un orologio evolutivo e per tale motivo è senza radici. Le distanze genetiche e l'albero sono stati ottenuti usando il programma Phylip (ver. 3.1). La visualizzazione dell'albero è stata fatta tramite il programma TREEVIEW (win 32).

Per stimare la significatività o robustezza dell'albero costruito si può utilizzare il metodo bootstrap. Questo metodo si basa sull'idea che se un certo set di dati supporta fortemente una determinata ricostruzione filogenetica allora ogni sottoinsieme di quegli stessi dati dovrebbe supportare la stessa ricostruzione. Il programma crea uno pseudoallineamento con il quale costruisce un nuovo albero e lo confronta con l'albero originale. Questo procedimento viene effettuato da un minimo di 100 volte a un massimo di 2000 volte per garantire un intervallo di confidenza del 95% nei valori di bootstrap. Alla fine del processo, oltre all'albero originale, compare la frequenza con la quale i clade sono stati individuati anche negli alberi generati dal programma. Solitamente si accettano nodi che presentano valori di bootstrap superiori al 50%.

2.11.6 Struttura genetica

La struttura genetica delle popolazioni è stata studiata anche mediante il programma STRUCTURE, ver. 2.3.1, (Pritchard et al., 2000) che utilizza un modello basato sulla clusterizzazione. Tale programma è scaricabile gratuitamente dal sito http://pritch.bsd.uchicago.edu/software/structure2_2.html.

Il modello assume la presenza di K popolazioni (dove K può essere sconosciuto), ognuna della quali è caratterizzata da un set di frequenze alleliche per ciascun locus. I singoli individui sono assegnati in modo probabilistico ad uno, a due o più cluster.

Le analisi sono state condotte in accordo con il modello “admixture”, con K da 2 a 4 e facendo fare al programma più analisi per ogni K (esattamente 4).

3. Risultati

3.1 Risultati Cromosoma Y

3.1.1. Frequenze alleliche

Le frequenze alleliche dei 17 loci STR localizzati sul cromosoma Y (DYS19, DYS385 a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439, DYS437, DYS448, DYS456, DYS458, DYS635 e Y GATA H4) sono riportate nelle tabelle 15, 16, 17, 18. In generale, si può dire che in tutte le popolazioni analizzate i marcatori si sono rivelati polimorfici. Nei campioni di Benetutti e del nord Sardegna hanno mostrato da 3 (DYS391, DYS437) a 8 alleli (DYS458), nei campioni di Carloforte da 3 (DYS391, DYS389I, YGATAH4) a 6 alleli (DYS456) e in quelli del Sulcis Iglesiente da 2 (DYS391) a 8 alleli (DYS458).

A Benetutti, ad eccezione dei loci DYS392, DYS19 e DYS458 che presentano una distribuzione bimodale, tutti gli altri loci STR hanno mostrato una distribuzione unimodale. Il locus DYS458 presenta l'allele non-consenso 18,2 (0,041). Una duplicazione è invece presente nei locus DYS19 (15-16) e DYS448 (19-20) con una frequenza rispettivamente di 0,020 e 0,041. Il locus DYS385a/b ha mostrato 19 genotipi, di cui i più frequenti sono 12-13 (0,187) e 12-12 (0,167). Nel campione del nord Sardegna i loci hanno tutti un andamento unimodale eccetto DYS438, DYS392, DYS456, DYS19, DYS448 e DYS635. Il locus DYS458 presenta l'allele non-consenso 18,2 (0,020). Il locus DYS19 presenta la duplicazione 15-16 con una frequenza del 0,020. Il locus DYS385a/b ha mostrato 23 genotipi, di cui i più frequenti sono 12-12 (0,145) e 12-13 (0,125). Il locus DYS385 mostra una variazione a due loci simultaneamente i quali non possono essere analizzati separatamente e per tale motivo

vengono considerati come un genotipo, come raccomandato dalla DNA Commission of International Society of Forensic Genetics.

A Carloforte hanno un andamento bimodale solo quattro STRs: DYS438, DYS392, DYS456, e DYS635. Il locus DYS385a/b presenta 18 genotipi di cui 11-14 è il più frequente (0,405).

Nel campione del Sulcis hanno un andamento bimodale gli STRs: DYS438, DYS392, DYS456, DYS19, DYS458 e DYS635. Anche in questo campione è presente la duplicazione 15-16 nel locus DYS19 (0,041). Il locus DYS385a/b ha mostrato 18 genotipi, di cui i più frequenti sono 12-12 (0,229) e 12-13 (0,146).

Allele	DYS393	DYS391	DYS438	DYS439	DYS392	DYS389I	YGATAH4	DYS456	DYS19	DYS437	DYS458	DYS448	DDYS635	DYS390	DYS389II	DYS385a/b	
9	0,042	0,042	0,167													aplotipo	
10	0	0,854	0,708	0,021	0,021		0,021									10-15	0,021
11	0	0,104	0,083	0,521	0,875	0,042	0,458									11-13	0,145
12	0,208		0,042	0,375	0,021	0,062	0,271									11-14	0,021
13	0,604			0,062	0,083	0,792	0,208	0,021	0,104							11-15	0,021
14	0,146			0,021		0,104	0,042	0,479	0,166	0,187	0,021					12-12	0,167
15								0,333	0,271	0,625	0,083					12-13	0,187
16								0,063	0,167	0,188	0,188					12-14	0,021
17								0,083	0,271		0,333					12-15	0,021
18								0,021	0		0,125	0,021				12-17	0,021
18,2									0	0,042						13-14	0,042
19									0,021		0,187	0,083				13-15	0,042
20											0,021	0,583	0,042			13-17	0,021
21												0,292	0,437	0,042		13-18	0,041
22												0,021	0,312	0,125		14-16	0,021
23													0,104	0,562		14-18	0,021
24													0,042	0,25		15-16	0,104
25													0,063	0,021		16-17	0,021
26																16-18	0,041
27																17-20	0,021
28															0,479		
29															0,208		
30															0,208		
31															0,042		
32															0,063		

Tabella 15. Frequenze alleliche in 17 Y-STR nella popolazione di Benetutti.

Allele	DYS393	DYS391	DYS438	DYS439	DYS392	DYS389I	YGATAH4	DYS456	DYS19	DYS437	DYS458	DYS448	DDY5635	DYS390	DYS389II	DYS385a/b	
9	0,042	0,042	0,146													aplotipo	
10	0	0,854	0,563	0,083	0,021											10-13	0,021
11	0,021	0,104	0,083	0,48	0,75	0,042	0,417									10-14	0,021
12	0,146		0,146	0,333	0,021	0,146	0,437									10-15	0,021
13	0,604		0,062	0,104	0,166	0,708	0,125	0,042	0,021		0,021					10-16	0,021
14	0,166				0,021	0,104	0,021	0,417	0,354	0,167	0					11-12	0,041
15	0,021				0,021			0,187	0,25	0,666	0,062					11-14	0,104
16								0,271	0,167	0,167	0,208					11-15	0,062
17								0,062	0,208		0,458					12-12	0,145
18								0,021			0,167	0,125				12-13	0,125
18,2											0,021	0				12-14	0,042
19											0,042	0,167				12-15	0,042
20											0	0,5	0,083			12-16	0,021
21											0,021	0,187	0,312			12-17	0,021
22												0,021	0,167	0,146		13-13	0,021
23													0,292	0,521		13-14	0,021
24													0,125	0,25		13-15	0,083
25													0	0,083		13-18	0,021
26													0,021			14-14	0,021
27															0,042	14-16	0,062
28															0,354	14-19	0,021
29															0,375	17-18	0,021
30															0,104	17-19	0,021
31															0,083	18-19	0,021
32															0,042		

Tabella 16. Frequenze alleliche in 17 Y-STR relativi al nord Sardegna.

Allele	DYS393	DYS391	DYS438	DYS439	DYS392	DYS389I	YGATAH4	DYS456	DYS19	DYS437	DYS458	DYS448	DDY5635	DYS390	DYS389II	DYS385a/b	
9			0,071													aplotipo	
10		0,619	0,357	0,024												11-13	0,048
11		0,357	0,024	0,357	0,286		0,428									11-14	0,405
12	0,143	0,024	0,548	0,452	0,166	0,214	0,524									11-15	0,024
13	0,619			0,167	0,5	0,714	0,048	0,024	0,191							11-17	0,024
14	0,143				0,024	0,072		0,167	0,333	0,238						12-14	0,024
15	0,095				0,024			0,405	0,333	0,571	0,143					12-16	0,024
16								0,095	0,143	0,167	0,238					12-18	0,047
17								0,262		0,024	0,452					13-13	0,024
18								0,047			0,119	0,024				13-14	0,047
19											0,048	0,488				13-15	0,047
20												0,366				14-14	0,024
21												0,049	0,214			14-15	0,024
22												0,073	0,167	0,143		14-16	0,095
23													0,476	0,238		15-15	0,024
24													0,119	0,524		15-16	0,024
25													0,024	0,071		15-17	0,024
26														0,024		16-17	0,047
27																16-16	0,024
28															0,119		
29															0,619		
30															0,238		
31															0		
32															0,024		

Tabella 17. Frequenze alleliche di 17 Y-STR nella popolazione di Carloforte.

Allele	DYS393	DYS391	DYS438	DYS439	DYS392	DYS389I	YGATAH4	DYS456	DYS19	DYS437	DYS458	DYS448	DDYS635	DYS390	DYS389II	DYS385a/b			
9	0,021		0,146														aplotipo		
10	0	0,771	0,562	0,062			0,042										10-14	0,042	
11	0	0,229	0,063	0,417	0,708		0,5										11-11	0,021	
12	0,167		0,208	0,354	0,042	0,188	0,271										11-13	0,021	
13	0,687		0,021	0,167	0,208	0,708	0,187	0,167	0,062		0,042						11-14	0,083	
14	0,125				0,042	0,104		0,333	0,313	0,083	0,021						11-15	0,042	
15								0,25	0,208	0,688	0,146						11-16	0,021	
16								0,187	0,167	0,229	0,021						12-12	0,229	
17								0,042	0,25		0,416						12-13	0,146	
18								0,021			0,312	0,083					12-14	0,021	
19											0,021	0,188					12-15	0,042	
20											0,021	0,479	0,042				12-16	0,042	
21												0,188	0,292	0,062			12-18	0,021	
22												0,062	0,229	0,125			13-13	0,042	
23													0,333	0,542			13-14	0,041	
24													0,083	0,229			13-15	0,041	
25													0,021	0,042			13-16	0,041	
26																	15-16	0,021	
27																	0,042	16-16	0,083
28																	0,354		
29																	0,375		
30																	0,166		
31																	0,021		
32																	0,042		

Tabella 18. Frequenze alleliche di 17 Y-STR relativi alla popolazione del Sulcis iglesiente

Dall'analisi degli STRs attraverso il programma Arlequin (ver.3.1) sono stati calcolati vari parametri, i cui risultati sono riportati nella tabella 19, che permettono di attestare una diversità genetica fra le popolazioni analizzate e il rispettivo campione di controllo. Carloforte condivide con il suo campione di confronto un solo aplotipo, mentre, Benetutti non condivide alcun aplotipo con il campione del nord Sardegna. I valori di diversità genetica prossimi a 1 riscontrati nelle popolazioni analizzate evidenziano una grande variabilità genetica interna.

Sia il numero medio di differenze fra coppie di sequenze che la media della diversità genetica tra loci sono risultati maggiori in entrambi i campioni di controllo rispetto ai due isolati indicando una maggiore eterogeneità nei campioni di controllo.

La differenziazione delle popolazioni, calcolata tramite l'FST, è risultata statisticamente significativa (p-value <0,005), avendo ottenuto un valore pari a 0 sia tra Carloforte e il Sulcis, che tra Benetutti e il nord Sardegna.

	Benetutti	N.Sardegna	Carloforte	Sulcis
N° campioni	48	48	42	48
N° aplotipi	43/48	47/48	39/42	45/48
Diversità genica	0,975±0,001	0,980	0,975	0,978
Average gene diversity	0,593±0,301	0,633±0,320	0,603±0,306	0,623±0,315
Mean number pairwise	9,481±4,357	10,135±4,636	9,650±4,429	9,962±4,562

Tabella 19. Parametri rilevati con il programma Arlequin per le due popolazioni analizzate e i rispettivi campioni di controllo.

3.1.2 Analisi degli aplogruppi

Nei grafici riportati nelle figure 13 e 14 sono indicate le frequenze degli aplogruppi per le popolazioni di Benetutti e del nord Sardegna. Confrontando i due grafici si evince che l'aplogruppo maggiormente rappresentato in entrambe le popolazioni è l'aplogruppo I_{2a_1} (Benetutti 50% e Nord Sardegna 34%).

Benetutti mostra la quasi totalità dei suoi aplogruppi in comune con quelli del Nord Sardegna, pur con frequenze differenti: $E_1b_1b^*$ (Nord Sardegna 4% e Benetutti 13%), J_{2a_1} (Nord Sardegna 8% e Benetutti 11%), R_1b^* (Nord Sardegna 21% e Benetutti 4%), G_{2a^*} (Nord Sardegna 15% e Benetutti 10%)

Nelle figure 15 e 16 sono riportate le frequenze degli aplogruppi per Carloforte e il suo campione di confronto. A Carloforte l'aplogruppo più frequente è R_1b (55%), mentre nel Sulcis l' I_{2a_1} (38%). Questo risultato è importante poiché l'aplogruppo R_1b presenta valori simili nel nord Italia (Ferri et al.,2009; Semino et al.,2000) ed ha una frequenza del 25% nel Sulcis, del 4% a Benetutti e del 21% nel nord Sardegna.

Invece, l'aplogruppo I_{2a} che identifica la mutazione M26, molto frequente in Sardegna, è assente nel resto d'Italia (Onofri et al., 2007) e presenta una frequenza bassissima a Carloforte (2%).

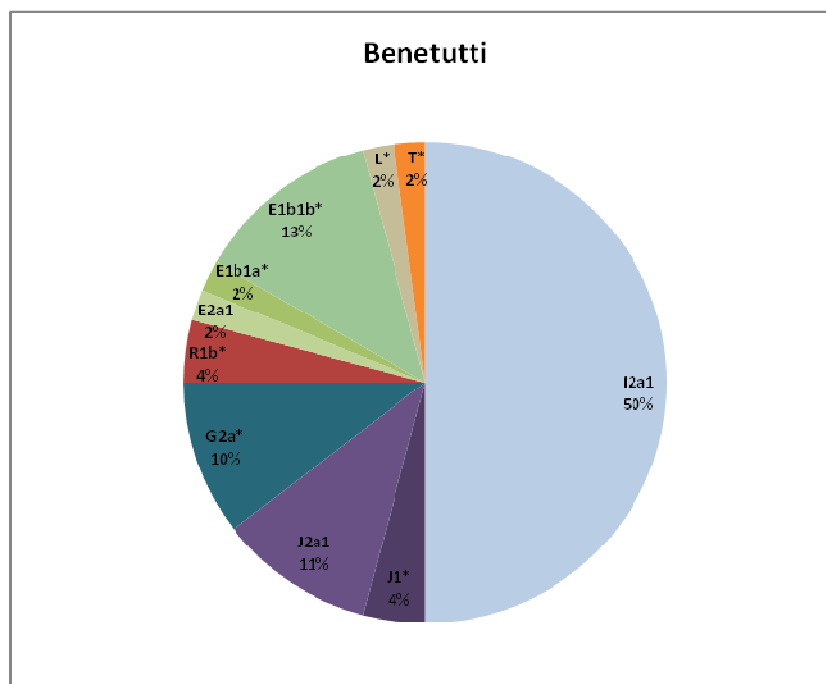


Figura 12. Frequenze degli aplogruppi per la popolazione di Benetutti

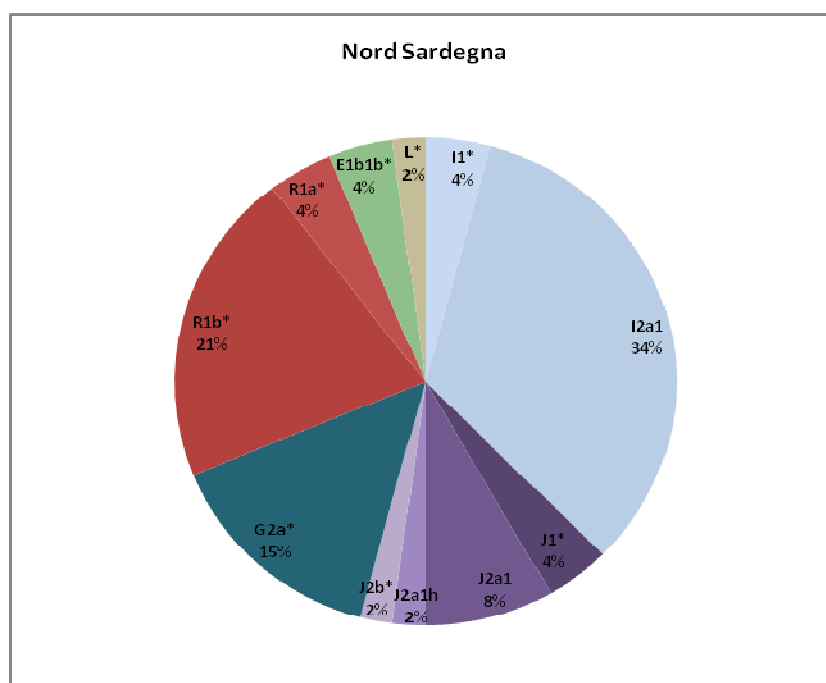


Figura 13. Frequenze degli aplogruppi per la popolazione del nord Sardegna

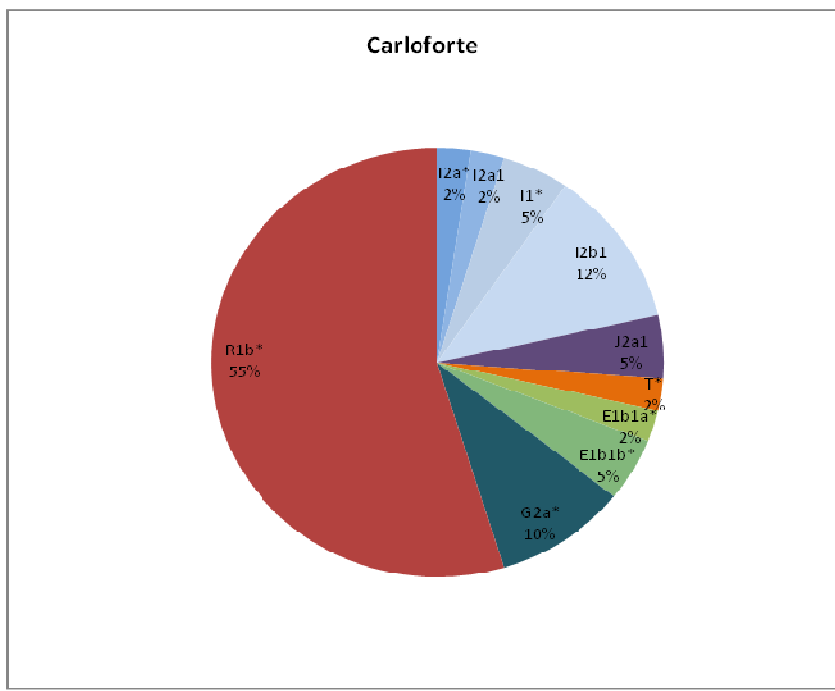


Figura 14. Frequenze degli aplogruppi per la popolazione di Carloforte

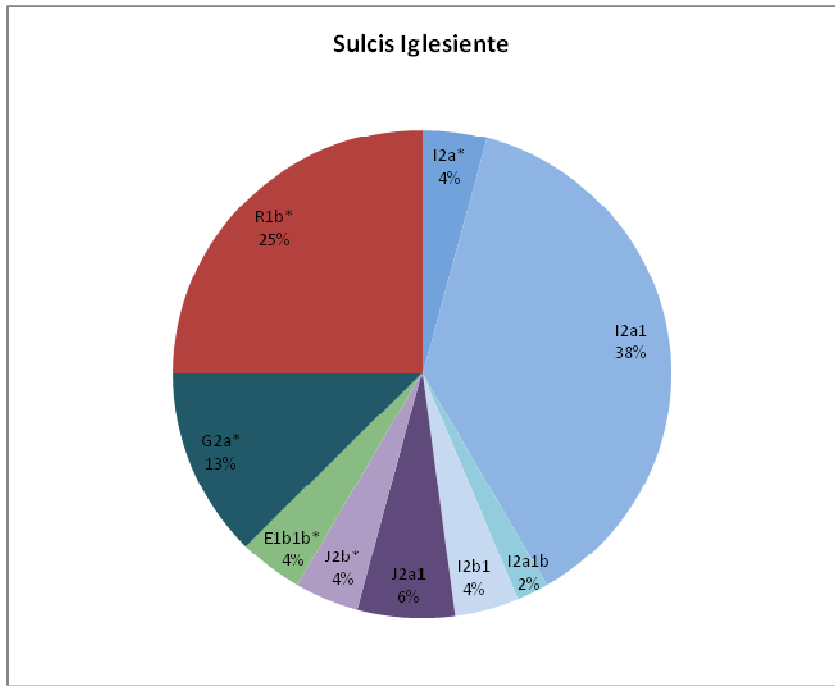


Figura 15. Frequenze degli aplogruppi per la popolazione del Sulcis Iglesiente

3.1.3 Network filogenetici

Per poter analizzare le relazioni filogenetiche degli aplotipi emersi dall'analisi degli STRs delle popolazioni è stato costruito un network che mostra la clusterizzazione degli aplotipi in base alla loro origine. Nel grafico i cerchi rappresentano gli aplotipi la cui area è direttamente proporzionale alla frequenza dell'aplotipo nella popolazione. Le mutazioni sono rappresentate con delle linee.

La forma del network filogenetico per la popolazione di Benetutti (Fig. 17) e di Carloforte (Fig. 19) mostra una eterogeneità della popolazione fondatrice, infatti, non vi è un aplotipo centrale più rappresentato degli altri, ma sono evidenti diversi nuclei (cluster aplotipici) da cui si dipartono i vari rami. Probabilmente tale risultato è dovuto alla strategia di campionamento che ha massimizzato la diversità genetica presente nei due isolati.

Successivamente si è andato a verificare quali aplogruppi presentassero i campioni che avevano clusterizzato e si è visto che vi è una corrispondenza fra cluster aplotipici e aplogruppi. A Benetutti vi è un cluster aplotipico principale che corrisponde all'aplogruppo I_{2a_1} ; a Carloforte, invece, il cluster principale presenta l'aplogruppo R_{1b} .

I network delle popolazioni di confronto mostrano una eterogeneità maggiore e un numero di rami superiore che si dipartono dai vari nuclei, pur presentando sempre una suddivisione per cluster principali (Fig. 18 e 20). Nel nord Sardegna vi è un cluster aplotipico maggiore cui corrisponde l'aplogruppo I_{2a_1} e altri due nuclei rilevanti che presentano rispettivamente gli aplogruppi R_{1b} e G_{2a} . Nel Sulcis Iglesiente al cluster aplotipico maggiore corrisponde l'aplogruppo I_{2a_1} e gli altri due nuclei di una certa rilevanza presentano l'aplogruppo R_{1b} e G_{2a} .

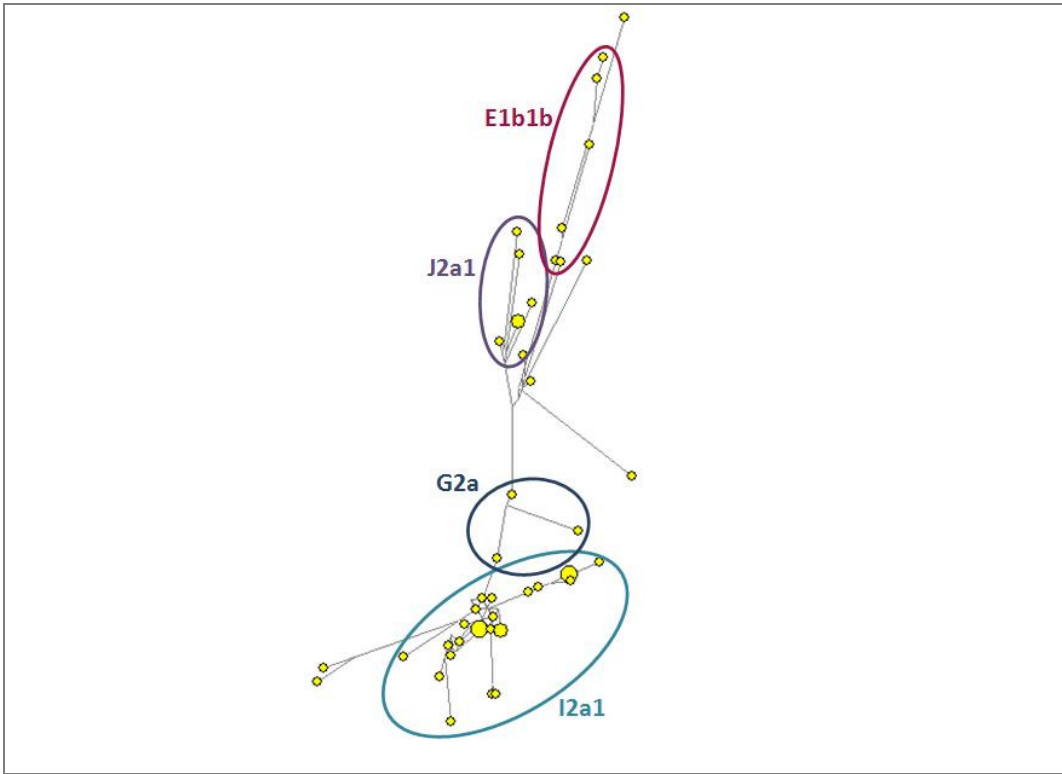


Figura 16. Network Y Benetutti

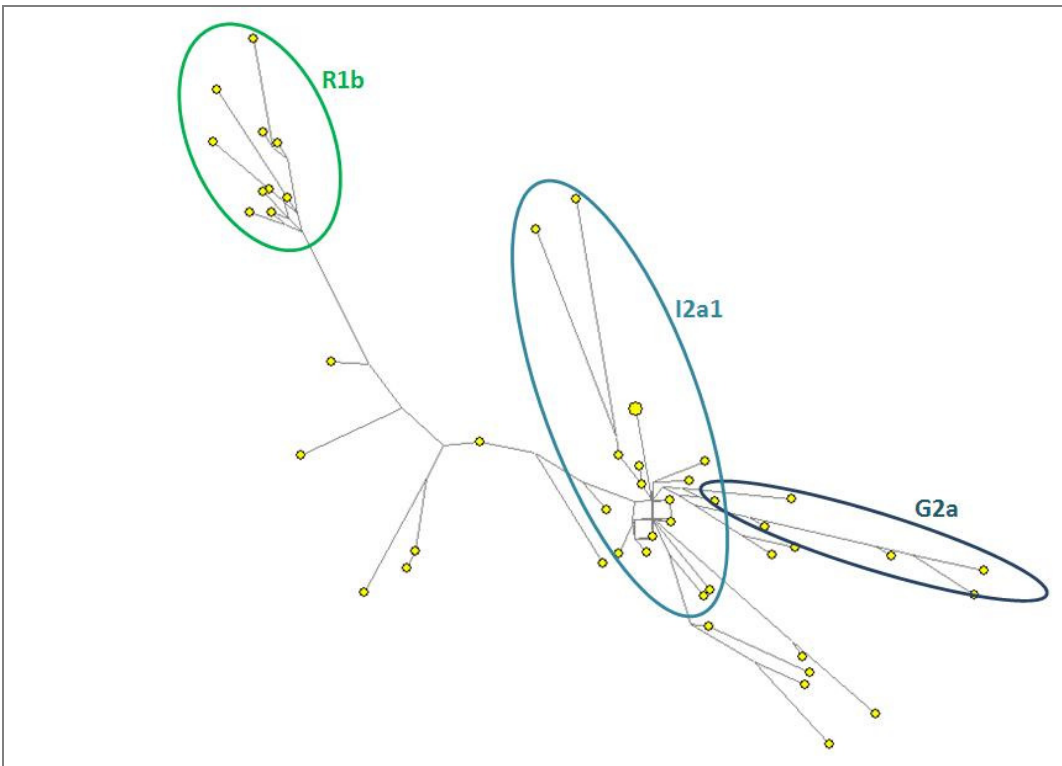


Figura 17. Network Y Nord Sardegna

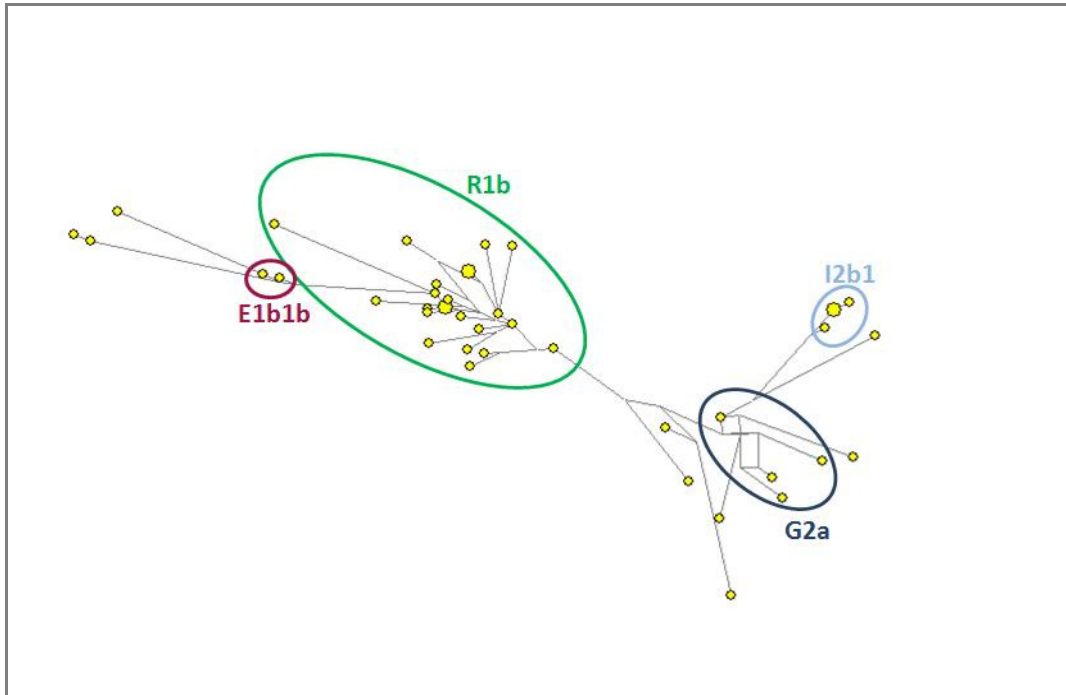


Figura 18. Network Y Carloforte

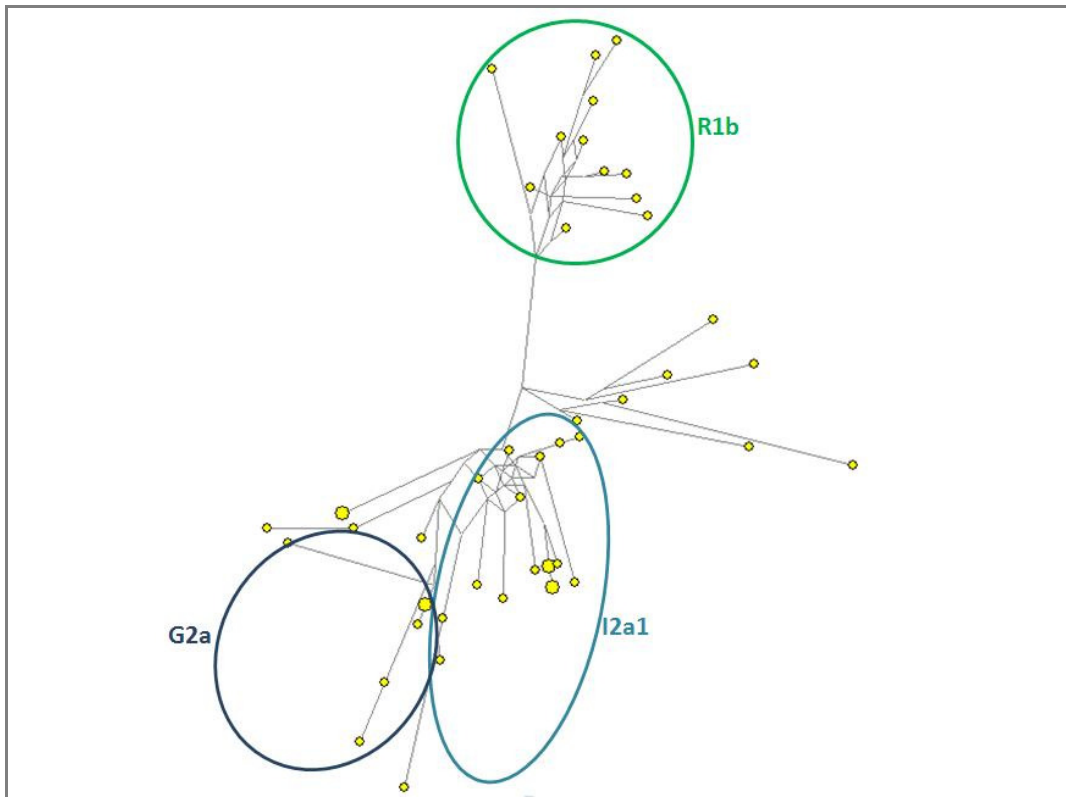


Figura 19. Network Y Sulcis Iglesiente

3.1.4 Distanze ed albero filogenetico

Sono stati effettuati dei confronti fra le nostre popolazioni e altre dell'area Mediterranea (nord-est Italia, centro Italia, Croazia, Marocco, Spagna, Tunisia e Grecia) mediante le frequenze di 12 STRs reperite in bibliografia (Turrina et al., 2006; Onofri et al., 2007; Pokupcic et al., 2008; Palet et al., 2010; Martin et al., 2004; Frigi et al., 2006; Kovatsi et al., 2009). La matrice delle distanze F_{ST} , elaborata con il programma Arlequin 3.1, è riportata in tabella 15.

Benetutti mostra le distanze maggiori con la Croazia (0,272), la Tunisia (0,226), la Spagna (0,184), Carloforte (0,138) e il Marocco (0,137) e le distanze minori con il nord Sardegna (0,023), il Sulcis Iglesiente (0,028), la Grecia (0,078) e il centro Italia (0,092).

Carloforte presenta le distanze maggiori con la Croazia (0,301), la Tunisia (0,264), il Marocco (0,178), e Benetutti (0,138) e le distanze minori con il nord-est Italia (0,018), il centro Italia (0,032) e la Spagna (0,034).

Dalla matrice delle distanze è stato costruito l'albero filogenetico rappresentato nella figura 25. Di tale ricostruzione è stata stimata la significatività o robustezza con il metodo bootstrap. Si può vedere dai valori di bootstrap riportati nei nodi dell'albero che tale ricostruzione è robusta e quindi significativa. Carloforte si colloca su un ramo separato come un outlier ed è più vicina alla Spagna e al nord Italia, invece, Benetutti ha clusterizzato con le altre popolazioni della Sardegna considerate. Inoltre, è interessante notare come le popolazioni sarde non abbiano clusterizzato con la popolazione italiana.

	N.E.Italia	C.Italia	Benetutti	N.Sardegna	Carloforte	Iglesiente	Croazia	Marocco	Spagna	Tunisia	Grecia
N.E.Italia	-										
C.Italia	0,019813	-									
Benetutti	0,118548	0,092407	-								
N.Sardegna	0,067091	0,052263	0,023309	-							
Carloforte	0,018505	0,03259	0,138546	0,084208	-						
Iglesiente	0,058875	0,048187	0,028697	0,007922	0,071825	-					
Croazia	0,249437	0,215418	0,27287	0,254902	0,301562	0,277481	-				
Marocco	0,145382	0,10471	0,137916	0,137037	0,178708	0,14583	0,207647	-			
Spagna	0,025209	0,039394	0,184535	0,124092	0,034235	0,111865	0,285428	0,169809	-		
Tunisia	0,230921	0,174122	0,226928	0,231307	0,264466	0,238817	0,280625	0,098117	0,245844	-	
Grecia	0,061648	0,038227	0,078284	0,067765	0,087223	0,069919	0,178608	0,068475	0,101769	0,126921	-

Tabella 15. Matrice delle distanze Fst per le popolazioni del Mediterraneo.

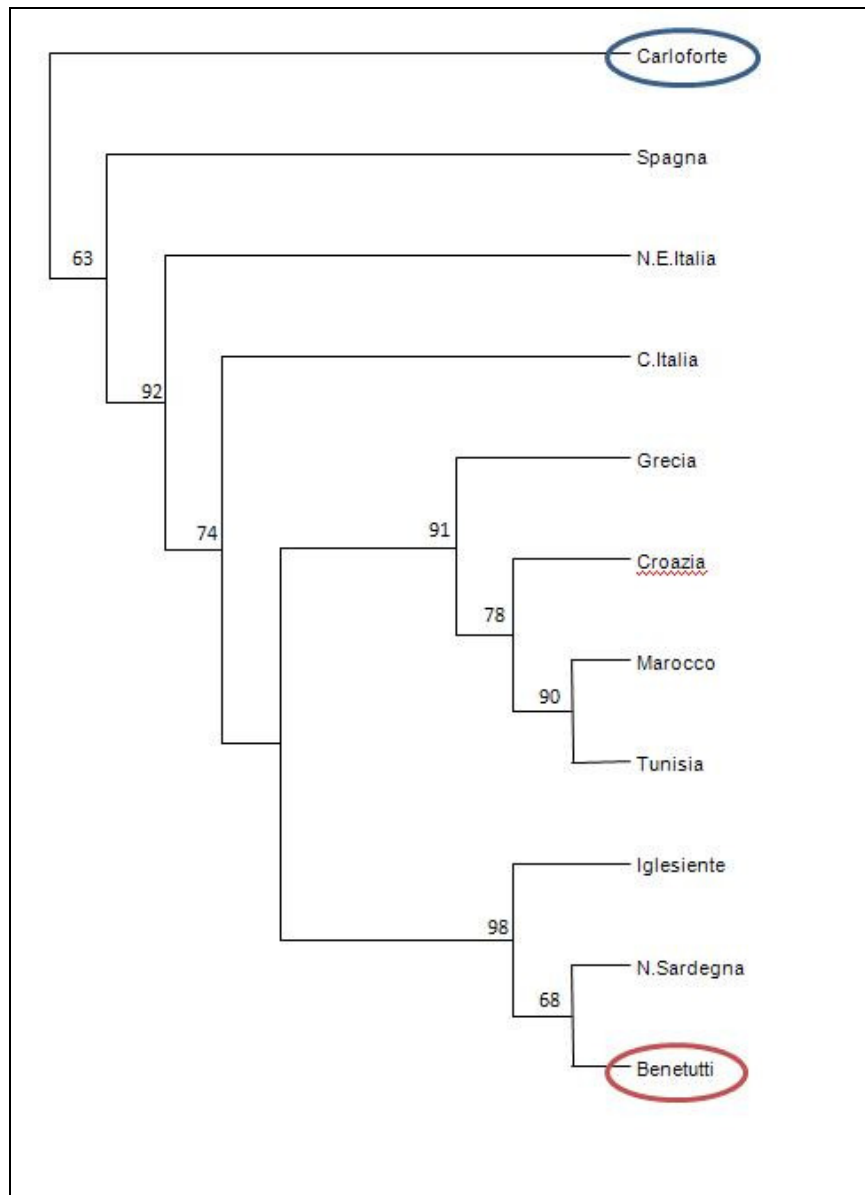


Figura 25. Albero del cromosoma Y per le popolazioni del mediterraneo

Per valutare ulteriormente la struttura genetica delle popolazioni è stato utilizzato il programma STRUCTURE . Si è usato un numero ipotetico di popolazioni (K) pari a 2, 3 e 4 , e per ciascuna di queste abbiamo ripetuto l'analisi quattro volte. La migliore interazione si è avuta per $K = 3$. poichè presenta i valori più bassi di $\ln P$. Nella figura 26 ogni individuo è indicato come una linea verticale ed è suddiviso in componenti K colorate (gruppi di popolazioni) che rappresentano l'appartenenza in cluster K. L'analisi della struttura della popolazione mostra la presenza di due cluster e la separazione delle nostre comunità oggetto di studio in due aree: nord e sud (Fig. 26), in accordo con la loro posizione geografica e la relativa vicinanza ($K = 2$). Il grafico ottenuto per $k= 3$ evidenzia la variabilità individuale all'interno della popolazione. Si può notare un basso grado di differenziazione fra i nostri isolati analizzati e il rispettivo campione di controllo, soprattutto per quanto riguarda Benetutti e il nord Sardegna.

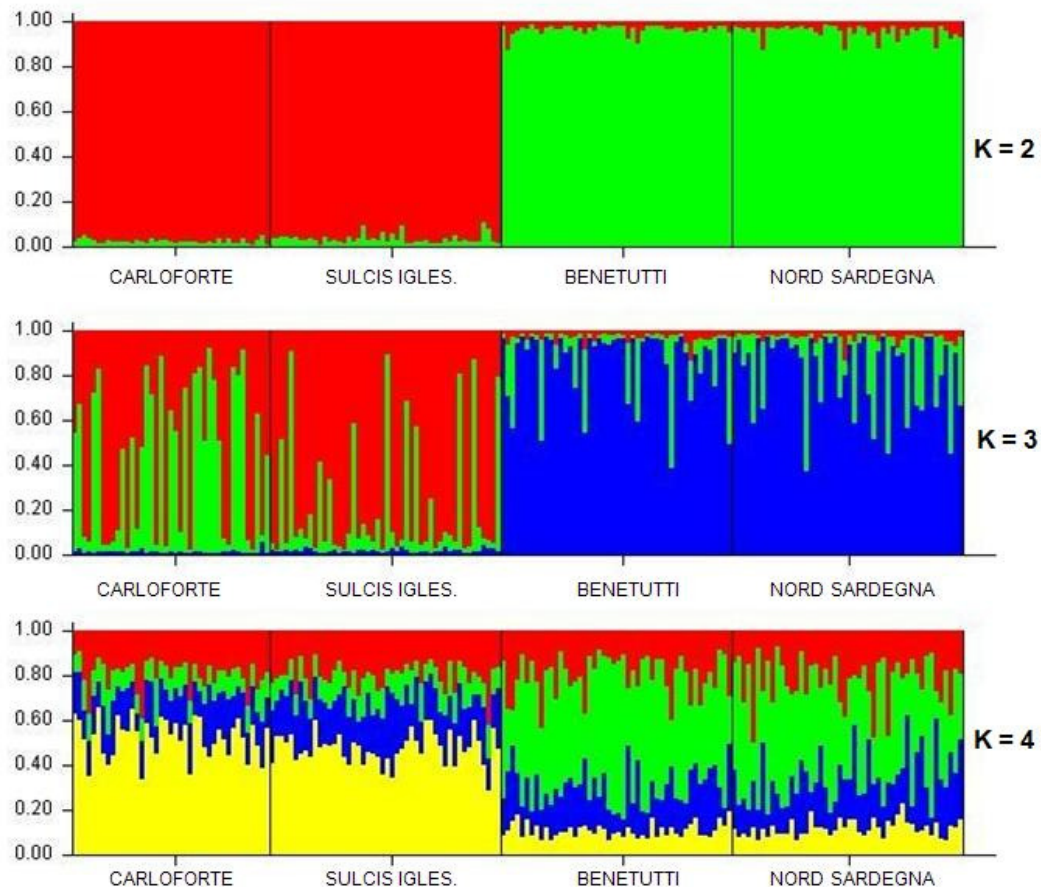


Figura 26. Risultati di STRUCTURE per il cromosoma Y con il presupposto di differenti gruppi di popolazioni (K=2, 3, 4). K=3 presenta il più alto valore di LnP(D)

3.2 Risultati DNA mitocondriale

3.2.1. Analisi delle sequenze della regione ipervariabile I e II.

Le sequenze del segmento ipervariabile I e II delle comunità prese in esame sono state confrontate con la sequenza di riferimento di Cambridge (CRS) e i risultati sono stati riportati nelle tabelle sottostanti (dalla 21 alla 28).

Il segmento HVRI è compreso fra la base 16024 e la 16365, mentre il segmento HVRII è compreso tra le posizioni 73 e 340. (Caramelli, 2009).

Per quanto riguarda la regione ipervariabile I sono stati analizzati 50 campioni di Benetutti, 41 del nord Sardegna, 49 di Carloforte e 50 del Sulcis iglesiente. Delle 50 sequenze analizzate di Benetutti, riportate nella tabella 21, 11 sono uguali a quella di Anderson.

Sono stati riscontrati 35 siti variabili per un totale di 128 mutazioni, di cui 113 sono transizioni (88,3%) e 15 transversioni (11,7%). Di tali transizioni l'84% sono pirimidiniche ($T \leftrightarrow C$). I siti con il più alto numero di mutazioni sono il 16311 e il 16356, entrambi con una frequenza nel campione del 22%, che mostrano lo stesso tipo di transizione $T \rightarrow C$. La transversione più frequente nel campione (18%) si trova nella posizione 16183 ($A \rightarrow C$).

Delle 41 sequenze analizzate della zona del nord Sardegna (tabella 22) solo una si presenta uguale alla sequenza di Anderson. Si sono rilevati 44 siti variabili che presentano un totale di 125 mutazioni, di cui 112 sono transizioni (89,6%) e 13 transversioni (10,4%). La transizione più frequente nel campione (39%) si trova nella posizione 16126 ($T \rightarrow C$), mentre la transversione più frequente (14,6%) è nella posizione 16033 ($G \rightarrow T$).

Fra i 49 campioni analizzati di Carloforte (Tab. 23) sono state trovate 10 sequenze uguali a quella di Anderson e un totale di 34 siti variabili. Le mutazioni rilevate sono 154, di cui 142 transizioni (92,2%) e 12 transversioni (7,8%). Il sito con il più alto numero di mutazioni (transizione $T \rightarrow C$) è il 16189 con una frequenza del 42,8% nella popolazione. La transversione più frequente (24,5%) si trova, invece, nella posizione 16183 ($A \rightarrow C$). Inoltre, 13 campioni presentano l'inserzione di una citosina dopo il sito 16193.

Nei 50 campioni della zona del Sulcis Iglesiente (tabella 24) sono state trovate 13 sequenze uguali a quella di Anderson, 44 siti variabili e un totale di 127 mutazioni, di cui 121 transizioni (95,3%) e 6 transversioni (4,7%). La mutazione più frequente nel campione (32%) è una transizione $T \rightarrow C$ nella posizione 16126. In generale, in tutte le quattro popolazioni oggetto di studio le transizioni più frequenti sono quelle pirimidiniche ($T \leftrightarrow C$), che si aggirano sempre intorno all' 80% delle mutazioni totali .

Per quanto riguarda il segmento ipervariabile II si hanno a disposizione 45 sequenze per Benetutti, 38 per il suo campione di controllo, 50 per Carloforte e per il rispettivo campione di confronto (Tab. 25, 26, 27, 28).

Le sequenze di Benetutti, riportate nella tabella 25, presentano 36 siti variabili per un totale di 292 mutazioni, di cui il 29,8% rappresentate da transizioni puriniche, il 36,6% da transizioni pirimidiniche e il 33,6% da trasversioni. Tutti i campioni analizzati presentano una transizione (A→G) nella posizione 263 e una trasversione (G→C) nella posizione 316.

Le sequenze del nord Sardegna, riportate nella tabella 26, presentano 24 siti variabili e un totale di 142 mutazioni. Il 66,2% di tali mutazioni è rappresentato da transizioni puriniche, il 32,4% da transizioni pirimidiniche e l'1,4% da trasversioni. Tutti i campioni presentano una transizione (A→G) nella posizione 263 e il 52,6% dei campioni presenta la mutazione A→G nella posizione 73.

Le sequenze di Carloforte, riportate nella tabella 27, presentano 17 siti variabili per un totale di 142 mutazioni. Tutti i campioni presentano la mutazione A→G nella posizione 263 e il 56% dei campioni hanno la transizione A→G nella posizione 73.

Le sequenze del Sulcis Iglesiente, riportate nella figura 28, presentano 24 siti variabili e un totale di 149 mutazioni che sono tutte transizioni. Le transizioni puriniche rappresentano il 63% del totale mentre quelle pirimidiniche il 37%. Tutti i campioni eccetto uno presentano la mutazione nella posizione 263 (A→G) e il 48% hanno la transizione A→G nel sito 73.

	16032	16033	16037	16038	16039	16040	16041	16042	16069	16074	16126	16153	16167	16183	16189	16192	16193	16206	16220	16223	16224	16256	16257	16266	16270	16274	16278	16290	16292	16294	16296	16298	16304	16311	16356			
Anderson	T	G	A	A	G	C	A	G	C	A	T	G	C	A	T	C	C	A	A	C	T	C	C	C	C	G	C	C	C	C	C	T	T	T	T			
002BEN	C	T	G		
004BEN	T	C	C	C	C	
005BEN	A	C	
009BEN	
013BEN	C	T	G		
017BEN	A	C	
018BEN	.	.	.	T	G	G	.	.	.	C	A	C	T	T	.	.	C		
021BEN	A	
023BEN	T	C	C	T	C	C	
024BEN	.	G	C	T	C	.		
025BEN	C	C	
027BEN	
031BEN	C	G	T		
035BEN	C	
036BEN	A	C	
038BEN	
041BEN	C	C	
044BEN	.	T	G	A	
045BEN	T	
046BEN	C	
078BEN	C	T	T	T	C	.	
079BEN	T	C	T	T	C	
081BEN	C	T	T	C	.	
084BEN	
085BEN	
087BEN	
088BEN	T	C	C	T	C	C	
089BEN	T	C	C	T	C	C	
090BEN	A	C	
092BEN	T	C	A	
094BEN	T	
095BEN	T	.	T	
096BEN	
097BEN	C	T	C	.	
100BEN	C	
101BEN	C	C	T	C	.	
102BEN	
103BEN	T	T	T	
105BEN	G	C	T	T	
106BEN	C	C	
111BEN	C	.
113BEN	T	C	C	T	C	C
114BEN	T	C	C	T	C	C
115BEN
116BEN	.	.	G	C	C	
118BEN	C	C	
121BEN
122BEN	A
123BEN	C	T	T	T	.	.	.	C	.	.	
125BEN

Tabella 21. Siti variabili nella regione HVRI della popolazione di Benetutti.

	16032	16033	16037	16038	16069	16093	16111	16114	16126	16129	16150	16153	16168	16169	16172	16182	16183	16189	16192	16193	16206	16209	16213	16223	16224	16235	16256	16257	16264	16266	16270	16274	16278	16290	16291	16292	16294	16295	16296	16298	16304	16311	16319	16362	
AND	T	G	A	A	C	T	C	C	T	G	C	G	C	C	T	A	A	T	C	C	A	T	G	C	T	A	C	C	C	C	G	C	C	C	C	C	C	C	T	T	T	G	T		
1 NS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2 NS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3 NS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4 NS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5 NS	-	-	-	T	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
6 NS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7 NS	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	T	-	T	-	C	-	-	-		
8 NS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	
9 NS	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	T	-	-	-	-	-	C		
10 NS	C	T	G	-	T	-	-	C	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
11 NS	-	-	-	-	T	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
12 NS	C	T	G	-	-	-	-	-	A	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C
13 NS	-	-	-	-	-	-	-	C	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A
14 NS	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	T	-	-	-	-	-	G	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
15 NS	C	T	G	-	-	-	-	-	-	-	-	-	-	-	-	C	C	C	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
16 NS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
17 NS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
18 NS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
19 NS	-	G	-	T	C	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
20 NS	C	T	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
21 NS	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
22 NS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
23 NS	-	-	-	-	T	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
24 NS	C	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
25 NS	-	-	-	-	T	-	G	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
26 NS	C	T	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
27 NS	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
28 NS	-	-	-	-	T	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
29 NS	-	-	-	-	T	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
30 NS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
31 NS	C	-	-	-	T	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
32 NS	-	-	-	-	T	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
33 NS	C	T	G	-	T	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
34 NS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
35 NS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
36 NS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
37 NS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
38 NS	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
39 NS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
40 NS	-	-	-	-	T	-	-	C	T	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
41 NS	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	

CAMPIONE	16069	16093	16126	16129	16145	16150	16163	16183	16184	16186	16189	16192	16193		16218	16223	16224	16231	16239	16249	16256	16261	16269	16270	16278	16288	16292	16294	16296	16304	16311	16320	16354	16356	16359			
ANDERSON	C	T	T	G	G	C	A	A	C	C	T	C	C	-	C	C	T	T	C	T	C	C	A	C	C	T	C	C	C	T	T	C	C	T	T			
1 CF	.	.	C	-	T	T	C	
2 CF	-	
3 CF	.	C	.	A	.	.	.	C	.	.	C	.	.	C	.	T	.	.	.	C	.	.	G	C	.	.	C	
4 CF	C	.	.	C	.	.	C	.	T	T	C	.	.	.	C
5 CF	.	C	.	A	.	.	.	C	.	.	C	.	.	C	.	T	.	.	.	C	C	.	.	.	C	
6 CF	C	.	.	C	.	.	C	.	T	T	
7 CF	-	
8 CF	.	C	C	.	.	.	G	.	.	T	C	.	.	-	T	
9 CF	T	.	-	T	.	.	T	
10 CF	.	C	.	A	.	.	.	C	.	.	C	.	.	C	.	T	.	.	.	C	C	.	.	.	C		
11 CF	-
12 CF	.	C	C	.	.	.	G	.	.	T	C	.	.	-	T
13 CF	C	.	.	C	.	.	C	.	T	T
14 CF	-	T
15 CF	C	.	.	C	.	.	C	.	T	T
16 CF	.	C	C	.	.	.	G	.	.	T	C	.	.	-	T
17 CF	-
18 CF	C	.	.	.	-	C	.
19 CF	-
20 CF	C	.	.	-	C	.
21 CF	C	.	.	C	.	.	C	.	T	T
22 CF	-
23 CF	.	.	C	-	C	T	T	T	C
24 CF	T	.	.	-	T	T	.	.	T	.	.	T	T	T	T	C	
25 CF	C	.	.	C	.	.	C	.	T	T
26 CF	C	.	.	C	.	.	C	.	T	T
27 CF	T	-	T	.	.	.
28 CF	-
29 CF	-
30 CF	.	C	.	A	.	.	.	C	.	.	C	.	.	C	.	T	.	.	.	C	C	.	.	.	C	
31 CF	T	.	C	T	.	-	T	C
33 CF	-	T	T	.	.	T
34 CF	T	.	.	-	T	T	.	.	T
35 CF	T	.	C	.	A	-	.	.	.	C	.	.	.	T
36 CF	C	.	.	C	.	.	C	.	T	T
37 CF	C	.	.	C	C	.
38 CF	C	.	.	-	C	.
39 CF	-
40 CF	T	.	C	-
41 CF	.	.	C	.	.	.	G	.	T	T	C	.	.	-	T
42 CF	-	.	.	C	C	T	
43 CF	-	.	.	.	T
44 CF	-
45 CF	-	T
46 CF	.	.	C	-	T	T	C
47 CF	.	C	C	.	.	.	G	.	.	T	C	.	.	-	T
48 CF	-	T
49 CF	-	T
50 CF	-	T

Tabella 23. Siti variabili nella regione HVRI della popolazione di Carloforte.

	73	93	146	150	151	152	153	185	189	194	195	199	204	207	228	263	272	295	310	311	312	316	317	319	320	322	324	326	327	328	329	330	332	335	338	339			
AND	A	A	T	C	C	T	A	G	A	C	T	T	T	G	G	A	A	C	T	C	C	G	C	T	C	G	C	A	C	A	G	C	C	A	C	A			
009 BEN	C	G	.	.	C	T	.	C		
013 BEN	C	.	.	.	G	.	.	C	.	.	C		
017 BEN	G	C	.	C		
018 BEN	G	.	C	G	G	.	C	-	T	C			
021 BEN	G	.	.	C	-	T	C	.	.	T			
023 BEN	G	.	.	T	G	.	.	-	T	C	C			
024 BEN	G	G	G	.	C	.	T	C	.	.	T			
025 BEN	G	.	C	G	.	.	-	T	C		
027 BEN	G	.	.	C	-	C	.	.	T	T	G	C	A	C	A	G	A	T	A	C				
031 BEN	.	G	G	.	.	-	T	C		
035 BEN	G	.	.	.	C	.	A	G	.	C	.	C	.	.	.	G	C	G		
036 BEN	G	C	G		
038 BEN	G	G	C	G		
041 BEN	G	.	.	C	-	C	.	.	T	.	G	C	.	.	A	G	.	T	.	.	.			
044 BEN	G	.	.	C	-	C	.	.	T	.	G	C	.	.	.	A	G	.	T	.	.			
045 BEN	G	.	.	C	.	T	C	.	.	T		
046 BEN	C	.	.	.	G	.	.	-	T	C		
079 BEN	G	.	.	T	.	C	G	.	T	C	T	.	C		
081 BEN	G	G	G	.	C	.	T	C	.	.	T		
084 BEN	T	G	.	.	-	-	T	C		
085 BEN	G	C	-	C	-		
087 BEN	.	G	G	.	.	.	-	T	C		
088 BEN	G	.	.	T	G	.	.	-	T	C		
089 BEN	G	.	.	T	G	C	G		
090 BEN	G	.	.	-	T	C		
092 BEN	G	.	.	.	C	.	A	A	G	.	T	-	T	.	C		
094 BEN	G	.	.	-	-	T	C	
095 BEN	G	G	.	.	C	-	T	C	
096 BEN	C	.	.	.	G	.	.	C	-	C	.	.	T	.	G	C	T		
097 BEN	G	G	.	.	C	-	T	C	.	.	T	
101 BEN	G	.	.	T	G	.	.	C	T	.	C	
102 BEN	G	.	.	-	-	T	C	
103 BEN	G	G	T	C	.	C	A	.	G	.	.	C	T	.	C	.	.	T	
105 BEN	G	.	.	T	G	.	.	-	-	T	C	T	
106 BEN	C	C	.	.	.	G	.	.	C	-	C	.	.	T	.	G	T	.	.	
111 BEN	C	G	.	.	C	-	C	.	.	T	.	G	C	.	.	A	G	A	T	A	.	.	.		
113 BEN	G	.	.	T	G	.	.	-	-	T	C	
114 BEN	G	.	.	T	G	.	.	-	-	T	C	
115 BEN	C	G	.	.	C	-	T	C	.	.	T	
116 BEN	G	.	.	C	.	T	C	.	.	T	
118 BEN	G	.	C	.	.	C	.	A	G	.	.	-	-	T	C	
121 BEN	C	.	.	.	G	.	.	C	-	C	.	.	T	
122 BEN	G	.	.	C	-	C	.	.	T	T	G	C	.	C	A	G	A	T	A	C	.	.		
123 BEN	G	.	.	.	T	C	G	.	.	C	-	T	C
125 BEN	G	.	.	C	-	C	.	.	T	.	G	C	.	.	A	G	A	T	A	C	.	.	.	

Tabella 25. Siti variabili nella regione HVRII della popolazione di Benetutti.

CAMPIONE	73	93	146	150	151	152	183	185	188	189	195	199	204	214	228	235	250	263	272	295	310	320	324	339
Anderson	A	A	T	C	C	T	A	G	A	A	T	T	T	A	G	A	T	A	A	C	T	C	C	A
1 NS	.	G	G	.	.	C	.	.	.
2 NS	G	.	.	.	C	T	.
3 NS	G	.	.	.	C	T	.
4 NS	G
5 NS	G	A	A	.	.	G	.	T	C	.	.	.
6 NS	G
7 NS	G	.	.	.	T	C	G
8 NS	G	.	.	.	T	.	.
9 NS	G	.	.	.	T	C	G
10 NS	G	A	G	A	.	.	G	.	T
11 NS	G	A	A	.	.	G	.	T	C	T	G	C
12 NS	G	.	.	C	T	.	.
13 NS	G
14 NS	G	.	.	T	G
16 NS	G
17 NS	C	G
18 NS	.	G	G
19 NS	G	A	A	.	.	G	.	T
20 NS	G	.	.	.	T	.	.
21 NS	G	.	C	G	.	G	.	G
22 NS	G	G
23 NS	G	A	A	.	.	G	.	T	C	T	.	.
24 NS	G
25 NS	G	A	A	.	.	G	G	T
26 NS	C	.	C	G	.	.	.	T	.	.
27 NS	G
28 NS	G	A	.	G	A	.	.	G	G	T
29 NS	G	A	.	G	A	.	.	G	G	T
30 NS	G	C	C	C	G
31 NS	G	A	A	.	.	G	G	T
33 NS	G	A	G	A	.	.	G	.	T
34 NS	.	.	C	T	.	.	G	G	.	.	.	T	.	.
35 NS	G
36 NS	G	.	C	G
38 NS	G	.	.	T	G	.	.	.	T	.	.
39 NS	G
40 NS	G	.	C	.	.	C	.	A	A	.	.	G	.	T
41 NS	G	G	G

Tabella 26. Siti variabili nella regione HVRII della popolazione del nord Sardegna.

CAMPIONE	73	114	146	150	151	152	153	185	195	200	215	225	226	228	263	295				323
AND	A	C	T	C	C	T	A	G	T	A	A	G	T	G	A	C	.	.	.	T
1 CF	G	G	.	C	.	.	.
2 CF	.	.	C	.	.	C	G	.	C	.	.	.
3 CF	G	G	.	C	C	C	.
4 CF	G	G	.	C	.	.	A	C	.	G	.	C	.	C	.
5 CF	G	G	.	C	C	C	.
6 CF	G	G	.	C	.	.	A	C	.	G	.	C	C	C	.
7 CF	G	.	C	C	C	.
8 CF	G	.	.	T	C	G	.	.	.	C	.
9 CF	G	A	C	G	.	T	.	C	.
10 CF	G	G	.	C	C	C	.
11 CF	G	.	C	C	C	.
12 CF	G	.	.	T	C	G	.	C	.	C	.
13 CF	G	G	.	C	.	.	A	C	.	G	.	C	.	C	.
14 CF	G	G	.	C	.	C	.
15 CF	G	G	.	C	.	.	A	C	.	G	.	C	.	C	.
16 CF	G	.	.	T	C	G	.	C	.	C	.
17 CF	G	.	C	.	C	.
18 CF	G	.	C	.	C	.
19 CF	G	.	C	C	C	.
20 CF	G	.	C	.	C	.
21 CF	G	G	.	C	.	.	A	C	.	G	.	C	.	C	.
22 CF	G	.	.	.	C	.
23 CF	G	.	.	.	T	C	G	.	C	.	C	.
24 CF	G	G	.	.	.	C	.
25 CF	G	G	.	C	.	.	A	C	.	G	.	C	.	C	.
26 CF	G	G	.	C	.	.	A	C	.	G	.	C	.	C	.
27 CF	G	.	C	C	C	.
28 CF	G	.	C	.	C	.
29 CF	G	.	C	.	C	.
30 CF	G	G	.	C	C	C	.
31 CF	G	.	.	T	.	C	G	T	C	C	C	.
32 CF	G	.	C	C	C	.
33 CF	G	G	.	.	.	C	.
34 CF	G	G	.	.	.	C	.
35 CF	G	.	.	T	.	C	.	.	C	.	G	.	.	.	G	T	.	.	C	C
36 CF	G	G	.	C	.	.	A	C	.	G	.	C	.	C	.
37 CF	G	.	C	C	C	.
38 CF	G	.	C	.	C	.
39 CF	G	.	C	.	C	.
40 CF	G	A	A	G	T	C	.	C	.
41 CF	G	T	G	.	C	.	C	.
42 CF	G	C	G	.	T	.	C	.
43 CF	G	.	C	.	C	.
44 CF	G	.	C	.	C	.
45 CF	G	G	.	C	.	C	.
46 CF	G	C	G	.	C	.	C	.
47 CF	G	.	.	T	C	G	.	C	.	C	.
48 CF	G	G	.	C	.	C	.
49 CF	G	.	C	C	C	.
50 CF	G	.	C	C	C	.

Tabella 27. Siti variabili nella regione HVRII della popolazione di Carloforte.

CAMPIONE	73	93	103	114	143	146	150	151	152	153	185	195		199	203	204	215	225	226	228	250	263	279	295				319			
ANDERSON	A	A	G	C	G	T	C	C	T	A	G	T		T	G	T	A	G	T	G	T	A	T	C				T			
1 SULC	.	G	G		
2 SULC	C	G	.	.	C	.	.	.		
3 SULC	C	G	.	.	C	.	.	.		
4 SULC	G	.	.	C	.	.	.		
5 SULC	G	.	.	C	C	C	.		
6 SULC	G	T	G	C	.		
7 SULC	G	.	A	G	C	.		
8 SULC	G	.	.	C	C	C	.		
9 SULC	G	A	A	.	G	.	T	C	.	.	C	.		
10 SULC	.	G	G	C	.	
11 SULC	C	G	.	.	C	C	C	.		
12 SULC	G	T	C	.	.	T	G	.	.	C	.	C	.		
13 SULC	G	C	.	
14 SULC	G	.	T	G	.	.	C	.	C	.		
15 SULC	G	C	A	C	C	G	.	.	C	.	C	.		
16 SULC	G	T	.	C	.	.	C	G	G	.	T	.	.	C	C	.		
17 SULC	G	C	.	
18 SULC	G	T	G	C	.		
19 SULC	.	G	G	C	.	
20 SULC	G	T	.	C	G	.	T	C	.	C	.		
21 SULC	G	T	.	C	G	.	T	C	.	C	.		
22 SULC	G	G	C	.	
23 SULC	G	T	.	C	.	.	C	G	G	.	T	C	.	C	C	.		
24 SULC	G	G	C	A	C	.	.	G	.	.	C	.	.	C	.		
25 SULC	G	.	.	C	C	C	.		
26 SULC	G	T	.	C	A	.	G	.	T	C	.	C	.			
27 SULC	G	C	T	.	C	C	A	.	G	.	T	C	.	C	.		
28 SULC	G	T	.	C	C	G	.	T	C	.	C	.			
29 SULC	G	A	A	.	G	.	T	.	.	.	C	.		
30 SULC	.	G	G	C	.	
31 SULC	G	C	.	
32 SULC	G	C	.	
33 SULC	G	.	.	C	C	.	.	.	
34 SULC	C	G	C	.	
35 SULC	G	G	C	.	
36 SULC	G	T	C	G	.	.	C	C	C	.		
37 SULC	.	G	G	.	.	C	.	C	.		
38 SULC	C	G	C	.	
39 SULC	G	C	.
40 SULC	G	T	G	C	.	
41 SULC	G	C	.	
42 SULC	.	G	G	.	.	C	.	C	.		
43 SULC	G	C	.	C	G	C	C	.	
44 SULC	G	A	.	G	.	T	.	.	C	.		
45 SULC	.	G	G	.	.	C	.	C	.		
46 SULC	G	T	G	C	.	
47 SULC	G	T	C	T	G	.	.	C	.	C	.		
48 SULC	A	G	C	.	
49 SULC	G	.	.	C	C	C	.		
50 SULC	G	.	.	C	.	C	.		

Tabella 28. Siti variabili nella regione HVRII della popolazione del Sulcis Iglesiente.

Mediante l’analisi delle sequenze delle popolazioni oggetto di studio, tramite il programma Arlequin (ver.3.1), si sono analizzati alcuni parametri di indagine intra-popolazionistica. Fra le popolazioni oggetto di studio è stato rilevato un alto livello di diversità genetica, come si può notare dai dati riportati nella tabella 29.

In tutte le popolazioni si è riscontrato un valore di diversità genetica pari a 1, che significa che tutti gli individui mostrano sequenze differenti. La diversità genetica risulta superiore alla diversità genetica media della popolazione sarda (0.930) (Simoni et al., 2000). Sia il numero medio di differenze fra coppie di sequenze che la media della diversità genica tra loci sono risultati maggiori nel campione del nord Sardegna rispetto a Benetutti, indicando una maggiore eterogeneità nel campione di controllo. Tali valori sono invece risultati più alti a Carloforte rispetto al Sulcis Iglesiente.

La differenziazione delle popolazioni, calcolata tramite l’FST, è risultata statisticamente significativa (p-value <0,005), avendo ottenuto un valore pari a 0,009 tra Carloforte e il Sulcis e di 0 fra Benetutti e il suo campione di confronto. Questo dato è confermato dall’assenza di aplotipi uguali condivisi fra le popolazioni.

	Benetutti	nord Sard	Carloforte	Sulcis Igles.
n° campioni	50	41	49	50
n° aplotipi diversi	50/50	41/41	49/49	50/50
n° polymorfic loci	128	125	154	127
gene diversity	1 ± 0,004	1 ± 0,005	1 ± 0,0041	1 ± 0,004
mean number of pairwise differences	7,083 ± 3,376	14,182 ± 7,473	5,257 ± 2,585	4,580 ± 2,290
average gene diversity over loci	0,025 ± 0,013	0,031 ± 0,016	0,015 ± 0,008	0,013 ± 0,007
transizioni	113	112	142	121
transversioni	15	13	12	6

Tabella 29. Parametri rilevati dalle sequenze per i campioni analizzati.

3.2.2 Analisi degli aplogruppi

I risultati della determinazione degli aplogruppi sono riassunti nei grafici a torta, riportati nelle figure 27, 28, 29, 30, che sono stati realizzati, per un più facile confronto, accorpando i sottoaplogruppi negli aplogruppi principali. Per vedere la variabilità riscontrata nelle comunità oggetto di studio a livello di sottoaplogruppi si possono, invece, guardare i grafici 31, 32, 33 e 34 .

A Benetutti gli aplogruppi più frequenti sono H (63%) e T (15%), nel campione di confronto sono H (35%), J (30%) e HV (10%).

Si può notare che gli aplogruppi riscontrati sono praticamente gli stessi con frequenze differenti, ma nel campione del nord Sardegna vi è un maggiore numero di aplogruppi diversi. Fra gli aplogruppi non riscontrati a Benetutti vi sono I (3%), A (3%) e R0 (2%). L'unica eccezione è rappresentata dall'aplogruppo W che è presente solo a Benetutti con una frequenza del 2%.

A Carloforte gli aplogruppi più frequenti sono H (33%), U (19%), X (16%) e T (16%), mentre, nel Sulcis Iglesiente sono H (40%), J (20%), U (12%), HV (10%) e T (10%). La popolazione di Carloforte presenta l'aplogruppo M1a1 (8%) non riscontrato nel campione di controllo e nelle altre popolazioni analizzate. Inoltre, presenta un'alta frequenza dell'aplogruppo X₂b (16%), presente - con bassissima frequenza - nella popolazione del Sulcis (2%). Nel campione di controllo è stata riscontrata una maggiore variabilità per numero di aplogruppi diversi riscontrati. Questi sono: HV (10%), I (2%) ed R0 (2%).

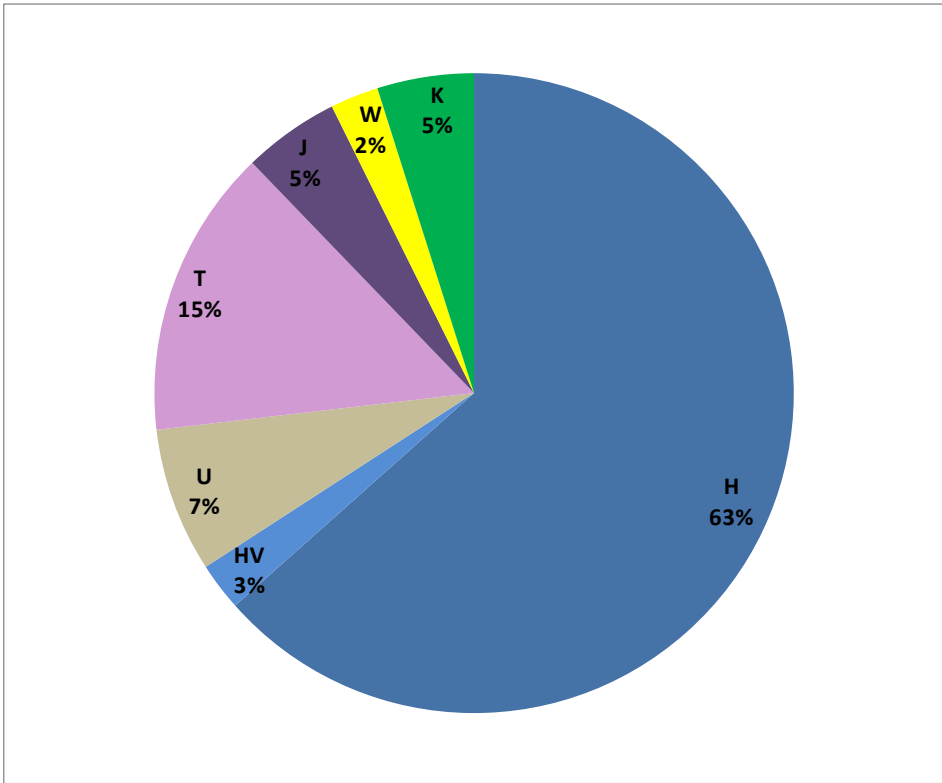


Figura 27. Aplogruppi per Benetutti

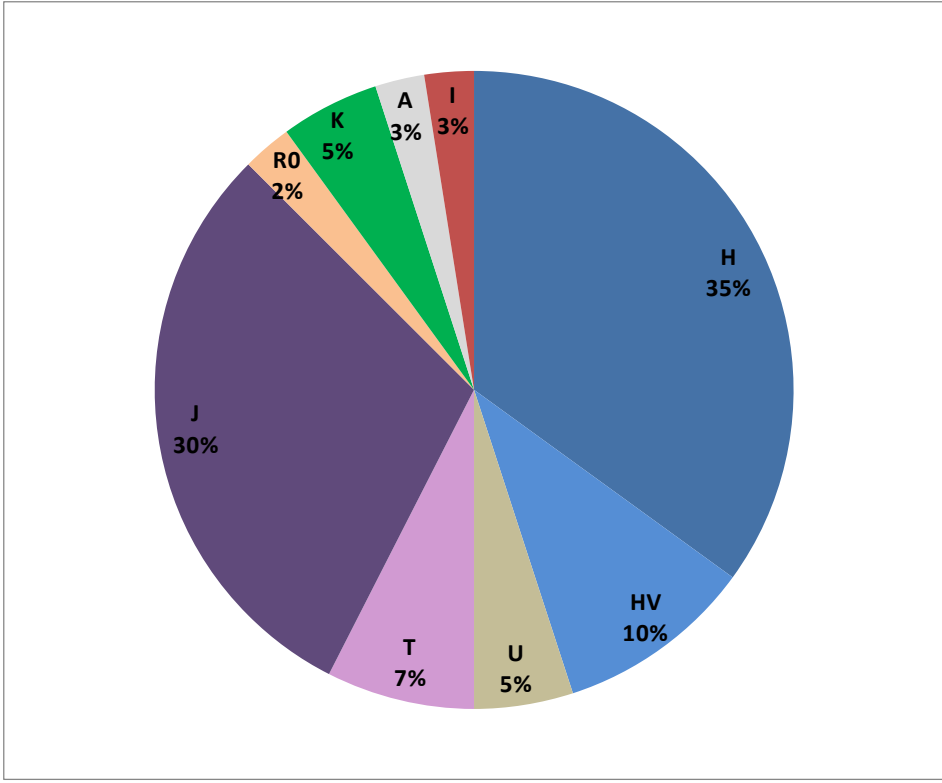


Figura 28. Aplogruppi del nord Sardegna

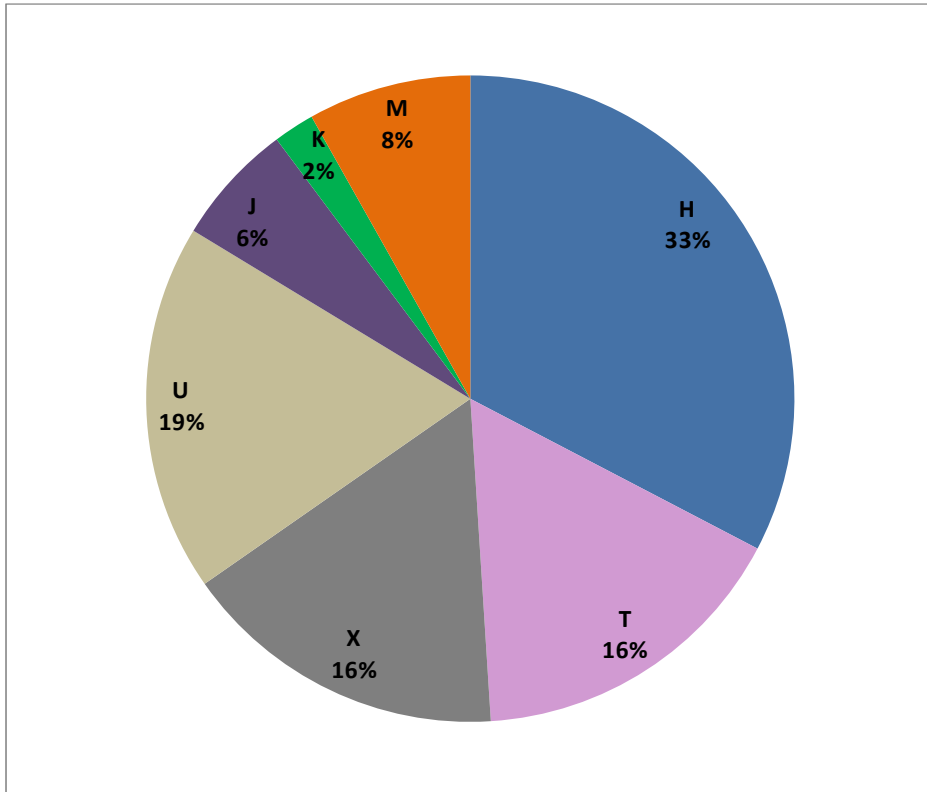


Figura 29. Aplogruppi per Carloforte

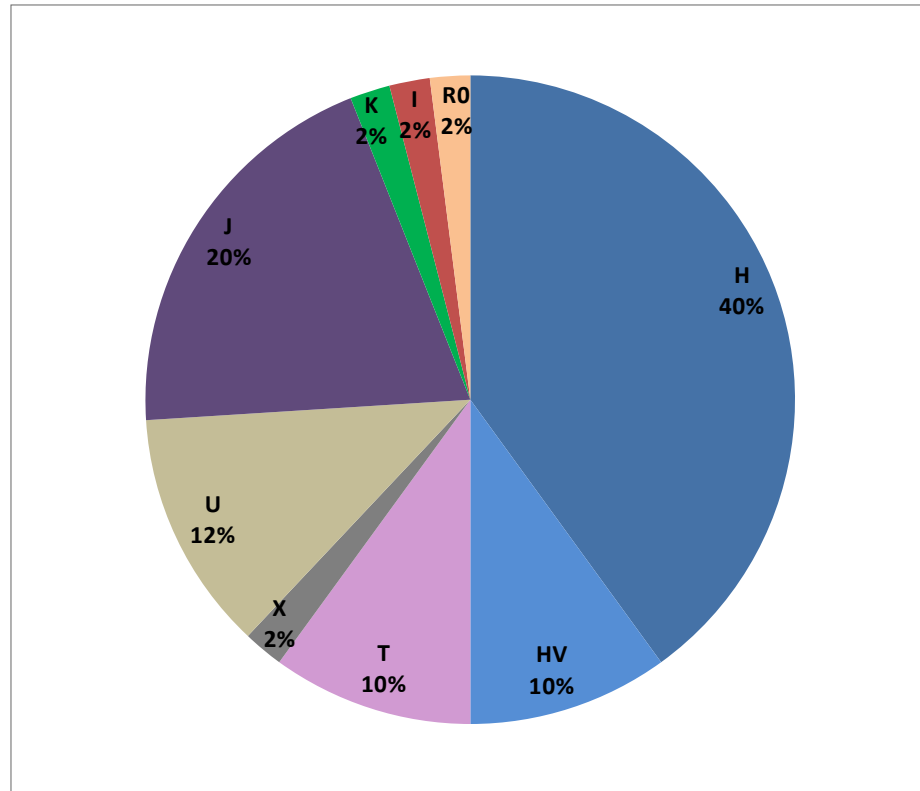


Figura 30. Aplogruppi mtDNA Sulcis Iglesiente

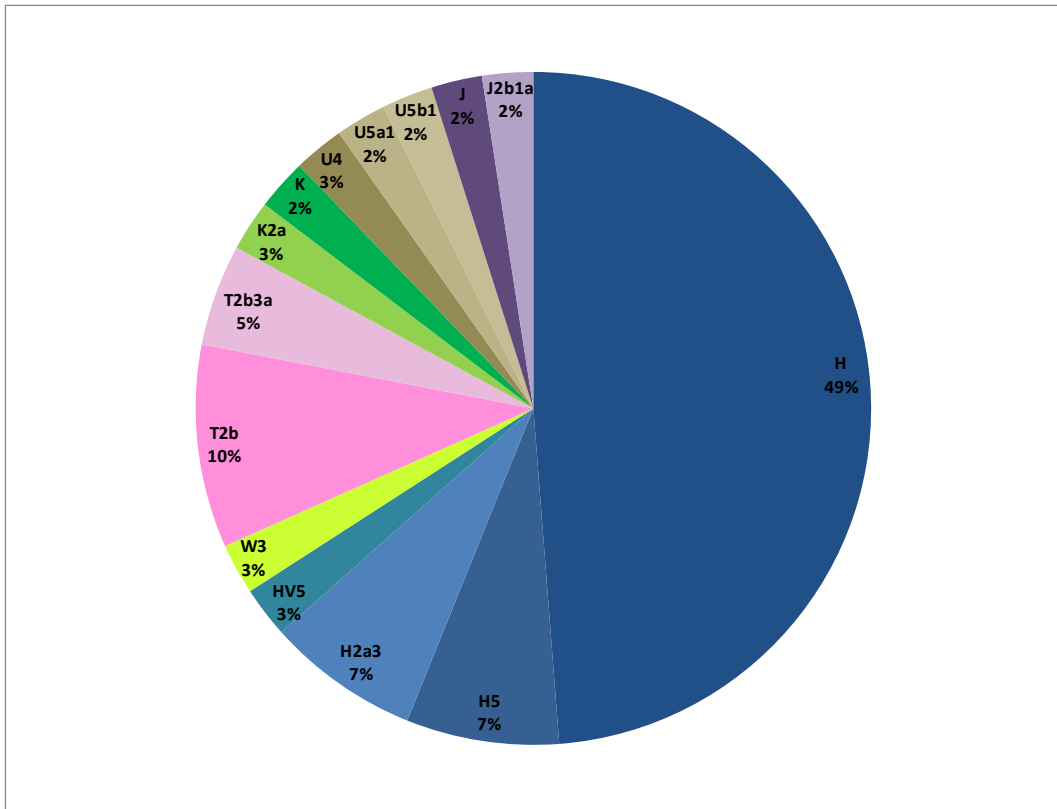


Figura 31 Aplogruppi dettagliati mtDNA Benetutti

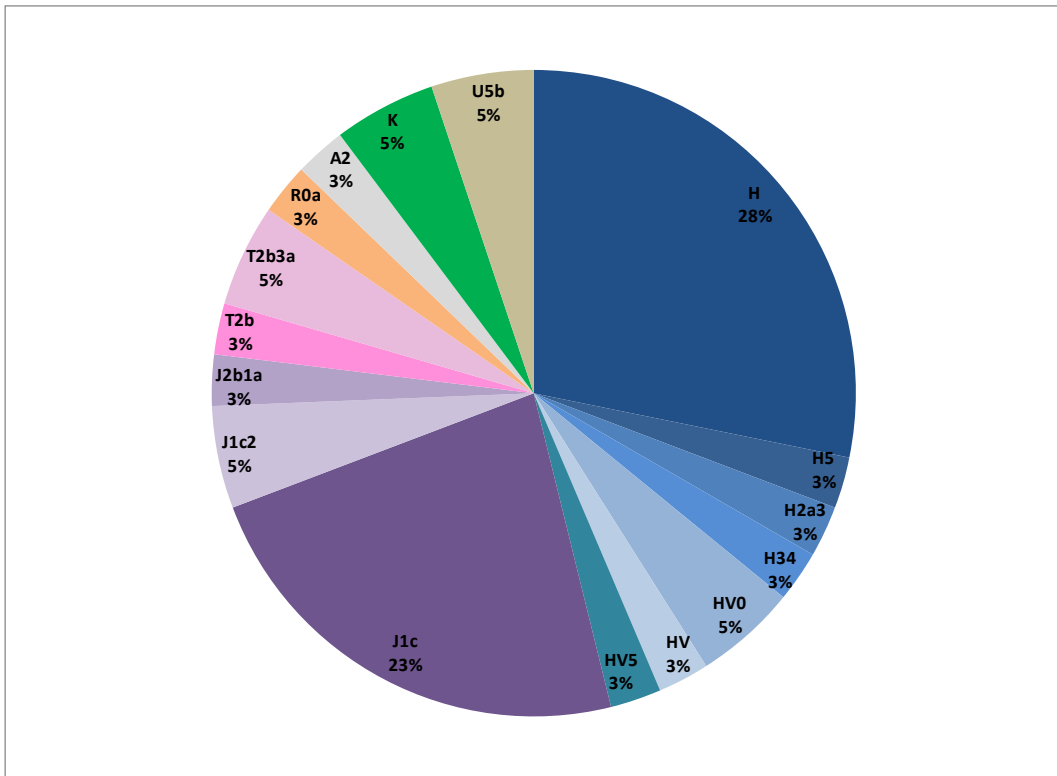


Figura 32. Aplogruppi dettagliati mtDNA nord Sardegna

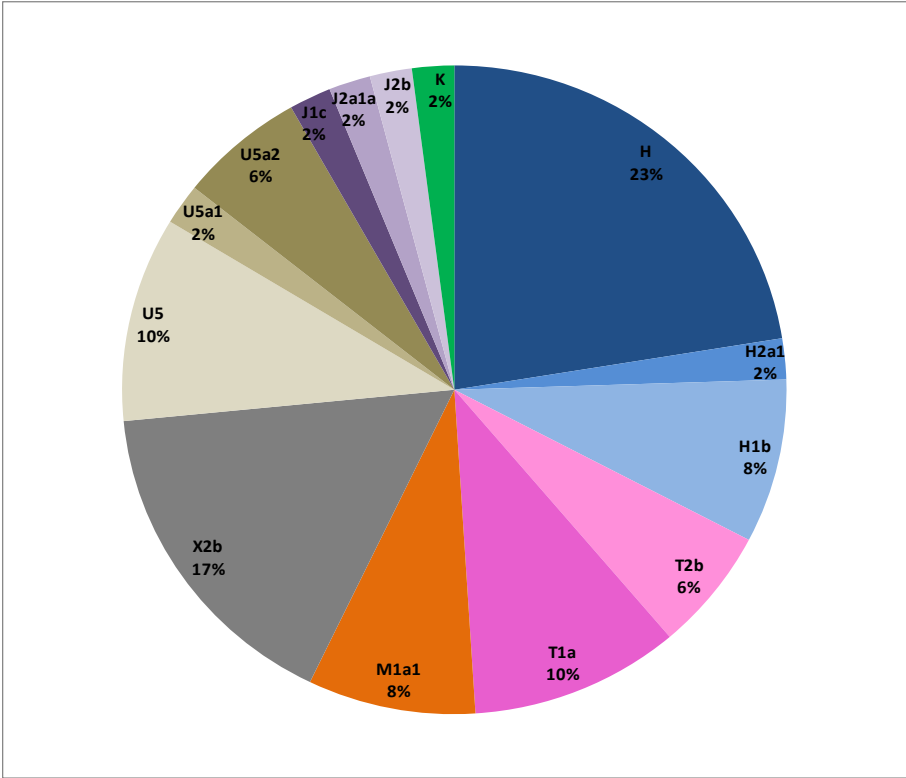


Figura 33. Aplogruppi dettagliati mtDNA Carloforte

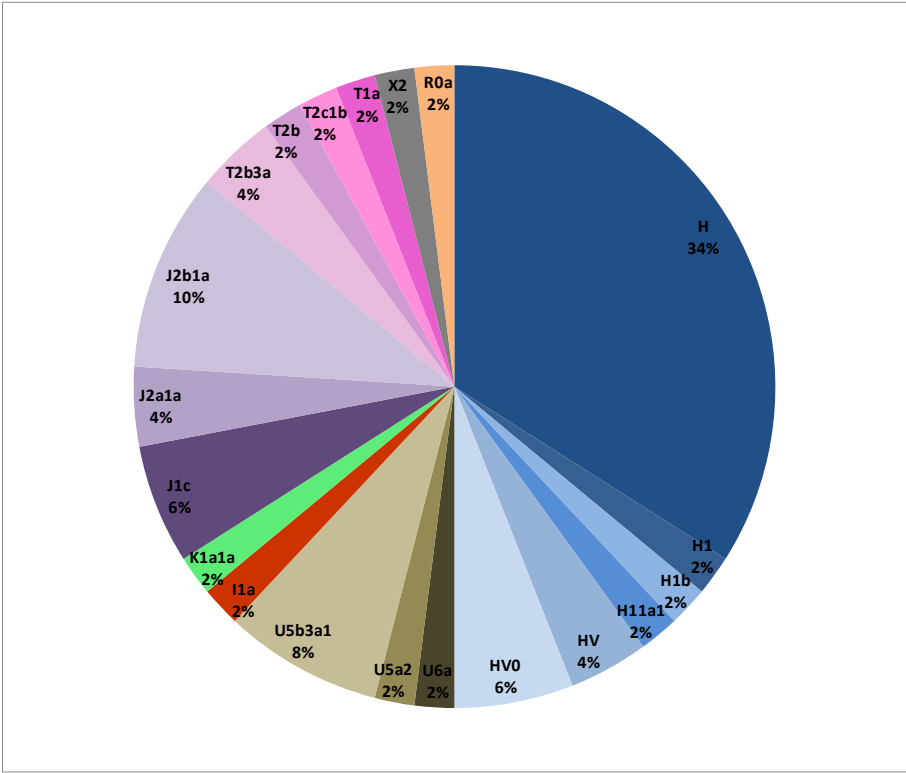


Figura 34. Aplogruppi dettagliati mtDNA Sulcis Iglesiente

3.2.3 Network filogenetici

Per poter analizzare le relazioni filogenetiche degli aplotipi emersi dall'analisi dei dati delle popolazioni è stato costruito un network che mostra la clusterizzazione degli aplotipi in base alla loro origine.

Nel grafico i cerchi rappresentano gli aplotipi la cui area è direttamente proporzionale alla frequenza dell'aplotipo nella popolazione. Le mutazioni sono rappresentate con delle linee.

Il network di Benetutti (Fig. 35) mostra un'evidente forma starlike, con un aplotipo centrale maggiormente rappresentato, da cui si dipartono tutte le ramificazioni.

Questo tipo di distribuzione indica un'omogeneità della popolazione iniziale ed un'origine monofiletica degli aplotipi.

Successivamente è stato associato ad ogni aplotipo il corrispettivo aplogruppo e si è visto che vi è una corrispondenza fra cluster aplotipici e aplogruppi (i due principali sono mostrati nel network in questione).

La forma del network filogenetico per la popolazione di confronto (Fig. 36) indica, invece, un'origine eterogenea della popolazione del nord Sardegna. Infatti, non vi è un aplotipo centrale più rappresentato degli altri, ma sono evidenti diversi nuclei da cui si dipartono i vari rami.

Riguardo alla distribuzione degli aplogruppi, essi sono sempre ben separati e raggruppati per cluster.

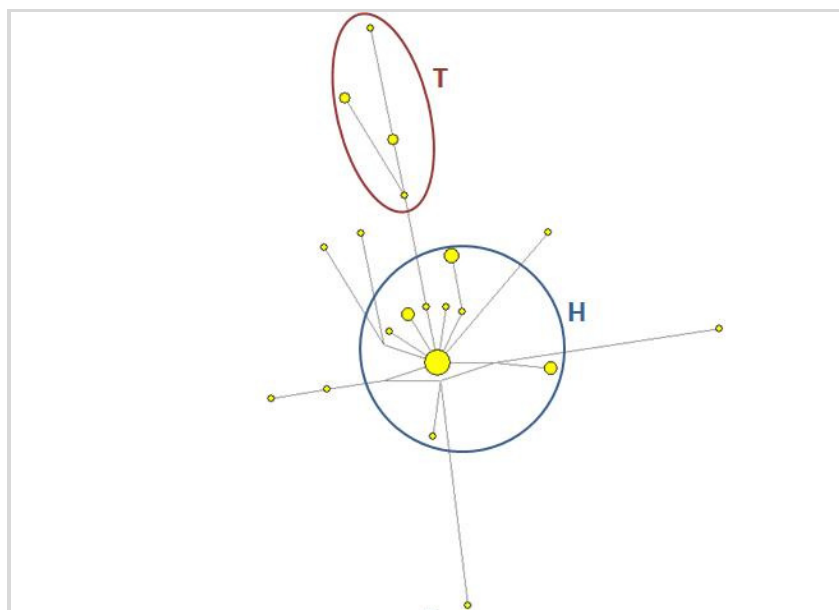


Figura 35. Network mtDNA Benetutti

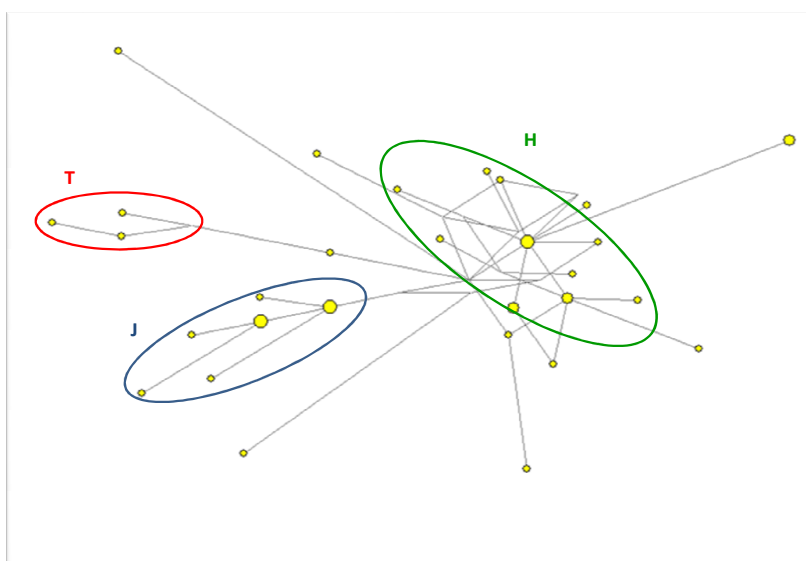


Figura 36. Network mtDNA nord Sardegna

Anche per l'isolato di Carloforte il network presenta un'evidente forma starlike con un cluster aplotipico centrale maggiormente rappresentato, corrispondente all'aplogruppo H, da cui si dipartono tutte le ramificazioni (Fig. 37). Quindi, anche per Carloforte si evidenzia una omogeneità della popolazione iniziale ed un'origine monofiletica degli aplotipi. Gli

aplogruppi appaiono discretamente separati gli uni dagli altri. Il network della popolazione di confronto presenta anch'esso una configurazione starlike, con un aplotipo centrale maggiormente rappresentato da cui gli altri hanno avuto origine, ma la forma denota una maggiore variabilità interna.

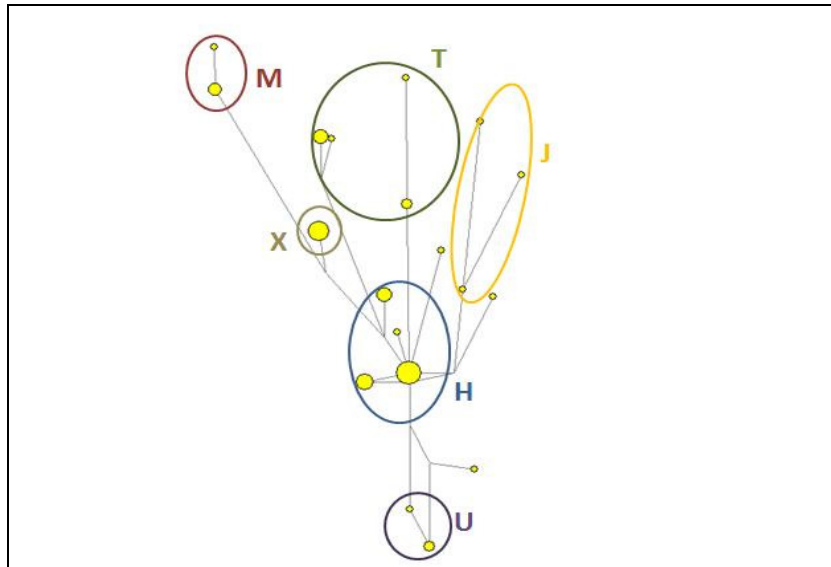


Figura 37. Network mtDNA Carloforte

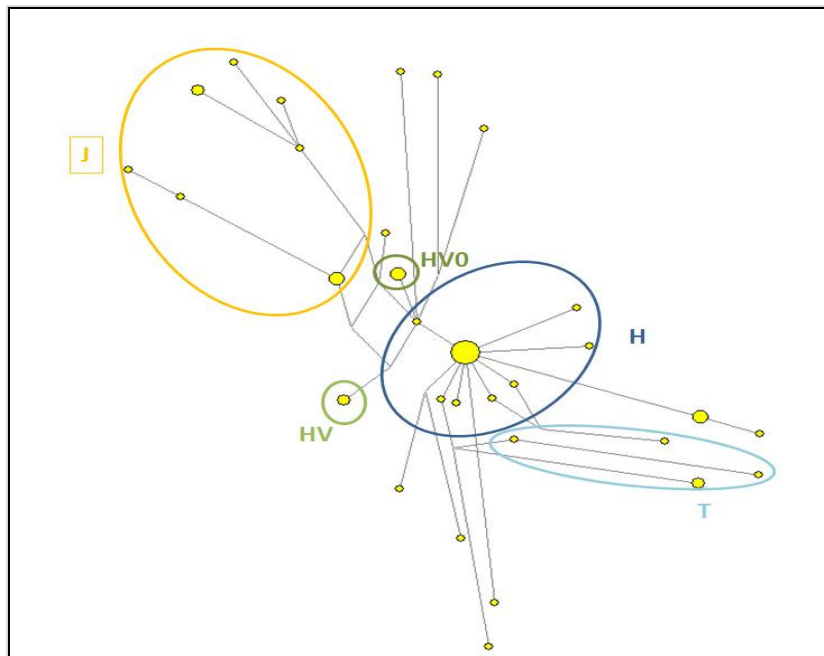


Figura 38. Network mtDNA del Sulcis Iglesiente

3.2.4 Distanze ed alberi genetici

È stato possibile attuare dei confronti tra le nostre comunità ed altre popolazioni dell'area del Mediterraneo grazie alle sequenze del segmento ipervariabile I disponibili in bibliografia e fruibili dai database di GenBank e Phylotree (Varesi et al., 2000; Krings et al., 1999; Fadhloui-Zid et al., 2004; Falchi et al., 2006; Côté-Real et al., 1996; Vona et al., 2001; Comas et al., 1996; Francalacci et al., 1996; Plaza et al., 2003; Rando et al., 1998; Stenico et al., 1996; Fraumene et al., 2006).

La matrice delle distanze F_{ST} , elaborata con il programma Arlequin 3.1, è riportata in tabella 30.

Benetutti mostra una distanza maggiore con: Algeria (0,130), Berberi (0,081), Marocco e Medio Oriente (0,069), e una distanza minore con: Trexenta (0,009), Sicilia (0,018) e Toscana (0,019).

Carloforte presenta una distanza maggiore con: Algeria (0,060), Berberi (0,055), Benetutti (0,051) e nord Sardegna (0,048), e una distanza minore con Berberi della Tunisia (0,013), Marocco (0,018) e Tunisia (0,019).

	Cen.Cors	Egitto	Berberi	Treventia	Algeria	M.Oriente	Sicilia	Turchia	Marocco	S.Antico	Gallura	Nuorese	Catòfort	S.Corsica	Andalusia	Baleari	Italia	Toscana	Berb.Tunis	Catalogna	Tunisia	N.Sardegna	Beretutti	Ogliastra
Cen.Cors	1	0,04536	0,04497	0,0879	0,1040	0,05474	0,01162	0,01189	0,03408	0,0222	0,00674	0,01065	0,03082	0,01691	0,00822	0,00386	0,03687	0,00689	0,02558	0,00857	0,03114	0,02351	0,02406	0,01032
Egitto	0,04536	1	0,04846	0,05483	0,06151	0,01346	0,03607	0,02431	0,02012	0,04593	0,04408	0,02398	0,02034	0,05264	0,04636	0,03894	0,04889	0,02759	0,01045	0,04173	0,01023	0,05097	0,06806	0,05404
Berberi	0,04497	0,04846	1	0,04844	0,0562	0,06876	0,05104	0,05333	0,05007	0,06162	0,0545	0,06385	0,05534	0,06699	0,06239	0,04484	0,06786	0,03983	0,03478	0,05106	0,04615	0,07126	0,08111	0,05889
Treventia	0,0879	0,05483	0,04844	1	0,08775	0,04666	0,0278	0,02104	0,04	0,00485	0,00135	0,00051	0,03957	0,00488	0,01002	0,0086	0,03024	0,0019	0,03691	0,00736	0,04031	0,0176	0,00902	0,00061
Algeria	0,1040	0,06151	0,0562	0,08775	1	0,08273	0,03867	0,08869	0,04363	0,0965	0,08961	0,0881	0,06019	0,09558	0,09176	0,084	0,11003	0,07909	0,04613	0,08972	0,06452	0,12031	0,13053	0,10065
M.Oriente	0,05474	0,01346	0,04866	0,08273	1	0,03055	0,01328	0,0309	0,03004	0,03936	0,03936	0,05579	0,02487	0,05073	0,0433	0,03744	0,03956	0,01743	0,02193	0,04988	0,02853	0,03375	0,06873	0,04932
Sicilia	0,01162	0,03607	0,05104	0,0278	0,03867	1	0,03055	0,0323	0,0236	0,00098	0,00166	0,00763	0,02828	0,01569	0,01051	0,00397	0,03076	-0,00287	0,02174	0,00821	0,03005	0,00395	0,01833	0,00929
Turchia	0,01189	0,05333	0,05333	0,02104	0,08869	0,0328	1	0,02148	0,01586	0,01028	0,02433	0,0232	0,02221	0,01889	0,01889	0,01275	0,03565	-0,00037	0,01306	0,01704	0,01406	0,01532	0,03218	0,02175
Marocco	0,03408	0,02012	0,05007	0,04	0,04363	0,0309	0,0236	0,02148	1	0,03341	0,03222	0,04351	0,01808	0,04656	0,02803	0,03354	0,06416	0,01953	0,01091	0,03565	0,0155	0,04988	0,06963	0,03851
S.Antico	0,0222	0,04593	0,06162	0,04665	0,0965	0,03004	0,00098	0,01586	0,03341	1	0,00043	0,01329	0,03465	0,03334	0,01805	0,01162	0,02761	-0,00007	0,0348	0,02719	0,04134	-0,00395	0,02311	0,00464
Gallura	0,00674	0,04408	0,0545	0,00135	0,08961	0,03936	0,00166	0,01028	0,03222	0,00043	1	0,00119	0,03202	0,01194	0,00379	0,00368	0,02878	-0,00449	0,02476	0,0094	0,02987	0,01306	0,02115	0,00295
Nuorese	0,01065	0,05298	0,06385	0,00051	0,0881	0,05579	0,00763	0,02433	0,04351	0,01329	0,00119	1	0,03539	0,00347	0,0031	0,0456	0,03088	0,00773	0,03134	0,00446	0,04398	0,02452	0,01993	0,00174
Catòfort	0,03082	0,02034	0,05534	0,03957	0,06019	0,02487	0,02828	0,02332	0,01808	0,03465	0,03202	0,03539	1	0,04132	0,02846	0,02399	0,03961	0,0286	0,01835	0,02135	0,01937	0,04832	0,05155	0,03952
S.Corsica	0,01691	0,05264	0,06699	0,04889	0,09558	0,05073	0,01569	0,02221	0,04656	0,03334	0,01194	0,00347	0,04132	1	0,01126	0,01009	0,03196	0,00937	0,03144	0,00609	0,04172	0,03702	0,02187	0,02357
Andalusia	0,00822	0,04636	0,06239	0,01002	0,09176	0,0433	0,01051	0,01889	0,02803	0,01805	0,00379	0,0031	0,02948	0,01126	1	0,00354	0,06315	0,00515	0,02633	0,00545	0,03466	0,02806	0,03227	0,01373
Baleari	0,00386	0,03894	0,04484	0,0066	0,084	0,03744	0,00397	0,01275	0,03354	0,01162	0,00688	0,0468	0,02399	0,01009	0,00654	1	0,01951	0,00167	0,01953	0,00264	0,03557	0,01539	0,02023	0,0159
Italia	0,03687	0,04889	0,06786	0,03024	0,11003	0,03956	0,03076	0,03565	0,06416	0,02761	0,02878	0,03088	0,03961	0,03196	0,03615	0,01951	1	0,02528	0,04229	0,0323	0,06021	0,03336	0,03748	0,03408
Toscana	0,00689	0,02759	0,03983	0,0019	0,07909	0,01743	-0,00287	-0,00037	0,01953	-0,00007	-0,00449	0,00773	0,0266	0,00937	0,00515	0,00167	0,02528	1	0,01843	0,00663	0,02054	0,00812	0,0189	0,00765
Berb.Tunis	0,02558	0,01045	0,03478	0,03591	0,04613	0,02174	0,01306	0,01306	0,01091	0,0348	0,02476	0,03134	0,01335	0,03144	0,02633	0,01953	0,04229	0,01943	1	0,01928	0,00433	0,04361	0,04811	0,03793
Catalogna	0,00857	0,04173	0,05106	0,00736	0,08972	0,04888	0,00821	0,01704	0,03565	0,02719	0,0094	0,0446	0,02135	0,00609	0,00545	0,00264	0,0323	0,00663	0,01928	1	0,0248	0,03633	0,01954	0,0146
Tunisia	0,03114	0,01023	0,04615	0,04031	0,06452	0,02853	0,03005	0,01406	0,0155	0,04134	0,02897	0,04398	0,01937	0,04172	0,03466	0,03557	0,06021	0,02054	0,00433	0,0248	1	0,05009	0,06271	0,0451
N.Sardegna	0,02351	0,05097	0,07126	0,0176	0,12031	0,03375	0,00395	0,01532	0,04986	-0,00395	0,01306	0,02462	0,04632	0,03702	0,02806	0,01539	0,03336	0,00812	0,04361	0,03633	0,05009	1	0,02308	0,02149
Beretutti	0,02406	0,06806	0,08111	0,00902	0,13053	0,06873	0,01833	0,03218	0,06963	0,02311	0,02115	0,01993	0,05155	0,02187	0,03227	0,02023	0,03748	0,0189	0,04811	0,03633	0,06271	0,02308	1	0,02547
Ogliastra	0,01032	0,05404	0,05889	0,00061	0,10065	0,04932	0,00929	0,02175	0,03951	0,00464	0,00295	0,00174	0,03952	0,02357	0,01373	0,01659	0,03408	0,00765	0,03793	0,0146	0,0451	0,02149	0,02547	1

Tabella 30. Matrice delle distanze Fst per le popolazioni del Mediterraneo

Dall'analisi dell'albero (Fig. 38) si può notare come Benetutti abbia clusterizzato con la Trexenta e si mostri, invece, differenziata dal nord Sardegna. È interessante rilevare come le varie zone della Sardegna non abbiano clusterizzato assieme denotando un certo grado di eterogeneità, peraltro già riscontrata in precedenti lavori (Gatti, 1990; Walter et al., 1991; Morelli et al., 1993; Moral et al., 1994; Zavattari et al., 2000; Angius et al., 2001; Calò et al., 2008; Vona e Calò, 2006; Pistis et al., 2009).

Carloforte ha invece clusterizzato con le popolazioni del nord Africa, mostrando una certa affinità con queste e una netta differenziazione sia con la attigua zona del Sulcis Iglesiente, sia verso tutte le altre zone della Sardegna e dell'Italia.

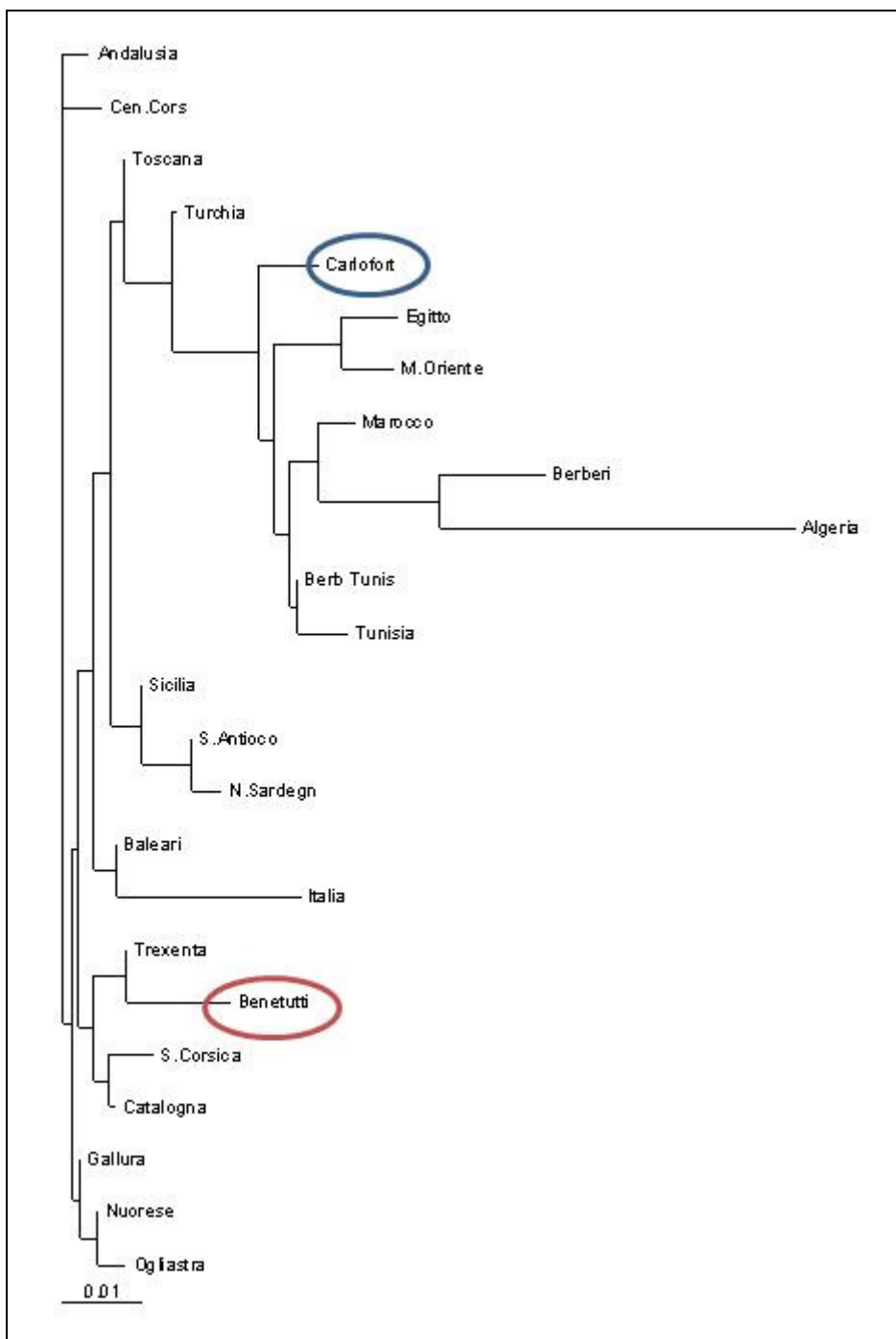


Figura 38. Albero mtDNA con le popolazioni del Mediterraneo

Per meglio indagare tale situazione è stata fatta un'analisi prendendo in considerazione le sole popolazioni della Sardegna. La matrice delle distanze Fst fra le popolazioni analizzate è riportata nella tabella 31.

Benetutti mostra la distanza minore con la Trexenta, il Nuorese, la Gallura e il nord Sardegna e la distanza maggiore con Carloforte.

Carloforte mostra la distanza minore con la Gallura e Sant'Antioco, e la distanza maggiore con Benetutti e il nord Sardegna.

	Trexenta	S.Antioco	Gallura	Nuorese	Carloforte	N.Sardegna	Benetutti	Ogliastra
Trexenta								
S.Antioco	0,00485							
Gallura	0,00135	0,00043						
Nuorese	0,00051	0,01329	0,00119					
Carloforte	0,03957	0,03465	0,03202	0,03539				
N.Sardegna	0,0176	-0,00395	0,01306	0,02452	0,04832			
Benetutti	0,00902	0,02311	0,02115	0,01993	0,05155	0,02308		
Ogliastra	0,00061	0,00464	0,00295	0,00174	0,03952	0,02149	0,02547	

Tabella 31. Matrice delle distanze Fst per la Sardegna

Dall'albero, costruito dalla matrice di tali distanze Fst (Fig. 39), si può notare come sia Carloforte che Benetutti si trovino su rami separati e distanti dalle restanti popolazioni della Sardegna denotando un certo grado di isolamento. Nord Sardegna e Sant'Antioco si trovano, invece, molto prossime.

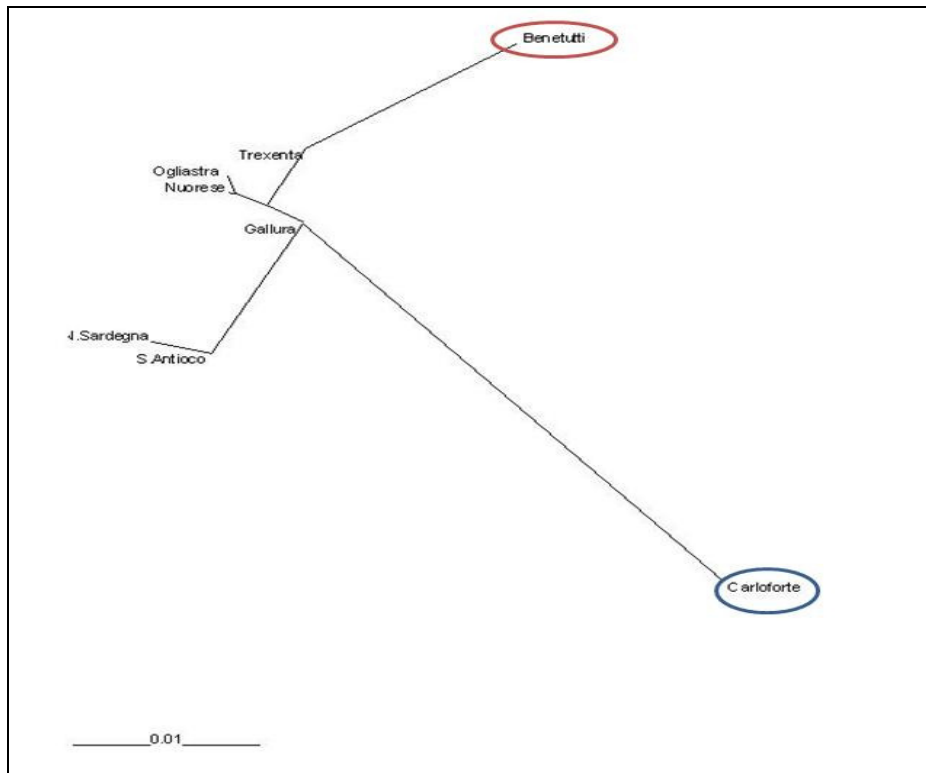


Figura 39. Albero mtDNA Sardegna.

Per valutare ulteriormente la struttura genetica delle popolazioni è stato utilizzato il programma STRUCTURE . Si è usato un numero ipotetico di popolazioni (K) pari a 2, 3 e 4 , e per ciascuna di queste abbiamo ripetuto l'analisi quattro volte. La migliore interazione si è avuta per $K = 3$. poichè presenta i valori più bassi di Ln P.

Questo risultato denoterebbe un alto grado di omogeneità genetica tra le popolazioni analizzate, anche se è necessario sottolineare che il programma è stato creato per i marcatori autosomici e solo successivamente adattato a sequenze del mitocondriale. Per questo motivo il risultato è da prendere con cautela.

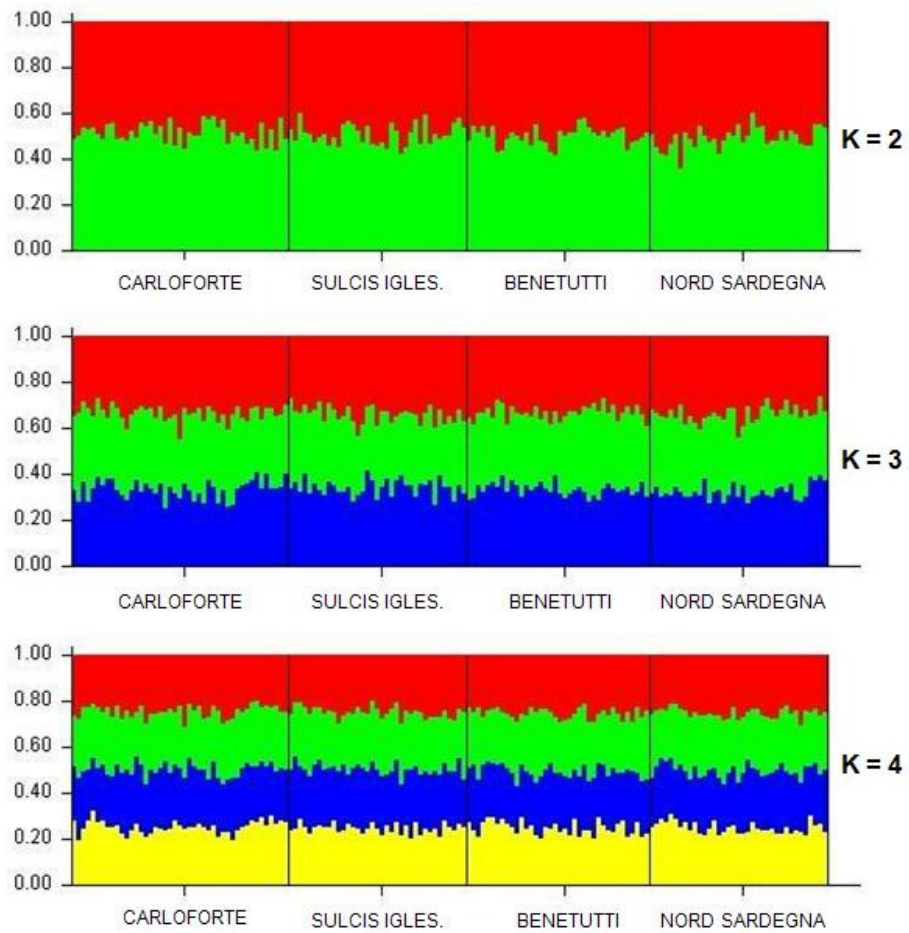


Figura 40. Risultati di STRUCTURE per il mtDNA con il presupposto di differenti gruppi di popolazioni (K=2, 3, 4). K=3 presenta il più alto valore di LnP(D)

4. Confronto fra due diversi metodi di campionamento

Per quanto riguarda Carloforte, abbiamo effettuato un confronto dei risultati ottenuti dal segmento HVRI del DNA mitocondriale acquisiti mediante due differenti strategie di campionamento: quello in base ai cognomi fondatori e quello secondo le genealogie. Questo è stato possibile poiché per studi precedenti, condotti nel nostro laboratorio, inerenti il segmento ipervariabile I del mtDNA, era stato effettuato un campionamento secondo le genealogie di individui appartenenti alla comunità Carlofortina. Queste persone erano originarie di Carloforte da almeno tre generazioni, non imparentate fra loro, di entrambi i sessi e apparentemente sane.

In un lavoro recente (Sanna et al., 2011) vi sono delle riflessioni interessanti su quanto sia importante in studi di genetica di popolazione la definizione della popolazione, la selezione del campione, la scelta del marcatore e la strategia di campionamento che dipende dalle domande che ci si pone.

Secondo tale ricerca, se una popolazione umana è identificata etnicamente con chiarezza e la recente commistione è trascurabile, i campionamenti che sono basati solo sul cognome e sul luogo di nascita sono quelli preferibili, poiché consentono l'esclusione dall'analisi degli immigrati recenti che non si sono ancora miscelati nel pool genico.

Inoltre, i criteri del cognome e del luogo di nascita possono essere estesi dai donatori ai loro antenati quando sono disponibili le informazioni genealogiche. Una strategia di campionamento più rigorosa è richiesta in studi di associazione "genome-wide" che cercano differenti distribuzioni alleliche fra individui con (detti casi) o senza (chiamati controlli) un fenotipo di interesse.

Mediante l'analisi delle sequenze del segmento ipervariabile I tramite il programma arlequin (ver.3.1), si sono rilevati alcuni parametri di indagine intra-popolazioneistica i cui risultati sono riportati nella tabella 32.

Con entrambe le strategie di campionamento è stato rilevato un alto livello di diversità genetica, ma con il metodo dei cognomi fondatori è stato riscontrato un maggior numero di aplotipi differenti (49/49) e un valore di diversità genetica pari a 1, che significa che tutti gli individui mostrano sequenze differenti. Anche il numero medio di differenze fra coppie di sequenze e la media della diversità genica tra loci sono risultati maggiori nel campionamento per cognomi fondatori. Si può quindi affermare che questa strategia di campionamento abbia massimizzato l'eterogeneità del campione.

	Camp.Cognomi fond.	Camp. Genealogie
n° campioni	49	50
n° aplotipi diversi	49/49	38/50
gene diversity	1	0,978
mean number of pairwise differences	5,257	4,459592
average gene diversity over loci	0,015	0,012388

Tabella 32. Parametri rilevati dall'analisi delle sequenze HVRI.

4.1 Analisi degli Aplogruppi

Come si può notare dai grafici riportati nelle figure 41 e 42, nel campionamento secondo le genealogie vi è una maggiore variabilità evidenziata da un più grande numero di aplogruppi differenti. Fra gli aplogruppi condivisi vi sono delle differenze a livello di frequenza. Nel campionamento secondo le genealogie vi è un aplogruppo che predomina per frequenza (H 53%) seguito da J (14%) e V (9%), mentre nel campionamento secondo i cognomi fondatori vi sono più aplogruppi con ragguardevole

frequenza: H (33%), U (19%), X e T (16%), M (8%). Uno dei predetti (T) non è presente nel campionamento secondo le genealogie. In generale, il campionamento secondo le genealogie presenta gli stessi aplogruppi e frequenze simili a quelle riscontrate nei campioni della Sardegna.

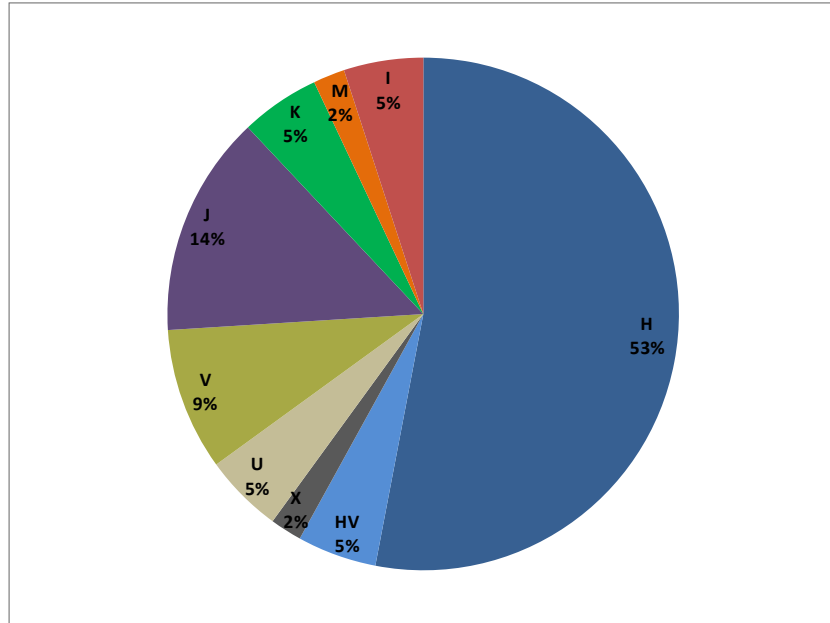


Figura 41. Aplogruppi mtDNA campionamento secondo le genealogie

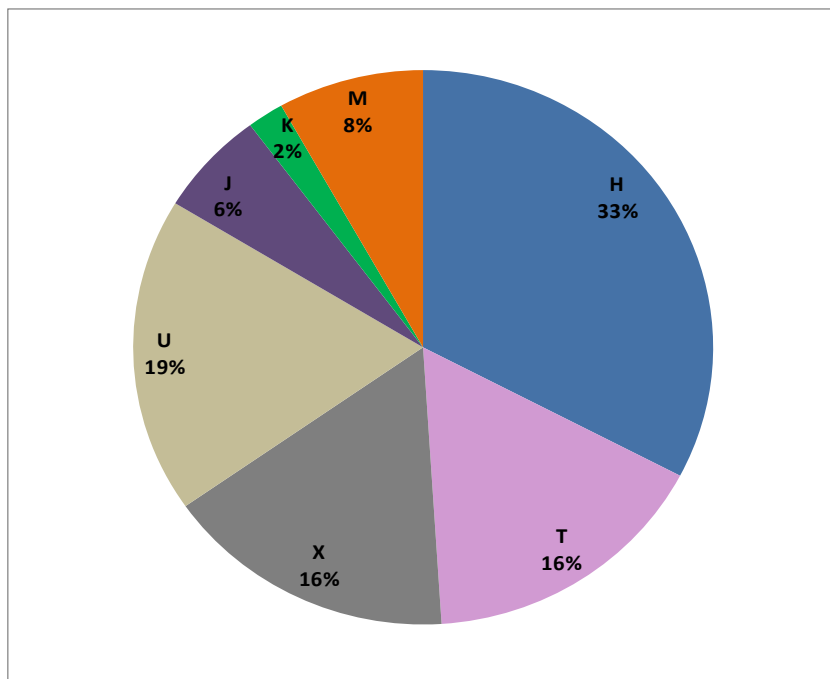


Figura 42. Aplogruppi mtDNA campionamento per cognomi fondatori

4.2 Network filogenetici

Entrambi i network presentano un cluster aplotipico centrale maggiormente rappresentato, corrispondente all'aplogruppo H, da cui si dipartono tutte le ramificazioni. Nel campionamento per cognomi fondatori vi sono altri cluster ben rappresentati che si dipartono dal nucleo principale (Fig. 43). Il network per il campionamento secondo le genealogie presenta sempre una forma starlike ma con una distribuzione degli aplotipi che evidenzia una maggiore variabilità interna (Fig. 44).

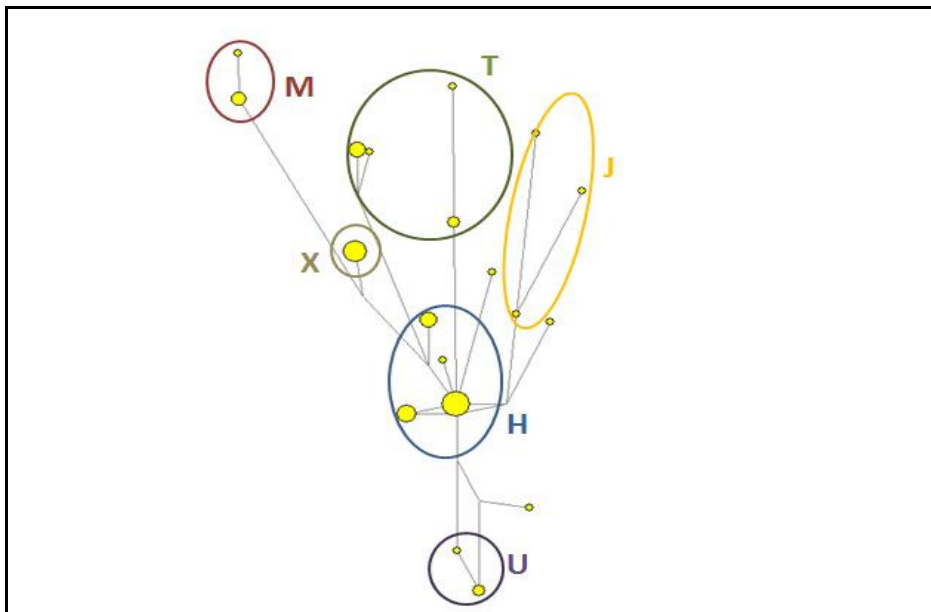


Figura43. Network mtDNA per il campionamento secondo i cognomi fondatori

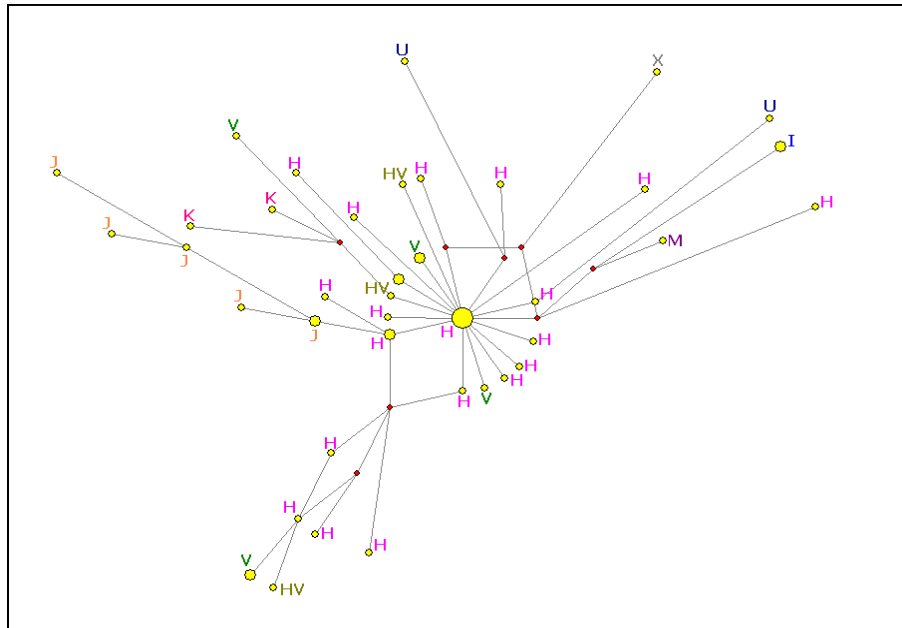


Figura 44. Network mtDNA per il campionamento secondo le genealogie.

4.3 Alberi filogenetici

Le sequenze del segmento ipervariabile I dei due diversi campionamenti effettuati per Carloforte sono state confrontate con quelle delle altre popolazioni sarde oggetto di studio e di altre popolazioni dell'area del Mediterraneo (Varesi et al.,2000; Krings et al.,1999; Fadhloui-Zid et al.,2004; Falchi et al.,2006; Côrte-Real et al.,1996; Vona et al.,2001; Comas et al., 1996; Francalacci et al., 1996; Plaza et al.,2003; Rando et al.,1998; Stenico et al.,1996; Fraumene et al., 2006). Nell'albero costruito con la popolazione carlofortina campionata secondo i cognomi fondatori Carloforte ha clusterizzato con le popolazioni del nord Africa e si presenta invece distante dall'Italia e dalle restanti comunità della Sardegna (Fig. 45). Nel campionamento secondo le genealogie, invece, Carloforte ha clusterizzato con l'Italia e si trova abbastanza vicina alle altre comunità sarde (Fig. 46). Abbiamo poi pensato di costruire un albero utilizzando i due distinti campionamenti per vedere come si collocassero l'uno rispetto

all'altro. È stato interessante osservare che questi non abbiano clusterizzato assieme ma si siano disposti l'uno lontano dall'altro (Fig. 47).

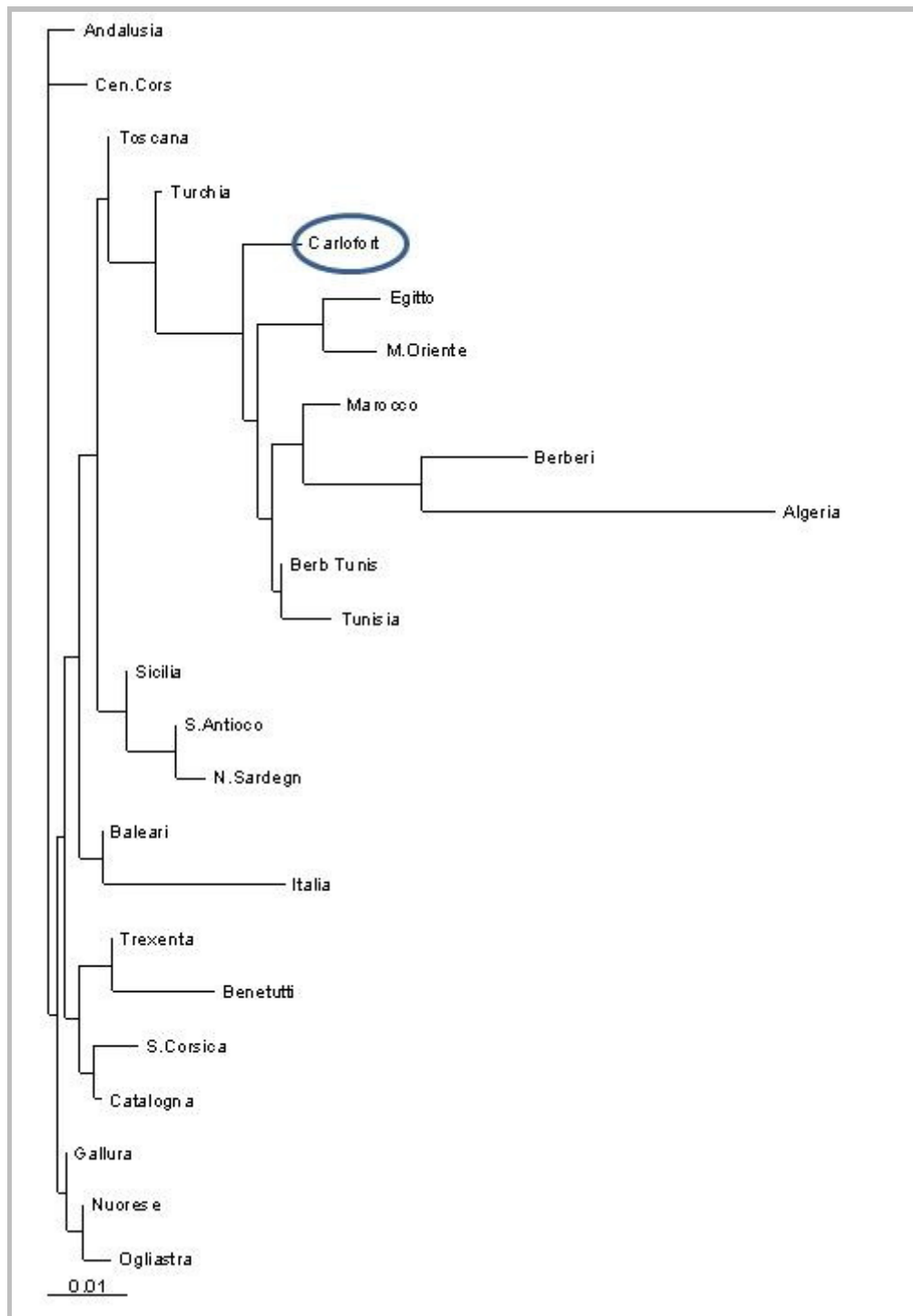


Figura 45. Albero mtDNA per le popolazioni del Mediterraneo (Carloforte è stata campionata secondo i cognomi fondatori).

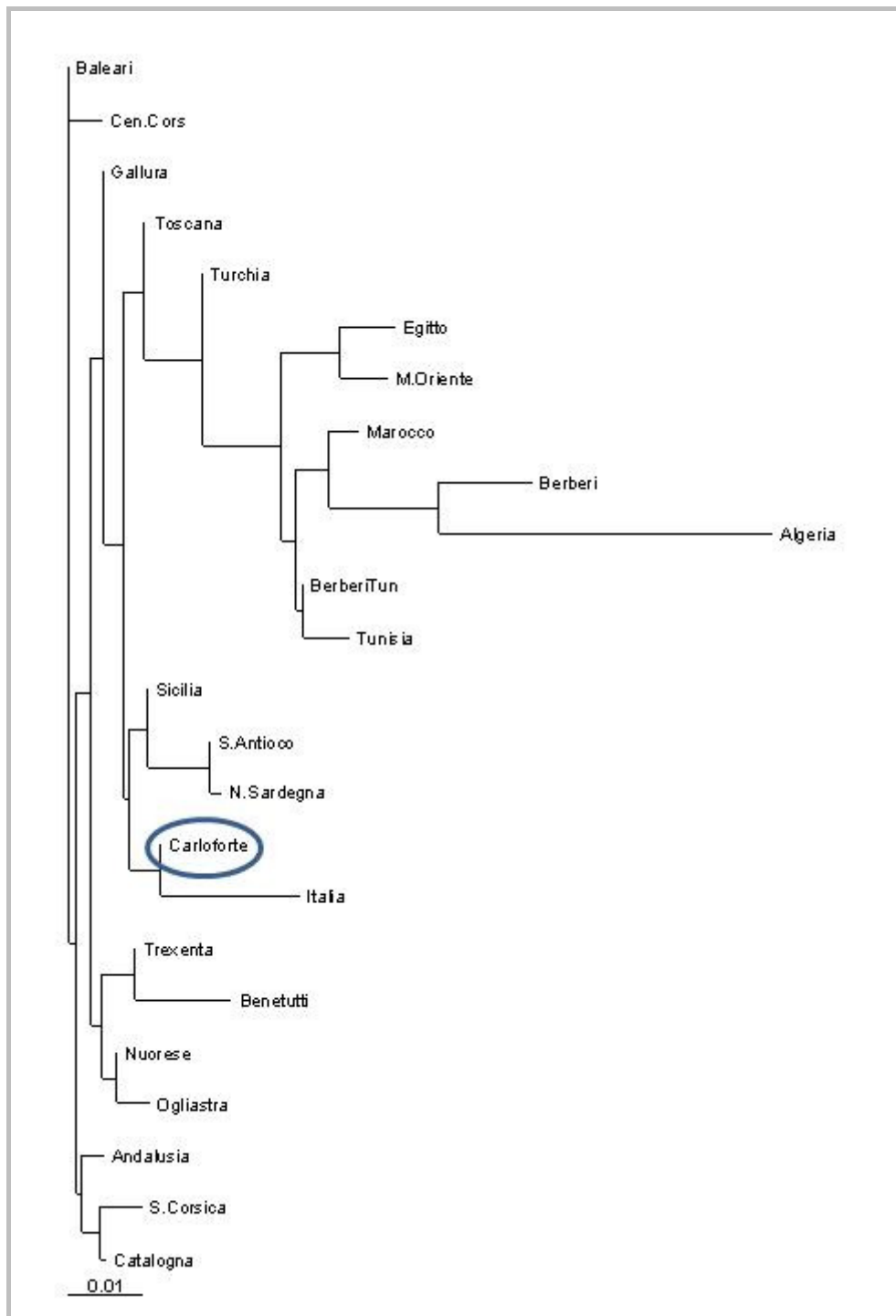


Figura 46. Albero mtDNA per le popolazioni del Mediterraneo (Carloforte è stata campionata secondo le genealogie).

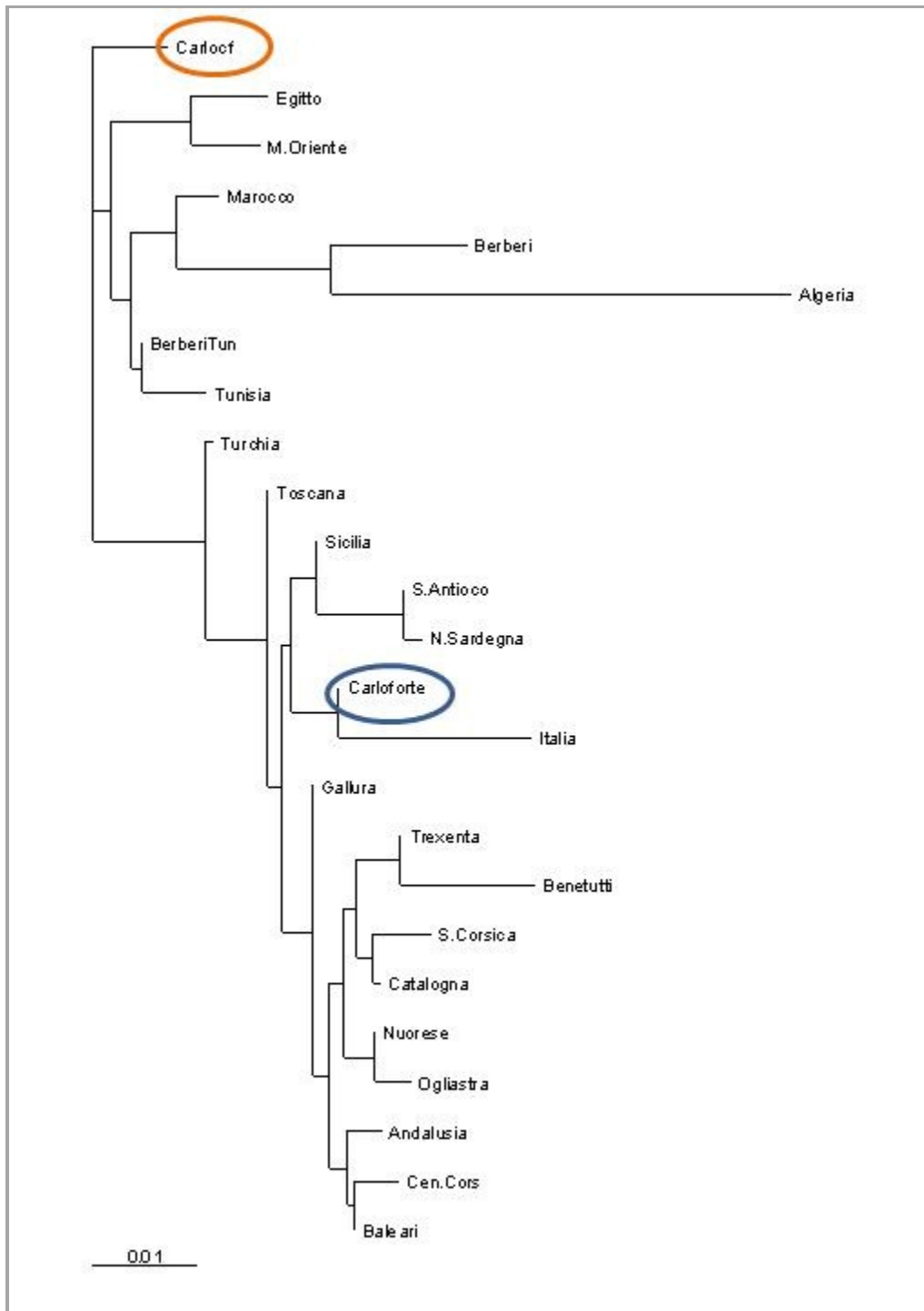


Figura 47. Albero mtDNA per le popolazioni del Mediterraneo. E' presente sia la popolazione di Carloforte campionata secondo le genealogie (chiamata Carloforte e indicata con il colore blu nella figura) sia quella campionata per cognomi fondatori (chiamata Carloof e indicata con il colore arancione nella figura).

5. Discussioni

Le ricerche su popolazioni isolate di piccole dimensioni sono un utile strumento per l'identificazione di associazioni genotipo-fenotipo, soprattutto considerando le enormi difficoltà incontrate nell'identificazione di loci per le malattie complesse (Heutink et al.,2002; Bourgain et al.,2001).

Inoltre, studiare gli effetti dell'isolamento sulla struttura genetica di una popolazione permette di valutare diversi aspetti: il grado di differenziazione rispetto alla popolazione di appartenenza; il grado di coinvolgimento nei flussi migratori che hanno investito l'area di interesse nei tempi preistorici e storici; l'effettiva "età" del pool genico degli isolati; la possibilità di riconoscere influssi di altre popolazioni; l'evoluzione autoctona con geni e aplotipi propri. Le migliori popolazioni candidate per gli studi di associazione sono quelle isolate che presentano un effettivo ridotto (10-100) di fondatori non imparentati tra loro. La Sardegna risponde perfettamente alle condizioni necessarie per questo tipo di studi, infatti, la sua popolazione è distribuita in pochi grandi aggregati urbani, di recente formazione e in molti piccoli villaggi isolati.

Per la realizzazione della presente ricerca sono state selezionate due comunità isolate: Carloforte (Sud ovest della Sardegna), che costituisce un isolato linguistico, e Benetutti (Centro Sardegna), un tipico isolato geografico. Il campionamento degli isolati è stato condotto secondo il principio dei cognomi fondatori e cioè scegliendo individui di sesso maschile che portavano i cognomi presenti al momento della fondazione della comunità. Inoltre, tali individui dovevano essere originari del paese e non imparentati fra loro da almeno tre generazioni. Per ogni cognome è stato campionato un solo individuo.

Lo scopo della ricerca è stato quello di analizzare i due isolati sardi, uno linguistico e uno geografico, attraverso i marcatori del DNA uniparentali, con lo scopo di valutare se le barriere culturali e linguistiche abbiano la stessa influenza delle barriere geografiche sull'isolamento e quindi nel determinare la struttura genetica di una popolazione. Si è poi voluto verificare se l'analisi comparativa dei risultati ottenuti con i marcatori ad eredità materna e paterna fornissero informazioni fra loro concordanti. Inoltre, per quanto riguarda Carloforte, si son messi a confronto i risultati del DNA mitocondriale ottenuti con due differenti metodi di campionamento: quello in base ai cognomi fondatori e quello secondo le genealogie, per capire come possano influire sui risultati differenti tecniche di campionamento.

L'analisi delle mutazioni presenti a livello della sequenza del cromosoma Y e del DNA mitocondriale ha fornito numerose informazioni sull'origine dell'uomo, la sua evoluzione e le migrazioni nei diversi continenti. Il motivo che rende i marcatori uniparentali indispensabili per studi filogenetici è la mancanza di ricombinazione; infatti, le mutazioni sono l'unica fonte di variabilità fra gli individui. Queste modificazioni sono avvenute durante e dopo il processo di colonizzazione e diffusione nelle differenti regioni geografiche, per cui esse costituiscono uno strumento fondamentale nell'analisi dei tempi di espansione, migrazione e altre forme di deriva genica in epoca preistorica e storica.

Per quanto riguarda l'analisi del cromosoma Y, in tutte le comunità oggetto di studio eccetto Carloforte è stata trovata la duplicazione 15-16 nel locus DYS19, già rilevata in un recente lavoro sulla popolazione del nord Sardegna (Ghiani et al., 2009). In questo lavoro è stato anche rilevato nel locus DYS458 l'allele non-consenso 18.2 che abbiamo trovato solo a Benetutti e nel rispettivo campione di controllo.

La differenziazione delle popolazioni, calcolata tramite l' F_{st} , è risultata statisticamente significativa (p -value $<0,005$) per entrambi gli isolati. A conferma di tale dato Carloforte condivide con il suo campione di confronto un solo aplotipo, mentre Benetutti non condivide nessun aplotipo con il campione del nord Sardegna. Inoltre, i valori di diversità genetica, prossimi a 1, sono indice di un alto grado di variabilità interna dei campioni. Sia il numero medio di differenze fra coppie di sequenze che la media della diversità genica tra loci sono risultati maggiori in entrambi i campioni di controllo rispetto ai due isolati indicando una maggiore eterogeneità nei campioni di confronto.

Questi dati suggeriscono che i campioni del Nord Sardegna e del Sulcis Iglesiente sono stati, nel corso della loro storia, caratterizzati da un minor grado di isolamento geografico e culturale rispetto a Benetutti e a Carloforte. Possiamo presumere che i due campioni di confronto abbiano avuto un maggior flusso genico che ha portato a un pool genico più eterogeneo. A tal proposito è necessario tenere in considerazione che si è confrontato un piccolo comune con un'area più estesa che comprende più paesi, per cui non è una sorpresa aver trovato una minore variabilità genetica nei piccoli comuni. Il maggior grado di isolamento di Benetutti e Carloforte è confermato anche dagli studi biodemografici sui tassi di endogamia (84,4% e 74,42% rispettivamente) e di consanguineità ($\alpha=3.85 \times 10^{-3}$ e $\alpha=1,63 \times 10^{-3}$ rispettivamente) (Sanna et al., 2004; Calò et al., 1994).

Questi fattori sono lo specchio del forte isolamento che caratterizza queste due comunità. Per Benetutti il motivo dell'isolamento è di tipo geografico, dovuto soprattutto alle cattive condizioni delle vie di collegamento stradale e all'orografia del territorio. Per Carloforte è attribuibile soprattutto a fattori culturali come la diversità della lingua, di origine ligure (dialetto pegliese), che ha impedito l'omologazione di

Carloforte alle altre popolazioni della Sardegna, conservando un maggior grado di affinità genetica con le popolazioni italiane continentali, contrariamente a quanto ci si attenda in base alla prossimità geografica (Vona, 1996).

L'analisi del network filogenetico ha mostrato un certo grado di eterogeneità della popolazione fondatrice sia per Benetutti che per Carloforte, dovuta verosimilmente alla strategia di campionamento che ha massimizzato la diversità genetica presente nei due isolati.

Risultati interessanti provengono dall'analisi degli aplogruppi; infatti, l'aplogruppo I_{2a}₁ ha presentato frequenze molto diverse tra le popolazioni esaminate: 50% a Benetutti, 34% nel nord-Sardegna, 2% a Carloforte e 38% nel Sulcis Iglesiente. La frequenza di tale aplogruppo a Benetutti è uguale a quella riscontrata da Zei e collaboratori (2003) in quella che è stata definita zona "arcaica" della Sardegna, e localizzata nella parte centrale e interna all'isola. Frequenze così differenti tra gli isolati e le zone limitrofe attestano scarso flusso genico e quindi un forte isolamento. A Carloforte, inoltre, tale aplogruppo è praticamente assente denotando una grande somiglianza con le popolazioni dell'Italia continentale. Infatti, la mutazione M26 che definisce l'aplogruppo I_{2a}₁ è assente nell'Italia settentrionale e presente a bassissime frequenze nell'Italia centro-meridionale (dallo 0,7% al 5%) (Semino et al., 2000; Scozzari et al., 2001; Francalacci et al., 2003; Rootsi et al., 2004).

Le peculiari frequenze di questo aplogruppo erano state studiate in precedenti ricerche (Zei et al., 2003) soprattutto in relazione alla mutazione M26, che definisce l'aplogruppo I_{2a}₁ secondo la nomenclatura utilizzata dall' International Society of Genetic Genealogy presa come riferimento in questa tesi. Infatti, il subclade I-M26 raggiunge frequenze molto alte in Sardegna (circa 37%) (Ghiani et al., 2009; Contu et al., 2008; Francalacci et al., 2003; Semino et al., 2000; Passarino et al., 2001) soprattutto nell'area arcaica (Zei

et al., 2003). Tale frequenza viene spiegata con l'effetto del fondatore, e potrebbe essere il retaggio del popolamento dell'isola tra il Paleolitico superiore (24.000-20.000 anni fa) ed il Mesolitico (20.000-15.000 anni fa) (Semino et al., 2001).

Questa mutazione è assente in Liguria come in tutto il nord Italia (Scozzari et al., 2001; Semino et al., 2000; Francalacci et al., 2003; Rootsi et al., 2004). In Europa è stata rilevata, anche se a frequenze molto basse (>7%), solo fra Baschi, Portoghesi, Andalusi, Spagnoli dei Pirenei, Francesi, Irlandesi, Scozzesi, Gallesi, Inglesi e in alcune zone limitate dell'Italia centro meridionale, (Rootsi et al., 2004; Onofri et al., 2007; Semino et al., 2000; Capelli et al., 2003; Capelli et al., 2005; Zei et al., 2003; Di Gaetano et al., 2009; Bosch et al., 2001, Francalacci et al., 2003; Cinnioglu et al., 2004; Lopez-Parra et al., 2009).

Un' altro aplogruppo ben rappresentato in Sardegna è risultato l'R_{1b} che presenta una frequenza del 4% a Benetutti, del 21% nel nord Sardegna, del 55% a Carloforte e del 25% nell'area del Sulcis Iglesiente. La frequenza riscontrata a Carloforte è la stessa trovata in Italia nelle regioni del centro-nord, a ulteriore riprova che la popolazione carlofortina ha subito un forte isolamento rispetto alle comunità sarde che le ha permesso di mantenere le sue caratteristiche genetiche ancestrali di origini Liguri (Ferri et al., 2009; Semino et al., 2000).

L'eterogenea distribuzione delle frequenze dell'aplogruppo R_{1b}, trova una conferma in un precedente studio di Zei e collaboratori (2003) che analizzava la distribuzione eterogenea del subcluster R-M269 in Sardegna. Infatti, se mediamente l'aplogruppo generico R-M173 (Eu 18 secondo una vecchia denominazione) si trova nell'isola con una frequenza di circa il 22%, la più bassa trovata in Europa Occidentale (50-80%) (Semino et al., 2000), la ripartizione del sub cluster R-M269 ha mostrato una forte eterogeneità presentando basse frequenze (10%) nella zona centro-orientale (zona

“arcaica”) e una maggiore frequenza nel nord (29,3%) e sud-ovest dell’Isola (25,5%). Tale distribuzione ha suggerito che questo sub-cluster sia arrivato verso le coste della Sardegna dal continente, probabilmente dopo l’insorgenza e la diffusione dell’autoctono subaplogruppo I-M26 (Zei et al., 2003).

Il presente lavoro ha quindi confermato che la maggior parte dei campioni sardi presentano gli aplogruppi E-M35, I-M170, J-12f2, G-201 e R-M173 (Zei et al., 2003). Pertanto, il pool degli aplogruppi del cromosoma Y in Sardegna è molto simile a quello degli italiani e degli altri Europei (Underhill et al., 2000; Underhill et al., 2001; Semino et al., 2000; Passarino et al., 2001). Le differenze osservate sono a livello di frequenze, dovute probabilmente all’isolamento e alla deriva genetica che hanno creato i presupposti affinché la frequenza di determinati aplogruppi aumentasse sensibilmente.

I confronti con altre popolazioni del Mediterraneo (Turrina et al., 2006; Onofri et al., 2007; Pokupcic et al., 2008; Palet et al., 2010; Martin et al., 2004; Frigi et al., 2006; Kovatsi et al., 2009) hanno confermato la peculiarità genetica della popolazione sarda. Carloforte si trova su un ramo a parte, vicina alla Spagna e all’Italia del nord e del centro, molto distante dalle popolazioni della Sardegna. Questo dato conferma, ancora una volta, il forte isolamento di questa popolazione che solo negli ultimi decenni sta andando diminuendo come dimostra un recente lavoro sui matrimoni endogamici ed esogamici dal 1738 al 2001 (Sanna *et al.*, 2006) in cui è evidenziata una certa tendenza al decremento dell’endogamia e della consanguineità. Un basso grado di omogeneizzazione tra sardi e carlofortini e una maggiore affinità di questi ultimi con gli italiani continentali era già stato evidenziato da studi sui marcatori classici e sulla struttura demografica (Vona et al., 1994; Vona et al., 1996). Carloforte si è collocata molto distante anche rispetto alla popolazione tunisina, dimostrando che le due popolazioni non si sono ibridate durante la loro permanenza nell’isola di Tabarka.

Le altre popolazioni della Sardegna hanno invece clusterizzato insieme, pur localizzandosi su rami differenti dell'albero denotando una certa differenziazione interna della popolazione sarda. Questo risultato trova riscontro con quanto riportato in lavori precedenti condotti sui marcatori classici e poi confermati da studi più recenti con i marcatori molecolari (Gatti, 1990; Walter et al., 1991; Morelli et al., 1993; Moral et al., 1994; Zavattari et al., 2000; Angius et al., 2001; Calò et al., 2008; Vona e Calò, 2006; Pistis et al., 2009). L'albero genetico ha comunque evidenziato una certa affinità tra Benetutti e il campione di confronto, un risultato non in pieno accordo con le analisi precedenti.

Le analisi del DNA mitocondriale trovano un parziale riscontro con quanto emerso per il cromosoma Y. Anche per la linea matrilineare è stato registrato un alto livello di diversità genetica. Infatti, la differenziazione delle popolazioni, calcolata tramite l'FST, è risultata statisticamente significativa (p-value <0,005), per entrambi gli isolati. Questo dato è confermato dall'assenza di aplotipi uguali condivisi fra le popolazioni.

Inoltre, l'alto livello di diversità genetica riscontrato fra gli isolati e il rispettivo campione di confronto indica che tutti gli individui mostrano sequenze differenti. La diversità genetica risulta superiore alla diversità genetica media della popolazione sarda (0.956) (Simoni, 2000), e questo potrebbe essere dovuto alla strategia di campionamento che ha massimizzato la variabilità genetica.

Per poter analizzare le relazioni filogenetiche degli aplotipi emersi dall'analisi dei dati delle popolazioni sono stati costruiti i network che mostrano la clusterizzazione degli aplotipi in base alla loro origine.

Il network di Benetutti e di Carloforte, mostra un'evidente forma *starlike*, con un cluster aplotipico centrale maggiormente rappresentato, da cui si dipartono tutte le ramificazioni. Tale distribuzione indica un'omogeneità della popolazione iniziale ed

un'origine monofiletica degli aplotipi. La forma del network filogenetico per la popolazione del nord-Sardegna evidenzia un'origine eterogenea della popolazione, essendo evidenti diversi nuclei da cui si dipartono i vari rami, mentre per la popolazione del Sulcis Iglesiente mostra una configurazione *starlike* la cui forma denota però un certo grado di variabilità interna.

Per quanto riguarda l'analisi degli aplogruppi è risultato un minor numero di aplogruppi diversi nei due isolati rispetto alle comunità di controllo. L'aplogruppo più frequente in tutti i campioni è l'H. Tale aplogruppo è il più rappresentato in Europa, con frequenze che si aggirano intorno al 40-50% (Quintans et al., 2004; Achilli et al., 2004). La frequenza riscontrata a Benetutti è risultata particolarmente alta (63%). Un valore simile è stato trovato in Barbagia (64,4%) (Morelli et al., 2000) e in Ogliastra (56,6%) (Fraumene et al., 2003; Fraumene et al., 2006) che come già detto fanno parte dell'“area arcaica” della Sardegna (Cappello et al., 1996). Le frequenze riscontrate nelle altre popolazioni oggetto di studio rientrano nell'ambito della variabilità della Sardegna (35% nord Sardegna, 33% Carloforte e 40% nel Sulcis). La frequenza di tale aplogruppo nell'Italia del nord è del 46,9% e in Tunisia del 26,5% (Achilli et al., 2004). La popolazione di Carloforte si caratterizza per la presenza dell'aplogruppo M1a1 (8%) non riscontrato nel campione di controllo e nelle altre popolazioni analizzate. L'aplogruppo M dovrebbe testimoniare la prima migrazione umana fuori dall'Africa verso le coste meridionali dell'Asia. La presenza di questo aplogruppo, infatti, si riscontra nelle popolazioni dell'Africa Orientale, l'Asia, le Americhe fino alla Melanesia (Caramelli, 2009). Uno studio recente fornisce la prova che l'aplogruppo M1, o il suo antenato, ha avuto un'origine asiatica e che la prima espansione dell'aplogruppo M1 in Africa si è verificata nella zona nord-ovest (Gonzalez et al., 2007). Secondo quanto riportato in tale studio l'aplogruppo M1 si trova con una

frequenza del 3,2% fra le popolazioni dell'area nord-ovest dell'Africa (3,9% fra i Berberi della Tunisia), dello 0,1% in Italia e dello 0,2% in Sardegna. È importante precisare che la frequenza riscontrata in Sardegna in tale lavoro fa riferimento a delle sequenze HVRI di mtDNA di carlofortini usate nello studio di Falchi e collaboratori (Falchi et al., 2006). Il lavoro di Morelli e collaboratori, invece, non ha rilevato la presenza di tale aplogruppo in Gallura e in Barbagia (Morelli et al., 2000).

Inoltre, Carloforte presenta un'alta frequenza dell'aplogruppo X₂b (16%), presente, con bassissima frequenza, nella popolazione del Sulcis (2%). L'aplogruppo X₂ è stato trovato con una frequenza del 2,9% nell'Italia continentale e Sicilia, dell'1,7 % in Tunisia (Reidla et al., 2003), 1,6% in Ogliastra (Fraumene et al., 2006). Una frequenza così alta a Carloforte può essere dovuta alla deriva genetica che ha agito per effetto del principio del fondatore. Nella zona del Sulcis Iglesiente è stata rilevata una frequenza dell'8% per l'aplogruppo U5b3. Una frequenza dell'11,1 % è stata trovata in un lavoro sull'Ogliastra (Fraumene et al., 2006). Tali frequenze sono più alte rispetto a quella riportata per la Sardegna nel lavoro di Pala e coll. (3,84%) (Pala et al., 2009). Secondo il suddetto studio la penisola italiana è la patria più probabile per l'aplogruppo U5b3, che si presenta con una frequenza molto bassa in tutta Europa (0,53% nell'Europa occidentale, 0,62 in Italia, 0,32% in Europa orientale). Sembrerebbe che l'espansione olocenica dell'aplogruppo U5b3 verso il Nord sia stata limitata dalle Alpi e si sia verificata solo lungo le coste del Mediterraneo, soprattutto verso la vicina Provenza. Da lì, circa 7.000-9.000 anni fa, un sottoclade di questo aplogruppo si trasferisce in Sardegna, probabilmente a causa del commercio di ossidiana che collegava le due regioni, lasciando una firma distintiva agli abitanti dell'isola (Pala et al., 2009).

Confrontando i dati delle sequenze del segmento ipervariabile I con dati analoghi di altre popolazioni dell'area mediterranea si sono trovati dei risultati interessanti (Varesi

et al.,2000; Krings et al.,1999; Fadhloui-Zid et al.,2004; Falchi et al.,2006; Côte-Real et al.,1996; Vona et al.,2001; Comas et al., 1996; Francalacci et al., 1996; Plaza et al.,2003; Rando et al.,1998; Stenico et al.,1996; Fraumene et al., 2006). Carloforte ha clusterizzato con le popolazioni del nord Africa fra cui la Tunisia mostrando una certa affinità con queste e una netta differenziazione sia con la attigua zona del Sulcis Iglesiente, sia verso tutte le altre zone della Sardegna e dell'Italia. Questo potrebbe significare che fra gli abitanti di Pegli trasferitisi nell'isola di Tabarka, dal 1542 al 1738, e gli individui autoctoni di tale zona della Tunisia vi siano state delle unioni che hanno apportato un contributo al pool genico della colonia pegliese da cui discendono gli individui fondatori di Carloforte.

E' degno di nota constatare che questo dato non era emerso attraverso l'analisi del cromosoma Y. Quindi, probabilmente attesta in diverso comportamento matrimoniale durante la permanenza a Tabarka. Con grande probabilità non vi è stato un libero scambio tra Pegliesi e Tabarkini, ma possiamo ipotizzare che alcune donne Tabarkine siano state inglobate nella comunità Pegliese, e avere così dato il loro contributo al pool genico dei Carlofortini.

Benetutti si è localizzata distante dalla popolazione del nord Sardegna, clusterizzando invece con la Trexenta. Non sappiamo se questo dato possa significare che tra le due popolazioni sia avvenuto uno scambio genico, oppure se il raggruppamento sia casuale e dovuto alla peculiarità genetica della popolazione della Trexenta, che un lavoro sui marcatori classici aveva dimostrato essere differenziata dalle restanti aree della Sardegna (Cappello et al., 1996).

È interessante notare come le varie zone della Sardegna non abbiano clusterizzato tutte assieme denotando un certo grado di eterogeneità interna, peraltro già riscontrata in diversi lavori (Gatti, 1990; Walter et al., 1991; Morelli et al., 1993; Moral et al., 1994;

Zavattari et al., 2000; Angius et al., 2001; Calò et al., 2008; Vona e Calò, 2006; Pistis et al., 2009).

Quando si sono prese in considerazione le sole popolazioni sarde sia Carloforte che Benetutti si sono disposte su rami separati dell'albero e distanti dalle restanti popolazioni della Sardegna sottolineando, ancora una volta, il loro alto grado di isolamento.

Inoltre, per quanto riguarda Carloforte, si sono messi a confronto i risultati del DNA mitocondriale ottenuti con due differenti metodi di campionamento: quello in base ai cognomi fondatori e quello secondo le genealogie, per valutare come possa influire la diversa strategia di campionamento. Da tale confronto è emerso un diverso quadro genetico. Dal numero di aplotipi diversi riscontrati si può notare che il campionamento secondo le genealogie ha evidenziato un certo grado di omogeneità all'interno della popolazione, mentre il campionamento secondo i cognomi fondatori ha massimizzato l'eterogeneità del campione. Dall'analisi degli aplogruppi è emersa una maggiore variabilità e un più grande numero di aplogruppi differenti nel campionamento secondo le genealogie, con un quadro generale simile a quello evidenziato per gli altri campioni della Sardegna. Sembra quindi che il campionamento in base ai cognomi fondatori abbia fatto emergere la struttura genetica della comunità al momento della costituzione del paese, mentre il campionamento secondo le genealogie abbia fatto emergere un certo grado di affinità con la popolazione sarda, dovuto sicuramente a un flusso genico con questa successivo alla fondazione. Un risultato molto interessante che ha riguardato i due differenti metodi di campionamento si è ottenuto attraverso gli alberi filogenetici. Nell'albero costruito con la popolazione carlofortina campionata secondo i cognomi fondatori Carloforte ha clusterizzato con le popolazioni del nord Africa e si presenta invece distante dall'Italia e dalle restanti comunità della Sardegna. Con il

campionamento secondo le genealogie, invece, Carloforte ha clusterizzato con l'Italia e si trova abbastanza vicina alle altre comunità sarde. Sembra quindi che la popolazione carlofortina si stia sempre più omogenando con la popolazione sarda e stia perdendo la componente genetica derivante dalla breve permanenza a Tabarka. Quando poi l'albero è costruito utilizzando i due distinti campionamenti è interessante osservare come questi non clusterizzino assieme ma si siano disposti l'uno lontano dall'altro come se si trattasse di due differenti popolazioni. Questo risultato riflette l'importanza che ha la scelta della strategia di campionamento nell'improntare correttamente una ricerca in base allo scopo prefissato.

Secondo quanto ottenuto con i marcatori uniparentali sembra di poter affermare che sia le barriere culturali e linguistiche che le barriere geografiche abbiano avuto un'influenza sull'isolamento e quindi nel determinare la struttura genetica di una popolazione, tuttavia le barriere linguistiche hanno avuto un maggior influsso nel differenziare dal punto di vista genetico la popolazione di Carloforte dalle comunità limitrofe. Inoltre, l'analisi comparativa dei risultati ottenuti con i marcatori ad eredità materna e paterna hanno fornito solo parzialmente informazioni concordanti. Tra i risultati più interessanti vi è una maggiore omogeneità genetica per il mtDNA tra le popolazioni sarde.

Questo potrebbe essere dovuto alla maggiore mobilità della donna, che ha contribuito a far diminuire la variabilità genetica a livello del DNA mitocondriale tra i vari paesi della Sardegna. Infatti, solitamente nei matrimoni esogamici è la donna che dopo il matrimonio si trasferisce nel paese del marito, contribuendo a diminuire la variabilità mitocondriale tra i paesi (Sanna et al., 2004).

Inoltre, avendo realizzato il campionamento per cognomi fondatori, ed essendo il cognome ereditato per via patrilineare come il cromosoma Y, possiamo essere certi che

i cromosomi Y selezionati siano i reali discendenti dei fondatori, non così per il DNA mitocondriale, che può invece risentire di nuovi flussi genici avvenuti nel corso dei secoli. Quindi il cromosoma Y dovrebbe aver ricreato quella che era la struttura genetica alla nascita del paese mentre il mtDNA potrebbe essere lo specchio di una storia più recente della popolazione.

Bibliografia

- ✓ Achilli A., Rengo C., Magri C., Battaglia V., Olivieri A., Scozzari R., Cruciani F., Zeviani M., Briem E., Carelli V., Moral P., Dugoujon J.M., Roostalu U., Loogvali E.L., Kivisild T., Bandelt H.J., Richards M., Villems R., Santachiara-Benerecetti A.S., Semino O., and Torroni A., (2004) *The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool.* Am J Hum Genet 75:910-918.
- ✓ Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG. *Sequence and organization of the human mitochondrial genome.* Nature., 1981; **290**:457–465. doi: 10.1038/290457a0.
- ✓ Angioni D., Loi S. e Puggioni G., 1997. *La popolazione dei Comuni Sardi dal 1688 al 1991*; CUEC, Cagliari.
- ✓ Angius A., Melis P. M., Morelli L., Petretto E., Casu G., Maestrale G. B., Fraumene C., Bebbere D., Forabosco P., Pirastu M., 2001. *Archival, demographic and genetic studies define a Sardinian sub-isolate as a suitable model for mapping complex traits.* Hum. Genet. 109: 198-209;
- ✓ Bernstein F., (1930): *Fortgesetzte Untersuchungen aus der Theorie der Blutgruppen.* Zeitschrift für Induktionem Abstammungs-und Vererbungslehre. **56**, 233-273.
- ✓ Bosch E., Calafell F., Comas D., Oefner P.J., Underhill P.A., Bertranpetit J. (2001) *High-resolution analysis of human Y chromosome variation shows a sharp discontinuity*

- and limited gene flow between Northwestern Africa and Iberian Peninsula. Am. J. Hum. Genet. 68:1019–1029.*
- ✓ Bourgain C., Genin E., Holopainen P., Mustalahti K., Maki M., Partanen J., Clerget-Darpoux F. *Use of closely related affected individuals for the genetic study of complex diseases in founder populations. (2001). Am. J. Hum Genet. 68:154-159.*
 - ✓ Calò C.M., Vona G. (1994). *Carloforte: evoluzione della struttura matrimoniale. Rendiconti seminario Facoltà di Scienze Università degli Studi di Cagliari, Vol. 64 Fasc.*
 - ✓ Calò C.M., Vona G.(2005). *Genetic and demographical analysis of San Pietro island (Sardinia, Italy). Human Evolution, Vol.20 - n. 2-3: 99-106.*
 - ✓ Calò C. M., Melis A., Vona G., Piras I. S. (2008). *Sardinian population (Italy) a Genetic Review. Int. J. Modern Anthropology. 1: 39-64.*
 - ✓ Capelli C., Redhead N., Abernethy J.K., Gratrix F., Wilson J.F., Moen T., Hervig T, Richards M., Stumpf M.P., Underhill P.A., Bradshaw P., Shaha A., Thomas M.G., Bradman N., Goldstein D.B. (2003). *A Y chromosome census of the British Isles. Curr Biol. 13:979–984.*
 - ✓ Capelli C., Redhead N., Romano V., Cali F., Lefranc, Delague V., Megarbane A., Felice A. E., Pascali V. L., Neophytou P. I., Poulli Z., G.Novelletto A., Malaspina P., Terrenato L., Berebbi A., Fellous M., Thomas M. G. and Goldstein D. B. (2005). *Population Structure in the Mediterranean Basin: A Y Chromosome Perspective Ann. Hum. Genet. doi: 10.1111/j.1529-8817.2005.00224.x.*
 - ✓ Cappello N., Rendine S., Griffò R., Mamei G.E., Succa V., Vona G., Piazza A. (1996). *Genetic analysis of Sardinia: I.data on 12 polymorphism in 21 linguistic domains. Ann. Hum. Genet. 60: 125-41;*

- ✓ Caramelli D., Vernesi C., Sanna S., Sampietro L., Lari M., Castri L., Vona G., Floris R., Francalacci P., Tykot R., Casoli A., Bertranpetit J., Lalueza-Fox C., Bertorelle G., Barbujani G. (2007). *Genetic variation in prehistoric Sardinia*. Hum. Genet. 122(3-4): 327-336.
- ✓ Caramelli D., (2009). *Antropologia molecolare*. Manuale di base. Firenze University Press.
- ✓ Cavalli-Sforza L.L., Piazza A., Menozzi P. (1994). *The history and geography of human genes*. Princeton University Press.
- ✓ Cinnioglu C., King R., Kivisild T., Kalfoglu E., Atasoy S., Cavalleri G.L., Lillie A.S., Roseman C.C., Lin A.A., Prince K., Oefner P.J., Shen P., Semino O., Cavalli-Sforza L.L., Underhill P.A. (2004) *Excavating Y-Chromosome Haplotype Strata In Anatolia*. Hum. Genet. 114:127-148.
- ✓ Colonna V., Nutile T., Astore M., Guardiola O., Antoniol G., Ciullo M., Persico M.G. (2007). *Campora: a young genetic isolate in South Italy*. Hum.Hered. 64(2): 123-135.
- ✓ Comas D., Calafell F., Mateu E., Perez-Lezaun A., Bertranpetit J. (1996). *Geographic variation in human mitochondrial DNA control region sequence: the population history of Turkey and its relationship to the European populations*. Mol. Biol. Evol. 13:1067-1077.
- ✓ Contu D., Morelli L., Santoni F., Foster J.W., Francalacci P., Cucca F. (2008). *Y.Chromosome based evidence for pre-neolithic origin of the genetically homogeneous but diverse Sardinian population: inference for association scan*. PLoS ONE. Issue 1, e1430.
- ✓ Côrte-Real H.B., Macaulay V.A., Richards M.B., Hariti G., Issad M.S., Cambon-Thomsen A., Papiha S., Bertranpetit J., Sykes B.C. (1996). *Genetic diversity in the*

- Iberian Peninsula determined from mitochondrial sequence analysis. Am. Hum. Genet.* 60:331-350.
- ✓ Cosseddu G.G., Fenu A., Floris G., Lucia G., Vona G. (1979). *Esame antropometrico di due etnie liguri in Sardegna. Antropologia contemporanea*, 2: 325-328.
 - ✓ Di Gaetano C, Cerutti N, Crobu F, Robino C, Inturri S, Gino S, Guarrera S, Underhill PA, King RJ, Romano V, Cali F, Gasparini M, Matullo G, Salerno A, Torre C, Piazza A. (2009). *Differential Greek and northern African migrations to Sicily are supported by genetic evidence from the Y chromosome. Eur. J. Hum. Genet.* 17(1):91-9.
 - ✓ Fadhlaoui-Zid K., Plaza S., Calafell F., Ben Amor M., Comas D., Bennamar Elgaaied A. (2004). *Mitochondrial DNA heterogeneity in Tunisian Berbers. Ann. Hum. Genet.* 68:222-233.
 - ✓ Falchi A., Giovannoni L., Calò C.M., Piras S.I., Moral P., Paoli G., Vona G., Varesi L. (2006). *Genetic history of some western Mediterranean human isolates through mtDNA HVRI polymorphisms. J Hum Genet* 51:9-14.
 - ✓ Ferraro G. (1989). *Da Tabarka a San Pietro. Nasce Carloforte. Tipografia Artigiana Musanti Editrice, Cagliari.*
 - ✓ Ferri G., Alù M., Corradini B., Radheshi E., Beduschi G. (2009). *Slow and fast evolving markers typing in Modena males (North Italy).Forensic science international. Gen.* 3(2):e31-3.
 - ✓ Flores C., Maca-Meyer N., Pèrez J.A., González A.M., Larruga J.M. and Cabrera V.M. (2003). *A Predominant European Ancestry of Paternal Lineages from Canary Islanders.* 67,138-152.
 - ✓ Floris G., Vona G. (1980). *Isonimie marital et coefficient de parenté entre six communes de l'île de Sardaigne. L'Anthropologie*, 84, 2, 300-306.

- ✓ Floris G., Sanna E. (2000). *On the biological structure of the Sardinian people*. In: *Biennial Books of EAA*. Volume 1. Human population genetics in Europe (Susanne C. and Bodzsar E.B. Eds) pp. 85-106. European Anthropological Association. Eötvös University Press, Budapest.
- ✓ Francalacci P., Bertranpetit J., Calafell F., Underhill P.A. (1996). *Sequence diversity of the control region of mitochondrial DNA in Tuscany and its implications for the peopling of Europe*. *Am. J. Phys. Anthropol.* 100:443-460.
- ✓ Francalacci P., Morelli L., Underhill P.A., Lillie A.S., Passarino G., Useli A., Madeddu R., Paoli G., Tofanelli S., Calò C.M., Ghiani M.E., Varesi L., Memmi M., Vona G., Lin A.A., Oefner P., Cavalli-Sforza L.L. (2003) *Peopling of Three Mediterranean Islands (Corsica, Sardinia, and Sicily) Inferred by Y-Chromosome Biallelic Variability*. *Am. J. Phys. Ant.* 121:270–279.
- ✓ Fraumene C., Petretto E., Angius A., Pirastu M. (2003). *Striking differentiation of sub-populations within a genetically homogeneous isolate (Ogliastra) in Sardinia as revealed by mtDNA analysis*. *Hum Genet.* 114(1):1–10.
- ✓ Fraumene C., Belle E.M.S., Castri` L., Sanna S., Mancosu G., Cosso M., Marras F., Barbujani G., Pirastu M. and Angius A. (2006). *High resolution analysis and phylogenetic network construction using complete mtDNA sequences in sardinian genetic isolates*. *Mol. Biol. Evol.* 23(11):2101–2111.
- ✓ Frigi S., Pereira F., Pereira L., Yacoubi B., Gusmao L., Alves C., Khodjet el Khil H., Cherni L., Amorim A., El Gaaied A. (2006). *Data for Y-chromosome haplotypes defined by 17 STRs (AmpFLSTR1 Yfiler™) in two Tunisian Berber communities*. *Forensic Science International* 160: 80–83.

- ✓ Gatti A.M. (1990). *L'area di scambi matrimoniali in Sardegna tra il XVII e XX secolo*. In AA.VV.: *Famiglia e matrimonio nella Società Sarda Tradizionale*. Nuoro: La tarantola Edizioni, pp 170-191.
- ✓ Gelehrter T.D., Collins F.S., Ginsburg D. (2002). *Genetica medica*. Masson. Milano.
- ✓ Ghiani M.E., Mameli A., Piras G., Berti A., Calò C.M. and Vona G. (2009). *Population data for Y-Chromosome haplotypes defined by AmpFISTR YFiler PCR Amplification Kit in north Sardinia (Italy)*. Coll. Antropol. 33, 2: 643-651.
- ✓ González A.M., Larruga J.M., Abu-Amero K.K., Shi Y., Pestano J. and Cabrera V.M. (2007). *Mitochondrial lineage M1 traces an early human backflow to Africa*. *BMC Genomics*, 8:223.
- ✓ Hammer M.F., Horai S. (1995). *Y chromosomal DNA variation and the peopling of Japan*. *Am. J. Hum. Genet.* 56: 951-962.
- ✓ Heath S., Robledo R., Geggs W., Rinaldi A., Siniscalco M. (2001). *A novel approach to search for identity by descent in small samples of patient and controls from the same mendelian breeding unit: a pilot study on myopia*. *Human Hereditary*, 52: 183-190.
- ✓ Helgason A., Sigurdardottir S., Gulcher J.R., Ward R., Stefansson K. (2000). *mtDNA and the Origin of the Icelanders: Deciphering Signals of Recent Population History*. *Am. J. Hum. Genet.* 66:999–1016.
- ✓ Heutink P., Oostra B.A. (2002). *Gene finding in genetically isolated populations*. *Hum. Mol. Gen.* 11: 2507-2515.
- ✓ Higgins DG, Bleasby AJ, Fuchs R. (1992). *CLUSTAL V: improved software for multiple sequence alignment*. *Comput. Appl. Biosci.*, 8:189-191.
- ✓ Jobling M.A, Tyler-Smith C. (2003). *The human y chromosome: an evolutionary marker comes of age*. *Nat. Rev. Genet.* 4:598-612.

- ✓ Karafet T.M., Mendez F. L., Meilerman M. B., Underhill P. A., Zegura S. L., Hammer M.F. (2008). *New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree*. Genome Res.
- ✓ Kovatsi L., Saunier J.L., Irwin J.A. (2009). *Population genetics of Y-chromosome STRs in a population of Northern Greeks*. Forensic Science International: Genetics 4: e21–e22.
- ✓ Krings M., Salem A.E., Bauer K., Geisert H., Malek A.K., Chaix L., Simon C., Welsby D., Di Rienzo A., Utermann G., Sajantila A., Paabo S., Stoneking M. (1999). *mtDNA analysis of Nile River Valley populations: a genetic corridor or a barrier to migration?*. Am J Hum Genet. 64:1166-1176.
- ✓ Lòpez-Parra A.M., Gusmao L., Tavares L., Baeza C., Amorim A., Mesa M.S., Prata M.J. and Arroyo-Pardo E. (2009). *In search of the Pre- and Post-Neolithic Genetic Substrates in Iberia: Evidence from Y-Chromosome in Pyrenean Populations*. Annals of Human Genetics. 73, 42-53.
- ✓ Lucia G. e Vona G. (1982). *Caratteristiche morfometriche della popolazione adulta attuale di Carloforte (Sardegna)*. Rend. Sem. Fac. Scienze Univ. Cagliari, 52, 1: 199-206.
- ✓ Luxoro E. (1977). *Tabarka e Tabarkini*. Edizioni della Torre, Cagliari.
- ✓ Mange A.P., Johansen Mange E. (1993). *Genetica e l'uomo*. Zanichelli. Bologna.
- ✓ Martin P., Garcia-Hirschfeld J., Garcia O., Gusmao L., Garcia P., Albarràn C., Sancho M., Alonso A. (2004). *A Spanish population study of 17 Y-chromosome STR loci*. Forensic Science International 139. 231–235.
- ✓ Moral P., Marogna G., Salis M., Succa V., Vona G. (1994). *Genetic data on Alghero population (Sardinia): contrast between biological and cultural evidence*. Am. J. Phys. Anthropol. 93(4): 441-453.

- ✓ Morelli L., Vona G. (1993). *Studio genetico della popolazione di Villanova Monteleone (Sardegna). Nota I. Struttura matrimoniale.* Rend. Semin. Fac. Sci. Univ. Cagliari. 63: 75-86.
- ✓ Morelli L., Grosso M.G., Vona G., Varesi L., Torroni A., Francalacci P. (2000). *Frequency distribution of mitochondrial DNA haplogroups in Corsica and Sardinia.* Hum Biol. 72(4):585-95.
- ✓ Moroni A., Anelli A., Anghinetti W., Lucchetti E., Rossi O., Siri E. (1972). *La consanguineità umana nell'isola di Sardegna dal XVIII al XX secolo.* Ateneo Parmense , 8, suppl. 1, 69-92.
- ✓ Onofri V., Alessandrini F., Turchi C., Fraternali B., Buscemi L., Pesaresi M., Tagliabracci A. (2007). *Y-chromosome genetic structure in sub-Apennine populations of Central Italy by SNP and STR analysis.* Int J Legal Med. 121:234–237.
- ✓ Oppes T. (1990). *Il Goceano*, EdiSar, Cagliari.
- ✓ Pala M., Achilli A., Olivieri A., Kashani B.H., Perego U.A., Sanna D., Metspalu E., Tambets K., Tamm E., Accetturo M., Carossa V., Lancioni H., Panara F., Zimmermann B., Huber G., Al-Zahery N., Brisighelli F., Woodward S.R., Francalacci P., Parson W., Salas A., Behar D.M., Villemes R., Semino O., Bandelt H-J. and Torroni A. (2009). *Mitochondrial Haplogroup U5b3: A Distant Echo of the Epipaleolithic in Italy and the Legacy of the Early Sardinians.* The American Journal of Human Genetics. 84, 814–821.
- ✓ Palet L., Coudray C., Galey C., Keyser C., Melhaoui M., Gagnor C., Sabatier M., Dugoujon J-M. (2010). *Y-STR genetic diversity in Moroccans from the Figuig oasis.* Forensic Science International: Genetics. 4 e139–e141.
- ✓ Passarino G., Underhill A.P., Cavalli-Sforza L., Semino O., Pes M.G., Carru C., Ferrucci L., Bonate M., Franceschi C., Deina L., Baggio G., De Benedictis G. (2001). *Y*

- chromosome binary markers to study the high prevalence of males in Sardinian centenarians and the genetic structure of the Sardinian population.* Hum. Hered. 52:136–139
- ✓ Pettener D. (1995). *Aspetti biodemografici nello studio delle popolazioni umane.* In Antropologia, Evoluzione, uomo, ambiente (di F.Facchini), 492-500. UTET, Torino.
 - ✓ Pistis G., Piras I., Pirastu N., Persico N., Sassu A., Picciau A., Prodi D., Fraumene C., Mocchi E., Manias M. T., Atzeni R., Cosso M., Pirastu M., Angius A. (2009). *High differentiation among eight villages in a secluded area of Sardinia revealed by genome-wide high density SNPs analysis.* PLoS One. 4(2): e4654.
 - ✓ Plaza S., Calafell F., Helal A., Bouzerna N., Lefranc G., Bertranpetit J., Comas D. (2003). *Joining the pillars of Hercules: mtDNA sequences show multidirectional gene flow in the western Mediterranean.* Ann. Hum. Genet. 67:312-328.
 - ✓ Pokupcic' K., Cukrov S., Klaric' I.M., Salihovic' M.P., Barac' Lauc L., Blazanovic' A., Janic'ijevic' B. (2008). *Y-STR genetic diversity of Croatian (Bayash) Roma.* Forensic Science International: Genetics 2. e11–e13.
 - ✓ Pritchard JK, Stephens M., Donnelly P. (2000). *Inference of population structure using multilocus genotype data.* Genetics 155: 945-959.
 - ✓ Quintans B., Alvarez-Iglesias V., Salas.A., Phillips C., Lareu M.V., Carracedo A. (2004). *Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaPshot minisequencing.* Forensic Science International 140: 251–257.
 - ✓ Rando J.C., Pinto F., Gonzalez A.M., Hernandez M., Larruga J.M., Cabrera V.M., Bandelt H.J. (1998). *Mitochondrial DNA analysis of northwest African populations reveal genetic exchanges with European, near- eastern, and sub- Saharan populations.* Ann. Hum. Genet. 62:531-550.

- ✓ Reidla M., Kivisild T., Metspalu E., Kaldma K., Tambets K., Tolk H.V., Parik J., Loogvali E.L., Derenko M., Malyarchuk B., Bermisheva M., Zhadanov S., Pennarun E., Gubina M., Golubenko M., Damba L., Fedorova S., Gusar V., Grechanina E., Mikerezi I., Moisan J.P., Chaventre' A., Khusnutdinova E. , Osipova L., Stepanov V., Voevoda M., Achilli A., Rengo C., Rickards O., De Stefano G.F., Papiha S., Beckman L., Janicijevic B., Rudan P., Anagnou N., Michalodimitrakis E., Koziel S., Usanga E., Geberhiwot T., Herrnstadt C., Howell N., Torroni A., Villems R. (2003). *Origin and Diffusion of mtDNA Haplogroup X*. Am. J. Hum. Genet. 73:1178–1190.
- ✓ Reynolds J., Weir B.S., Cockerham C.C. (1983). *Estimation of the coancestry coefficient: basis for a short-term genetic distance*. Genetics, 105:767–779.
- ✓ Rootsi S., Magri C., Kivisild T., Benuzzi G., Help H., Bermisheva M., Kutuev I., Barac' L., Peric'Ic' M., Balanovsky O., Pshenichnov A., Dion D., Grobei M., Zhivotovsky L.A., Battaglia V., Achilli A., Al-Zahery N., Parik J., King R., Cinniog'Lu C., Khusnutdinova E., Rudan P., Balanovska E., Scheffrahn W., Simonescu M., Brehm A., Goncalves R., Rosa A., Moisan J. P., Chaventre A., Ferak V., Furedi S., Oefner P.J., Shen P., Beckman L., Mikerezi I., Terzic R., Primorac D., Cambon-Thomsen A., Krumina A., Torroni A., Underhill P.A., Santachiara-Benerecetti A.S., Villems R., Semino O. (2004). *Phylogeography of Y-Chromosome Haplogroup I Reveals Distinct Domains of Prehistoric Gene Flow in Europe*. Am. J. Hum. Genet. 75:128–137.
- ✓ Saitou, N., Nei, M. (1987). *The neighbor-joining method: a new method for reconstructing phylogenetic trees*. Molecular and biological evolution, 4: 406-425.
- ✓ Sanna D., Pala M., Cossu P., Dedola G., Melis S., Fresu G., Morelli L., Obinu D., Tonolo G., Secchi G., Triunfo R., Lorenz J.G., Scheinfeldt L., Torroni A., Robledo R., Francalacci P. (2011). *Mendelian Breeding Units versus Standard Sampling Strategies:*

mitochondrial DNA variation in southwestern Sardinia. Genetics and molecular Biology. In press.

- ✓ Sanna E., Iovine M.C., e Floris G. (2004). *Evolution of marital structure in 20 Sardinian villages from 1800 to 1974*. *Anthrop. Anz.* 62: 2,169-184.
- ✓ Sanna E., Melis A., Iovine M.C. (2006). *Endogamy and exogamy in an alloglot linguistic community of Sardinia (Carloforte) from 1738 to 2001*. *Research Signpost* (Ed.) Trivandrum, India in *Human Genetic Isolates*, 121-130.
- ✓ Scozzari R., Cruciani F., Pangrazio A., Santolamazza P., Vona G., Moral P., Latini V., Varesi L., Memmi M.M., Romano V., De Leo G., Gennarelli M., Jaruzelska J., Villems R., Parik J., Macaulay V., Torroni A. (2001) *Human Y-chromosome variation in the western Mediterranean area: implications for the peopling of the region*. *Hum. Immunol.* 62:871–884.
- ✓ Semino O., Passarino G., Oefner P.J., Lin A.A., Arbuzova S., Beckman L.E., De Benedictis G., Francalacci P., Kouvatsi A., Limborska S., Marcikiae M., Mika A., Mika B., Primorac D., Santachiara-Benerecetti A.S., Cavalli-Sforza L.L., Underhill P.A. (2000). *The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective*. *Science*; 290(5494):1155-9.
- ✓ Simoni, L., Calafell, F., Pettener, D., Bertranpetit, J., Barbujani, G. (2000). *Geographic Patterns of mtDNA Diversity in Europe*. *The American Journal of Human Genetics*, Vol. 66:1, 262-278.
- ✓ Slatkin M. (1995). *A measure of population subdivision based on microsatellite allele frequencies*. *Genetics*, 139:457-462.
- ✓ Skaletsky H., Kuroda-Kawaguchi T., Minx P. J., Cordum H. S., Hillier L.D., Brown L. G., Repping S., Pyntikova T., Ali J., Bieri T., Chinwalla A., Delehaunty A., Delehaunty K., Du H., Fewell G., Fulton L., Fulton R., Graves T., Hou S.F., Latrielle P., Leonard

- S., Mardis E., Maupin R., Mcpherson J., Miner T., Nash W., Nguyen C., Ozersky P., Pepin K., Rock S., Rohlfing T., Scott K., Brian Schultz, Cindy Strong, Aye Tin-Wollam, Shiaw-Pyng Yang, Robert H. Waterston, Wilson R. K., Rozen S., Page D. C. (2003). *The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes*, Nature 423, 825-837.
- ✓ Stenico M., Nigro L., Bertorelle G., Calafell F., Capitanio M., Corrain C., Barbujani G. (1996). *High mitochondrial sequence diversity in linguistic isolates of the Alps*. Am. J. Hum. Genet. 59:1363-1375.
 - ✓ Turrina S., Atzei R., De Leo D. (2006). *Y-chromosomal STR haplotypes in a Northeast Italian population sample using 17plex loci PCR assay*. Int J Legal Med. 120: 56–59.
 - ✓ Underhill, P.A., Shen, P., Lin, A., Jin A.A., Passarino G., Yang W.H., Kauffman E., Bonn -Tamir B., Bertranpetit J., Francalacci P., Ibrahim M., Jenkins T., Kidd J.R., Mehdi S.Q., Seielstad M., Wells S., Piazza A., Davis R.W., Feldman M.F., Cavalli-Sforza L.L., Oefner P.J. (2000). *Y Chromosome Sequence Variation And The History Of Human Populations*. Nat. Genet. 26,358-361.
 - ✓ Underhill PA, Passarino G, Lin AA, Shen P, Mirazon Lahr M, Foley RA, Oefner PJ, Cavalli-Sforza LL. (2001). *The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations*. Ann Hum Genet. 65(Pt 1):43-62.
 - ✓ Vallebona G. (1974). *Carloforte. Storia di una colonizzazione*. Edizioni della Torre, Cagliari.
 - ✓ Varesi L., Memmi M., Cristofari M.C., Mameli G.E., Cal  C.M., Vona G. (2000) *Mitochondrial control region sequence variation in Corsican population (France)*. Am J Hum Biol. 12:339-351.
 - ✓ Vona G., Floris-Masala R., Mameli G.E., Succa V. (1992). *Red cell and serum protein polymorphism in Sardinia*. Inter. J. Phys. Anthr. 7:25-33.

- ✓ Vona G., Mameli G.E., Succa V., Lucia G. (1994). *Carloforte: un'etnia ligure in Sardegna. Dati sulla struttura genetica*. *Antropologia contemporanea*, 17: 103-111.
- ✓ Vona G., Calò C.M., Esteban E., Moral P. (1996). *Genetics, geography and Culture: the population of S.Pietro Island (Sardinia, Italy)*. *American Journal of physical anthropology*, 100: 461-471.
- ✓ Vona G., Francalacci P., Paoli G., Latini V., Salis M. (1996). *Study of the matrimonial structure of the population of central Sardinia (Italy)*. *Anthrop. Anz.* 54(4): 317-29.
- ✓ Vona G., Ghiani M.E., Calò C.M., Memmi M., Varesi L. (2001). *Mitochondrial DNA sequence analysis in Sicily*. *Am. J. Hum Biol.* 13: 576-589.
- ✓ Vona G., Calò C.M. (2006). *History of Sardinian population (Italy, Western Mediterranean) as inferred from genetic analysis*. In *Human Genetic Isolates, Research Signpost, India*, pp 1-28.
- ✓ Walter H., Matsumoto H., De Stefano G.F. (1991). *Gm and Km allotypes in four Sardinian population samples*. *Am. J. Phys. Anthropol.* 86:45-50.
- ✓ Workman P.L., Lucarelli P., Agostino R., Scarabino R., Scacchi R., Carapella E. (1975). *Genetic differentiation among Sardinian villages*. *Am. J. Phys. Anthropol.* 43: 165-176.
- ✓ Y Chromosome Consortium (2002). *A nomenclature system for the tree of human Y-chromosomal binary haplogroups*. *Genome Research.* 12: 339-348.
- ✓ Zavattari P., Deidda E., Whalen M., Lampis R., Mulargia A., Loddo M., Eaves I., Mastio G., Todd J. A., Cucca F. (2000). *Major factors influencing linkage disequilibrium by analysis of different chromosome regions in distinct populations: demography, chromosome recombination frequency and selection*. *Hum. Mol. Gen.* 9: 2947-2957.

- ✓ Zei G., Guglielmino C.R., Siri E., Moroni A., Cavalli-Sforza L.L. (1983). *Surnames as neutral alleles: observations in Sardinia*. Human Biology, 55, 357-408.
- ✓ Zei G., Lisa A., Fiorani O., Magri C., Quintana-Murci L., Semino O., Santachiara-Benerecetti A S. (2003). *From surnames to the history of Y chromosomes: the Sardinian population as a paradigm*. Eur. J. Hum. Genet. 11, 802–807.
- ✓ <http://www.isogg.org/tree/index10.html>.
- ✓ <http://www.hprg.com/hapest5/index.html#>.
- ✓ <http://www.hprg.com/hapest5/>.
- ✓ <http://web.hku.hk/~xxia/software/software.htm> Xia, X. 2000.
- ✓ <http://nnhgtool.nationalgeographic.com/classify/index.html>.
- ✓ <http://www.phylotree.org/tree/main.htm>.
- ✓ http://pritch.bsd.uchicago.edu/software/structure2_2.html.
- ✓ <http://www.cstl.nist.gov/strbase/ystrpos1.htm>.
- ✓ <http://www.mitomap.org/MITOMAP/HumanMitoSeq>.
- ✓ <http://anthro.unige.ch/software/arlequin/>.

Ringraziamenti

Un sentito ringraziamento:

Alla Dott.ssa Carla Maria Calò per i preziosi insegnamenti impartiti nei tre anni di dottorato e per il suo continuo supporto, non solo didattico e professionale, ma anche umano in tutte le fasi di questo lavoro;

Al Prof. Giuseppe Vona per aver costantemente messo a mia disposizione la sua esperienza e le sue competenze scientifiche e per avermi guidato con i suoi suggerimenti durante tutto il percorso formativo;

Al Prof. Emanuele Sanna per le sue consulenze in materia demografica e per i tanti preziosi consigli;

Al Prof. John Mitchell per l'aiuto fornito nel lavoro di analisi dei campioni, indispensabile per la realizzazione di questo lavoro;

Al Prof. Renato Robledo per le sue consulenze scientifiche;

Alla Prof.ssa Donata Luiselli per la disponibilità offerta con affettuosa amicizia per le analisi del DNA mitocondriale;

Al Dott. Ignazio Stefano Piras per i consigli forniti per l'elaborazione statistica dei dati;

Al Reparto Investigazioni Scientifiche dei Carabinieri di Cagliari nelle persone del Comandante Giovanni Delogu e del Maggiore Alessandro Mameli per l'aiuto nelle analisi del cromosoma Y;

Alle Amministrazioni Comunali per la disponibilità mostrata nel sensibilizzare le comunità affinché fornissero il loro apporto alla realizzazione di questo studio ed a tutte le persone che con la loro partecipazione al campionamento hanno reso possibile il mio lavoro di tesi.

Un ringraziamento particolare all'Assessore Daniele Arca, al Sig. Efsio Giovanni Cuccurazzu dell'Ufficio Anagrafe ed a Don Michele Satta del Comune di Benetutti per l'accoglienza e per l'ausilio fornito;