



Unione Europea



REGIONE CAMPANIA



PUBBL. Campania
FSE 2007-2013

La tua
Campania
cresce in
Europa



UNIVERSITÀ DEGLI STUDI DI SALERNO



UNIVERSITÀ DEGLI STUDI DI SALERNO
Dipartimento di Farmacia

Dottorato di ricerca
in Biologia dei Sistemi

Ciclo XII — Anno di discussione 2014

Coordinatore: Chiar.mo Prof. Antonietta Leone

***Studio della biodiversità mediante marcatori
molecolari di popolazioni naturali e di ecotipi
orticoli della regione Campania***

settore scientifico disciplinare di afferenza: BIO/01

Dottorando

Dott. Francesco Guarino

Tutore

Chiar.mo Prof. Stefano Castiglione

Indice

Capitolo I

Introduzione	1
1.1 Biodiversità	1
1.2 Marcatori molecolari	4
1.2.1 Simple Sequence Repeat (SSR).....	4
1.2.2 Amplified Fragment Length Polymorphism (AFLP).....	5
1.3 Epigenetica	7
1.4 Biodiversità agraria	11
1.5 Biodiversità forestale	13
1.6 Salicaceae	16
1.7 Conservazione della biodiversità	17
1.7.1 Conservazione e mantenimento della biodiversità agraria.....	19
1.7.2 Conservazione e mantenimento della biodiversità forestale.....	21

Capitolo II

Materiali e Metodi	22
2.1 Campionamenti	22
2.1.1 Raccolta del materiale vegetale di ecotipi orticoli.....	22
2.1.2 Raccolta del materiale vegetale di pioppo bianco.....	26
2.2 Estrazione del DNA	28
2.3 ANALISI MOLECORARI	29
2.3.1 Analisi molecolari sugli ecotipi orticoli campani.....	29
2.3.2 MSAP.....	34
2.4 Analisi bioinformatica	37
2.4.1 Analisi dei marcatori SSR e MSAP.....	37
2.4.3 Analisi della Varianza Molecolare (AMOVA).....	40
2.4.5 Analisi marcatori microsatelliti degli ecotipi orticoli campani.....	40
2.4.6 Analisi bioinformatica per popolazioni naturali di <i>P. alba</i>	41

Capitolo III

Risultati	44
3.1 Ecotipi orticoli Campani – Simple Sequence Repeat (SSR)	44
3.1.1 Aglio.....	44
3.1.2 Cipolla.....	48
3.1.3 Mais.....	54
3.1.4 Scarola.....	59
3.1.5 Zucchino.....	61
3.2 Popolazioni naturali di <i>P. alba</i>	67
3.2.1 <i>Populus alba</i> – Methylation Sensitive Amplified Polymorphism.....	69

Capitolo IV

Discussione87

Ringraziamenti97

Bibliografia98

Capitolo I

Introduzione

1.1 Biodiversità

La presente tesi di dottorato è incentrata sullo studio della biodiversità genetica di diversi ecotipi orticoli della regione Campania inseriti nel programma di salvaguardia del patrimonio orticolo regionale, finanziato attraverso il progetto AGRIGENET, e di popolazioni naturali di *Populus alba* L.. Il progetto di ricerca ha inoltre indagato lo stato di metilazione del DNA estratto dalle foglie di pioppo bianco sardo allo scopo di individuare eventuali alterazioni epigenetiche indotte dall'ambiente (Lukens and Zhan 2007). Prima di approfondire i temi citati occorre introdurre alcuni concetti chiave che contestualizzino la ricerca. Il concetto di biodiversità è più antico del termine che lo definisce. Infatti, con il sostantivo biodiversità, usato per la prima volta nella conferenza internazionale di Rio de Janeiro del 1992, si intende l'insieme di tutte le forme viventi geneticamente diverse. Darwin (1872) nella sua teoria dell'evoluzione considerava gli esseri viventi gli uni diversi dagli altri, e, proprio grazie a queste differenze, alcuni individui si adattavano meglio alle alterazioni degli habitat e alle mutazioni dell'ambiente intervenute nel corso degli anni, dei secoli, dei millenni. L'importanza di preservare la biodiversità si traduce non solo nella tutela dei singoli soggetti, ma nel preservare le relazioni che i diversi individui di specie differenti possono instaurare in un ecosistema. Grazie all'attività di numerosi ricercatori è stato stimato che, ai giorni nostri, quasi la totalità delle specie vegetali sono minacciate di estinzione con un tasso annuale prossimo allo 0,5 % sul totale delle specie esistenti. Ciò sarebbe causato da numerosi fattori, tra i quali, quello antropico ha il ruolo indubbiamente preponderante; l'uomo, infatti, attraverso alcune tecniche colturali (quali ad esempio la riduzione del numero di specie agricole coltivate, le monocolture, l'uso delle serre, etc.), oppure attraverso fenomeni di urbanizzazione incontrollata, o ancora a causa della sottrazione di territorio agricolo a favore di attività industriali e/o commerciali sovradimensionate rispetto alla richiesta, contribuisce alla riduzione della ricchezza di specie vegetali presenti nel mondo interrompendo il flusso genico e creando delle "isole" di speciazione. In questo modo il patrimonio genetico risulta depauperato rendendo le piante meno capaci di adattarsi alle alterazioni ambientali sia a livello globale che locale.

Dal 1992, grazie alla sottoscrizione da parte di oltre 150 nazioni della convenzione di Rio de Janeiro, gli stati si sono impegnati a intraprendere strategie e politiche atte a studiare, conservare e tutelare la biodiversità. A questo scopo sono stati indicati tre livelli di diversità biologica: ecosistema, specie e genetico.

Prima di approfondire il concetto di biodiversità a livello di ecosistema è doveroso ricordare il significato biologico del termine stesso: coniato nel 1935 dall'ecologo George Tansley, questi lo definì come "un sistema formato dall'insieme di organismi vegetali e animali che popolano un dato luogo circoscritto e caratterizzato da specifici fattori ambientali". Un ecosistema è il risultato delle interazioni tra tutti gli organismi che vi vivono e, tra questi e l'ambiente fisico, chimico e geo-morfologico. Per la stima della diversità biologica a livello di ecosistema è stato introdotto il concetto di bioma, il quale raggruppa l'insieme degli ecosistemi con caratteristiche comuni distribuiti in ogni particolare area geografica del pianeta. Ricordiamo, per esempio, il bioma del mediterraneo descritto come una zona di transizione tra gli ambienti temperati e quelli caldi presenti nell'emisfero boreale tra la fascia temperata e quella tropicale: questo ambiente è caratterizzato da associazioni di alberi e arbusti che vanno sotto il nome di macchia mediterranea. Le specie erbacee

meglio adattate a questo clima sono le annuali terofite che, alla fine della stagione vegetativa, muoiono e superano la stagione arida (l'estate) sotto forma di seme. In queste associazioni vegetali un posto di rilievo è occupato da un insieme di arbusti xerofili, sempreverdi e spesso aromatici, che costituiscono tipici elementi del bioma mediterraneo. Per stimare la biodiversità degli ecosistemi ci si limita a classificare il bioma, che caratterizza un'area delimitata, e a descrivere eventuali fattori che possono turbare l'equilibrio del bioma e dei suoi ecosistemi. Definiamo quindi la biodiversità a livello di ecosistema come la diversità di tutte le forme viventi che compongono l'ecosistema stesso, includendo sia i rapporti esistenti tra queste che quelli con le componenti abiotiche. La biodiversità a livello di specie è quella più comunemente conosciuta perché associata all'eterogeneità di forme viventi che popolano il pianeta. La diversità di specie può essere espressa come numero di specie in una determinata area (ricchezza di specie), o in base alle relazioni evolutive tra le specie che condividono uno stesso habitat (diversità filogenetica). Nonostante la diversità a livello di specie sia la più conosciuta e intuibile, essa non è di facile stima perché moltissime specie viventi sono, di fatto, sconosciute e, di conseguenza, non classificate. A ciò va aggiunta la difficoltà nel definire una specie in base ai caratteri fenotipici e morfologici; in alcuni casi, infatti, risulta davvero difficile per gli studiosi discriminare tra i taxa. Il concetto di specie maggiormente accettato si deve a Mayr (1963), che coniò la seguente definizione basata sulla compatibilità sessuale: "la specie è un insieme di individui in grado di incrociarsi fra di loro, effettivamente o potenzialmente, per produrre una discendenza a sua volta fertile, riproduttivamente isolata da altre popolazioni simili". Per quanto questa definizione sia considerata da molti studiosi la migliore in grado di definire il concetto di specie, è evidente come in alcuni casi, in particolare quelli riconducibili al mondo vegetale, tale definizione non risulti esaustiva a causa, per esempio, della capacità che hanno le piante di attuare meccanismi di riproduzione vegetativa e di generare ibridi. Nel primo caso, le popolazioni geneticamente identiche tra loro tendono ciascuna a riprodursi senza cambiare nel tempo, formando complessi di micro-specie identiche. Mentre nel secondo caso spesso gli ibridi sono inizialmente sterili per poi diventare fecondi per poliploidizzazione, ed evolvere quindi in una nuova specie (Kingsley R. Stern et al. 2009).

Gli indici usati per la stima della biodiversità a livello di specie sono relazionati al numero di specie esistenti in una determinata area e al numero di individui presenti per ogni specie. Su queste basi si fonda l'indice di biodiversità di Shannon e Wiener (Shannon 1948) descritto dalla seguente formula matematica:

$$H_{sh} = - \sum_{i=1}^s p_i \log_2 p_i$$

dove s = numero di specie identificate; p_i = probabilità che un individuo preso a caso dalla comunità appartenga alla specie $i = 1, 2, 3 \dots s$.

I valori ottenuti saranno compresi nell'intervallo zero-infinito; i valori tenderanno a zero quando tutti gli individui campionati appartengono alla stessa o a pochissime specie, al contrario per valori tendenti a infinito, sarà molto improbabile campionare individui appartenenti alla medesima specie.

Una seconda equazione impiegata per la stima della diversità biologica a livello di specie è quella di Simpson (Simpson 1949). La formula riportata di seguito descrive la probabilità che due individui campionati appartengano alla stessa specie:

$$H_{SI} = 1 - \sum_{i=1}^s p_i^2$$

dove: s = numero di specie identificate; p_i = probabilità che un individuo preso a caso dalla comunità appartenga alla specie $i = 1, 2, 3 \dots s$.

In questo caso i valori saranno compresi nell'intervallo da zero a uno; H_{SI} tendente a uno rivela una biodiversità maggiore.

Data la difficoltà di ottenere informazioni univoche e riproducibili circa la biodiversità a livello di specie, e grazie allo sviluppo e alla diffusione delle metodiche di indagine molecolare, i ricercatori hanno focalizzato il loro interesse sul livello genetico di biodiversità. La diversità a livello genetico si riferisce alle differenze esistenti nella sequenza nucleotidica che costituisce il patrimonio genetico di ciascun individuo. Questo tipo di biodiversità è rilevabile a livello di specie, ma anche di popolazione o addirittura di singolo individuo. La variabilità genetica, responsabile anche dei diversi fenotipi, rende di fatto possibile la continua evoluzione di tutti gli organismi viventi, e perciò anche delle piante, attraverso errori casuali nella replicazione del DNA (inserzioni, traslocazioni, delezioni, etc.). Questi errori potrebbero avere come conseguenza la comparsa o la scomparsa di una determinata caratteristica biologica su cui la selezione naturale può liberamente agire nel corso dei millenni. La maggior parte delle mutazioni risultano o neutrali o negative per l'adattamento dell'organismo, altre invece possono fornire all'organismo che le ha acquisite caratteristiche tali da permettergli un migliore adattamento alle condizioni ambientali. Inoltre, durante le fasi riproduttive, avvengono ulteriori eventi di modificazione dell'informazione genetica (*crossing over*) che aumenta notevolmente la variabilità della popolazione. Questa variabilità genetica, che si riscontra nelle popolazioni naturali e che ne garantisce la sopravvivenza, è quella che distingue gli organismi di una stessa specie. Le popolazioni dove la biodiversità genetica si riduce al minimo (nel caso di popolazioni a riproduzione vegetativa, deriva genetica, etc.) sono particolarmente sensibili a eventi critici, quali: cambiamenti climatici, ambientali, diffusione di patogeni, etc., e potrebbero avere limitate capacità di adattarsi (Grassi et al. 2006). Questa limitata variabilità genetica tipica delle popolazioni, che si riproducono per via vegetativa non permette alle piante, infatti, di rispondere alle pressioni selettive in maniera differenziata andando così incontro a una probabile estinzione, in caso di situazioni ambientali estreme e alle quali l'individuo e le popolazioni con limitata biodiversità non sono certamente in grado di adattarsi. L'analisi del DNA mediante varie metodiche molecolari messe a punto negli ultimi decenni quali i RAPD (*Random Amplified Polymorphic DNA*), gli SSR (*Simple Sequence Repeat*), gli AFLP (*Amplified Fragment Length Polymorphism*), il sequenziamento del DNA bar code, permette di discriminare anche i singoli individui, oltre che le popolazioni, le specie e i generi (Medina et al. 2006). Tali tecniche permettono anche di colmare una lacuna per le specie vegetali creata dalla definizione di specie proposta da Mayr (1963). Come detto in precedenza questa definizione non è sempre adatta per gli organismi vegetali, infatti, all'interno della stessa specie esistono differenze tra le piante che hanno reso indispensabile la creazione di nuove entità tassonomiche di livello inferiore alla specie, quali: sottospecie, ecotipo o forma (Grassi et al. 2006).

Per studiare la biodiversità a livello genetico esistono numerosi marcatori molecolari, di cui discuteremo nel seguente paragrafo. Ovviamente i dati ottenuti dalle analisi genetiche hanno necessità di essere elaborati, per tale motivo alcuni studiosi hanno messo a punto indici bioinformatica ad hoc per la stima della biodiversità a livello genetico. Anche questi indici saranno discussi e approfonditi nei seguenti paragrafi.

1.2 Marcatori molecolari

La biodiversità genetica può essere indagata attraverso numerosi marcatori molecolari, per la maggior parte di essi si tratta di *tools* basati sulla PCR (*Polymerase Chain Reaction*) (Mullis et al. 1992). Una delle peculiarità che ha reso i marcatori molecolari particolarmente adatti alla caratterizzazione dei polimorfismi genetici per lo studio delle diversità o similarità genetiche è riconducibile alla loro neutralità. Secondo la cosiddetta teoria della neutralità, formulata nel (1983) da Kimura e collaboratori, la maggior parte delle modificazioni molecolari sono il risultato di mutazioni selettivamente neutrali. Tale teoria assume che solo una piccolissima frazione dei cambiamenti del DNA abbia significato in termini adattativi, con effetto sulla fitness. La maggior parte delle mutazioni sarebbe, infatti, silente (sostituzione sinonima), ovvero tale da lasciare il prodotto genico inalterato e quindi con effetti così piccoli sulla capacità di sopravvivenza e riproduttività di un organismo, che la selezione naturale sarebbe incapace di influenzarne in modo apprezzabile la frequenza. La costruzione di mappe genetiche e il sequenziamento di cloni genomici di molte specie vegetali hanno fornito molte informazioni circa i meccanismi di evoluzione genica a livello molecolare. Nel progetto di dottorato discusso in questa tesi sono stati usati due differenti marcatori, gli SSR e gli AFLP. Inoltre, grazie ad una evoluzione della metodica AFLP, la MSAP (*Methylation Sensitive Amplification Polymorphism*), è stato possibile indagare anche lo stato di metilazione delle popolazioni naturali di pioppo bianco della Sardegna allo scopo di valutare eventuali alterazioni epigenetiche di cloni cresciuti in condizioni ambientali diverse e in aree poste anche a decine di chilometri di distanza.

1.2.1 Simple Sequence Repeat (SSR)

Gli SSR sono marcatori molecolari basati sull'identificazione di polimorfismi a livello delle sequenze di DNA ripetute, come nel caso di SSR formati da 2-5 bp ripetuti n volte, [e.g. (CA) n , (GCC) n , (GATA) n]. È stato stimato che esista, in media, una regione SSR ogni 50-750 Kb, in relazione alla sequenza ripetuta. La dimensione delle sequenze di DNA ripetitiva è generalmente molto ridotta: in mais, ad esempio, si aggira mediamente tra 100 e 200 bp (Taramino and Tingey 1996). Nelle piante, le ripetizioni AT risultano presenti in misura significativamente superiore a quanto atteso in base alla frequenza riscontrata nei genomi per ciascuno dei possibili dinucleotidi, mentre quelle AC/TG sono risultate piuttosto rare, contrariamente a quanto riportato per i mammiferi (Tautz and Renz 1984). Attraverso l'utilizzo di oligonucleotidi innesco (*primer*), complementari alle regioni che fiancheggiano le sequenze ripetute, è possibile amplificare questi frammenti genomici e renderli quindi rilevabili mediante tecniche elettroforetiche.

Gli SSR presentano un elevato polimorfismo, un'ampia distribuzione nel genoma, sono codominanti, facili da analizzare e hanno un'elevata riproducibilità. Questi marcatori, inoltre, sono spesso trasferibili da una specie a un'altra/e filogeneticamente molto prossima; questa caratteristica li rende anche ideali per la costruzione di mappe genetiche e "fingerprinting" (Kalia et al. 2010).

La dimensione del genoma delle piante è molto diversificata, si va dalle 145 Mb di *Arabidopsis* alle 5.400 Mb di mais, per arrivare fino alle 33.000 Mb del frumento (Kingsley R. Stern et al. 2009). Una caratteristica comune dei genomi di grandi dimensioni, come quelli tipici di molte specie di interesse agrario, è rappresentata dal possedere un numero relativamente contenuto di geni e un numero elevatissimo di elementi ripetuti. In tempi recenti si è assistito a una rapida evoluzione delle tecniche molecolari e al loro crescente utilizzo sia per l'acquisizione di conoscenze di base che per la soluzione di problemi di rilevanza applicativa. Dati relativi alla distribuzione delle dimensioni degli alleli e delle distanze genetiche calcolate sulla base di polimorfismi SSR hanno suggerito l'esistenza di

un meccanismo mutazionale particolare, noto come *Stepwise Mutational Model*, proposto da Valdes e collaboratori (1993). Tale meccanismo prevede per un singolo evento mutazionale l'inserzione o la delezione, con uguale probabilità, di una singola ripetizione nell'allele mutato rispetto a quello originario. In questo modo, alleli di dimensioni paragonabili risulterebbero simili anche in termini di eventi mutazionali subiti e conseguentemente di età evolutiva; tale conclusione invece non dovrebbe essere considerata valida nel caso di alleli aventi dimensioni diverse. Il tasso di mutazione sembra, inoltre, essere correlato con la dimensione della sequenza SSR: a un dato locus, alleli di dimensioni maggiori presentano una maggiore incidenza di mutazioni e, quindi, un maggiore livello di polimorfismo. In definitiva, gli SSR sono una classe di marcatori molecolari capaci di rilevare un'ampia gamma di diversità genetica in quanto i polimorfismi sono dovuti a varianti alleliche del medesimo locus: il numero di alleli riportato per singolo locus SSR varia da 3-5 in orzo, 6-8 in soia, 5-11 in riso, fino a oltre 20 in mais (Kalia et al. 2010). Tale particolarità rende gli SSR adatti alla caratterizzazione genotipica e all'identificazione e tipizzazione varietale. Un'altra peculiarità dei marcatori SSR, che ne giustifica il largo impiego nell'analisi finalizzata alla costruzione di mappe genetiche, è connessa alla loro natura codominante: l'analisi di singoli loci rende possibile rivelare sia gli alleli omozigoti che quelli eterozigoti. Tuttavia, questi marcatori presentano alcune limitazioni nello studio della struttura e della ripartizione della variabilità genetica intra-popolazione. Le mutazioni che originano tali polimorfismi, infatti, possono avere una frequenza variabile a seconda della specie, inficiando così lo studio delle relazioni filogenetiche tra specie diverse nell'ambito dello stesso genere e tra popolazioni nell'ambito della stessa specie. Inoltre, il meccanismo mutazionale può essere reversibile comportando omoplasia tra individui, cioè identità di situazione (*identity-in-state*), ma non identità di derivazione (*identity-by-descent*). Questo significa che specie, o popolazioni nell'ambito di una specie, apparentemente vicine, perché presentano alleli analoghi allo stesso locus SSR, possono in realtà differire per l'origine delle sequenze ripetute e amplificate. Nonostante questa limitazione, gli SSR hanno trovato e trovano tuttora numerose applicazioni in indagini filogenetiche ed evolutive.

1.2.2 Amplified Fragment Length Polymorphism (AFLP)

Gli AFLP (Vos et al. 1995) si basano sull'amplificazione selettiva mediante PCR di frammenti di DNA derivanti dalla digestione del DNA genomico con specifici enzimi di restrizione. Il protocollo sperimentale per la rilevazione di tali marcatori prevede: i) la digestione del DNA genomico con due differenti enzimi di restrizione, un *rare cutter* (in genere, *EcoRI* o *PstI*) e un *frequent cutter* (in genere, *MseI* nelle piante e *TaqI* negli animali); ii) la ligazione di adattatori oligonucleotidici ai frammenti di restrizione; iii) una prima amplificazione selettiva (o preamplificazione) mediante due *primer* complementari alle sequenze dei siti di restrizione e degli adattatori, aventi ciascuno una base selettiva in posizione 3'OH; iv) una seconda amplificazione selettiva con due *primer* complementari alle sequenze dei siti di restrizione e degli adattatori, ma aventi due o tre nucleotidi selettivi ciascuno. Uno dei *primer* deve essere preventivamente marcato in 5' ricorrendo a coloranti fluorescenti (<http://www3.appliedbiosystems.com>); v) la separazione dei frammenti amplificati attraverso elettroforesi capillare su sequenziatore automatico che permette la discriminazione anche di una singola base.

La numerosità dei marcatori molecolari evidenziabili per combinazione di *primer* è comunque una funzione del numero di basi selettive adottate e della dimensione del genoma analizzato. Una formula che consente di calcolare il numero di prodotti di amplificazione (Npa) teoricamente ottenibili in un genoma di dimensione nota è la seguente:

$$Npa = \frac{2G}{4,096} \times \left(\frac{1}{4}\right)^n$$

dove G è la dimensione del genoma espressa in Kb e n è il numero totale di basi selettive.

Tale formula è valida quando si adottano combinazioni di enzimi di restrizione che riconoscono quattro e sei coppie di basi, come ad esempio *EcoRI-MseI* o *PstI-TaqI*. La prima parte della formula fornisce una stima del numero totale di frammenti di restrizione utili per la ligazione degli adattatori, mentre la seconda parte misura l'incidenza del numero totale di basi selettive sulla riduzione di frammenti amplificabili. Naturalmente la scelta delle endonucleasi può influenzare considerevolmente il numero di frammenti di DNA realmente amplificabili in quanto ogni genoma è caratterizzato da un contenuto peculiare di guanine e citosine e, quindi, da una abbondanza o carenza di specifici siti di restrizione. L'origine dei polimorfismi AFLP è molteplice dal momento che questi possono essere causati da riarrangiamenti del DNA, che determinano la perdita o la presenza di siti di restrizione, ma anche a cambiamenti nella sequenza interna dove agiscono le basi selettive di uno o entrambi i *primer*, oppure da variazioni della lunghezza totale del frammento digerito e amplificato. In generale, i marcatori AFLP riescono a saggiare anche le regioni telomeriche, laddove gli RFLP (*Random Fragment Length Polymorphism*) (Williams et al. 1990) forniscono invece scarse informazioni. Le regioni cromosomiche indagate da tali marcatori sono comunque strettamente dipendenti dall'enzima di restrizione usato, per esempio, *EcoRI* e *PstI* hanno una diversa sensibilità alla metilazione. Il primo appartiene alla categoria degli enzimi di restrizione insensibili alla metilazione (*methylation insensitive*), tagliano cioè la doppia elica del DNA anche in corrispondenza di basi metilate, mentre il secondo appartiene alla categoria di quelli sensibili alla metilazione (*methylation sensitive*), vale a dire che non operano tagli in presenza di basi metilate. La presenza di basi metilate nella catena del DNA, per l'aggiunta enzimatica di gruppi CH₃ a specifici residui di citosina o adenina, rende alcuni enzimi incapaci di operare tagli. È stato dimostrato che lo stato di metilazione di alcune sequenze CG in regioni di controllo sia in relazione con l'espressione genica (Fujimoto et al. 2012). Esiste un'evoluzione della tecnica AFLP, denominata MSAP, che consente di analizzare il polimorfismo dovuto alla metilazione a livello della citosina contenuta in specifiche sequenze di DNA. Questa tecnica è basata sull'impiego degli enzimi di restrizione *HpaII* e *MspI*, due isoschizomeri che differiscono in termini di sensibilità alla metilazione a carico della sequenza riconosciuta come sito di taglio. Entrambe le endonucleasi eseguono la restrizione del DNA in corrispondenza della stessa sequenza tetranucleotidica (5'-CCGG-3'), ma con modalità di azione condizionata dallo stato di metilazione del residuo citosinico esterno o interno: l'enzima *HpaII* non agisce quando una o entrambe le C della sequenza bersaglio sono metilate, ovvero quando a entrambi i filamenti sono legati covalentemente gruppi CH₃, mentre è attivo quando la sequenza di riconoscimento risulta emimetilata, ovvero uno soltanto dei filamenti presenta almeno un gruppo CH₃ a livello della C esterna o interna. Viceversa, l'enzima *MspI* è in grado di operare la restrizione in corrispondenza della sequenza di riconoscimento sia quando è emimetilata o completamente metilata, purché il gruppo CH₃ sia legato alla C interna (5'-C^mCCGG-3'), mentre non è attivo nel caso in cui la C interessata sia quella esterna (5'-^mCCGG-3'). In Fig. 1 sono riportati, in maniera schematica, le situazioni relative a una sequenza non metilata, emimetilata e completamente metilata. La demetilazione tende a fornire prodotti di amplificazione di lunghezza ridotta in quanto aumenta la frequenza dei siti cromosomici soggetti a normale restrizione.

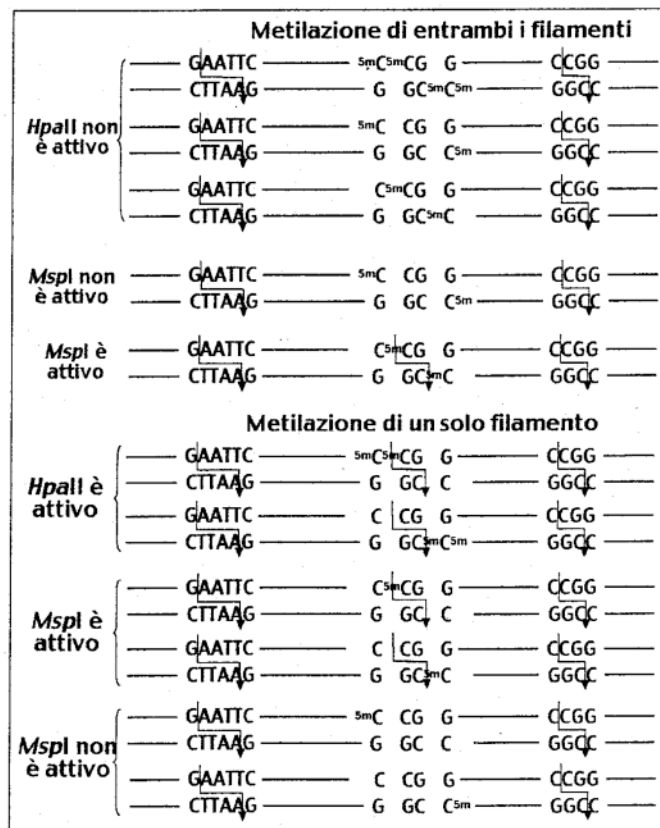


Fig. 1. La figura riporta lo stato di attivazione degli enzimi *MspI* e *HpaII* in relazione allo stato di DNA non metilato, emimetilato e metilato.

Allo stato attuale si ritiene che la metilazione delle C rappresenti un meccanismo capace di svolgere un ruolo importante nella regolazione dell'espressione genica (Putiri and Robertson 2011). In particolare, la metilazione è in grado di inibire l'espressione dei geni corrispondenti, così come la demetilazione artificiale dei geni può determinare una loro riattivazione (Zhang et al. 2006). Nelle piante la metilazione del DNA è stata spesso associata a importanti fenomeni epigenetici come, ad esempio, i meccanismi di regolazione attivi durante lo sviluppo e la differenziazione cellulare, la risposta a stress pedoclimatici e ambientali (Ahuja et al. 2010; Grativol et al. 2012; Karan et al. 2012; Wang et al. 2012).

1.3 Epigenetica

L'epigenetica è una branca della biologia molecolare che studia i cambiamenti della regolazione genica che non dipendono da mutazioni del DNA, ma da cambiamenti della struttura chimica dei residui di citosina. Negli ultimi anni è stato chiaramente dimostrato che modificazioni epigenetiche quali la metilazione del DNA, le modificazioni delle code istoniche e la presenza di varianti istoniche possono agire a livello della cromatina per alterare e regolare spazialmente e temporalmente l'espressione genica, specialmente in risposta a fattori ambientali (Vaillant and Paszkowski 2007; Boyko and Kovalchuk 2008; Mazzio and Soliman 2012). Queste modifiche reversibili possono essere modulate in risposta a stimoli ambientali, oppure essere regolazioni previste nello sviluppo di un

organismo. E' stato dimostrato che alcune modifiche epigenetiche possono essere ereditate attraverso la mitosi e la meiosi, portando quindi alla creazione di una nuova variabilità genetica che può essere trasmessa alla progenie (Boyko and Kovalchuk 2008). La regolazione epigenetica dell'espressione genica è modulata attraverso tre distinte vie biosintetiche o *pathway*: il *pathway* autonomo, inizialmente individuato come *pathway* di regolazione della fioritura, la metilazione del DNA e il *pathway* degli small-RNA (sRNA). Questi tre diversi meccanismi epigenetici sono strettamente interconnessi tra di loro e insieme sono coinvolti nella regolazione dello sviluppo delle piante. Il *pathway* autonomo comprende geni e prodotti genici inizialmente identificati perché coinvolti nella regolazione della fioritura della pianta modello *Arabidopsis thaliana*. In *Arabidopsis* l'induzione alla fioritura è regolata dal silenziamento mediante metilazione del gene FLC (*flowering locus C*), che funge da regolatore negativo dell'induzione a fiore. I geni del *pathway* autonomo, insieme all'influenza di fattori ambientali (luce e freddo in primis) e ormonali, portano al silenziamento di FLC permettendo l'induzione a fiore (Buchanan et al. 2003). Il *pathway* di metilazione delle citosine è invece responsabile del silenziamento trascrizionale delle regioni eucromatiche e degli elementi trasponibili. Mediante l'azione degli enzimi metil-transferasici i gruppi metilici vengono legati covalentemente alle citosine del DNA (Zhang et al. 2010). La metilazione delle citosine, interferendo con i complessi proteici che effettuano la trascrizione, causa il silenziamento dell'espressione dei geni (Fig. 2).

Le piante possiedono tre tipi di DNA metiltransferasi con localizzazioni e funzioni diverse. La MET1 (*methyltransferase1*) mantiene una metilazione simmetrica sui siti 5'-CG-3' durante la replicazione del DNA, copiando l'informazione dal filamento stampo al filamento di nuova sintesi (Zhang et al. 2010). Questo tipo di metilazione è utilizzata per il silenziamento degli elementi trasponibili. Le CMT3 (*chromomethylase3*) metilano le sequenze 5'-CHG-3' e possono effettuare una metilazione de novo di siti del DNA in corrispondenza di regioni con particolari modificazioni istoniche (Zhang et al. 2010). Le proteine DRM1 (*domains rearranged methyltransferase 1*) e DRM2 (*domains rearranged methyltransferase 2*) sono strettamente correlate tra loro e hanno funzioni simili nel metilare le sequenze CHH (dove H può essere una A, C o T) non simmetriche (Fig. 3). Questo tipo di metilazione porta alla perdita di informazione durante la replicazione del DNA, in quanto un filamento non dispone della C metilata. Le proteine CMT3, DRM1 e DRM2 sono necessarie per la regolazione dell'espressione di geni endogeni, specialmente per quelli legati al controllo dello sviluppo. La rappresentazione schematica delle diverse attività delle DNA metiltransferasi è riportata in figura 3.

Il *pathway* di produzione degli siRNA (*small interfering RNA*) include invece una serie di meccanismi di silenziamento genico post-trascrizionale. Gli siRNA sono RNA di lunghezza variabile tra i 20-25 nucleotidi formati dal processamento di più lunghi RNA a doppio filamento (dsRNA). Il doppio filamento di siRNA è aperto da un'elicasi e il filamento antisense è legato dal complesso RISC (*RNA-interfering silencing complex*), mentre il filamento senso è degradato. Il complesso così formato può legarsi a una sequenza complementare presente sull'mRNA (RNA messaggero), se l'appaiamento risulta perfetto una componente di RISC (detta *argonaute protein* o Ago) effettua un taglio sull'mRNA (Chitwood and Timmermans 2010). Così facendo si ottengono due frammenti di mRNA privi di protezione e che sono rapidamente degradati dalle RNAsi. Nel caso in cui l'appaiamento non sia perfetto il complesso RISC inibisce comunque la traduzione del trascritto in proteina, bloccando quindi l'espressione del gene.

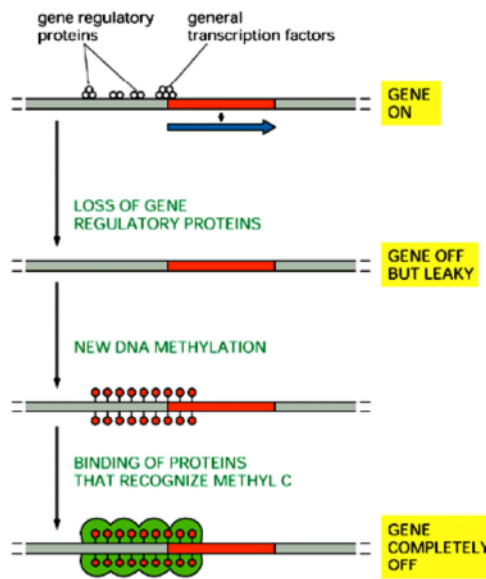


Fig. 2. Rappresentazione schematica dell'effetto della metilazione sull'espressione genica. La metilazione interferisce con i gruppi proteici che effettuano la trascrizione causando il silenziamento genico.

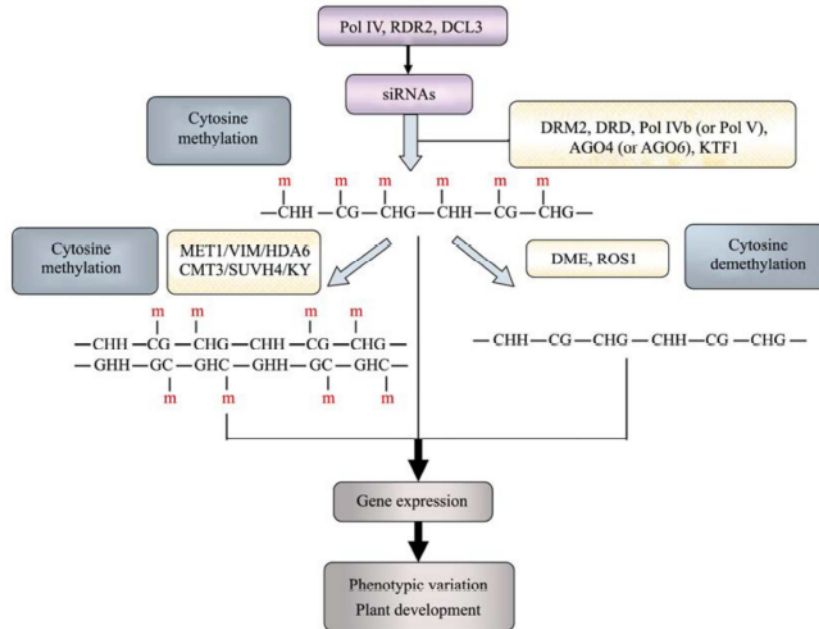


Fig. 3. Nella figura qui riportata è illustrato il meccanismo di metilazione da parte degli enzimi citati precedentemente. In particolare si può notare come la metilazione non sia necessaria soltanto per il silenziamento genico, ma anche per la regolazione dell'espressione genica mediante metilazione di mantenimento, specialmente per i geni legati al controllo dello sviluppo (Zhang et al. 2010).

La metilazione delle C nelle regioni promotrici può influenzare l'espressione genica e contribuire quindi alla variabilità fenotipica. I meccanismi epigenetici, come il *pathway* di metilazione e quello autonomo, influenzano l'espressione genica anche agendo direttamente a livello della cromatina, o meglio, sullo stato di condensazione della cromatina, costituita da DNA associato a proteine istoniche e non istoniche, principale componente del nucleo della cellula eucariote.

1.4 Biodiversità agraria

Il valore della biodiversità è particolarmente evidente in agricoltura. Per millenni i popoli hanno coltivato una vasta gamma di specie e varietà per stabilizzare la produzione e diversificare l'alimentazione, valore che viene nuovamente apprezzato oggi, come dimostrato dalla ricerca di processi e sistemi produttivi a basso impatto ambientale. La crescita della popolazione mondiale sta riportando l'attenzione anche su specie che erano state trascurate e che di conseguenza erano pressoché scomparse. Esse possono fornire alimenti anche in zone in cui le specie più diffuse non sembrano in grado di produrre in modo soddisfacente. Molte specie selvatiche possiedono geni per migliorare la resistenza delle piante coltivate a malattie e parassiti, evitando l'uso di prodotti inquinanti. Molte specie selvatiche forniscono, inoltre, i principi attivi per trattare malattie come il cancro, la leucemia e il morbo di Parkinson (Panickar 2013).

Negli ultimi anni, i prodotti dell'agricoltura sono stati utilizzati anche per fornire biocarburanti, sebbene la comunità scientifica sia divisa su questo punto; se da una parte il loro uso può aiutare a non fare aumentare le emissioni di gas serra, dall'altra pone delle gravi conseguenze economiche e sociali. La conversione di terreni precedentemente usati per la coltivazione di prodotti alimentari per generare biocarburanti è stata uno dei fattori che ha contribuito alla recente crisi dei prezzi delle derrate alimentari. La biodiversità agraria, o agro-biodiversità, è un sottoinsieme della biodiversità che include tutte le forme viventi direttamente connesse all'ecosistema agricolo. Gli ecologi hanno sostenuto che l'aumento del numero di specie presenti e rappresentate da un numero elevato di piante aumenta ovviamente la biodiversità, che a sua volta incrementa la stabilità ecologica e la resistenza a organismi patogeni delle piante (Cohen and Tilman 1996). In particolare è stato accertato che, a livello di azienda agricola, l'impiego diversificato di colture cerealicole aumenta anche la biodiversità della fauna del suolo con effetti positivi sulla produttività, resistenza alle malattie e ai patogeni; in questo modo anche il reddito degli agricoltori è più stabile e sicuro (Di Falco and Perrings 2005; Asrat et al. 2010; Di Falco et al. 2010)

Negli ultimi cinquant'anni, in particolare, il progresso tecnico e lo sviluppo economico mondiale, accompagnati dall'aumento esponenziale della popolazione umana, hanno portato a molteplici e gravi modificazioni ambientali, che, nel lungo periodo, potrebbero comportare seri problemi agli agro-ecosistemi, con conseguenti estinzioni di specie vegetali e animali selvatiche, che non si raffigura come un evento isolato, ma come risultato di una vera e propria trasformazione ambientale di portata globale (Grassi et al. 2006). Oggi nel mondo sono a rischio oltre 15.500 specie, di cui la metà sono animali (7.266), secondo quanto si può leggere nella Lista Rossa (*Red Book*) dell'Ufficio Internazionale per la Conservazione della Natura (IUCN) (<http://www.iucn.org>). Negli ultimi 50 anni si sono registrati purtroppo anche cambiamenti epocali nelle tecniche di coltivazione dei terreni agrari, con forti ripercussioni sulla scelta delle specie di piante da coltivare o degli animali da allevare. Sebbene si conoscano almeno 80mila piante commestibili, l'alimentazione mondiale dipende ormai solo da una ventina di specie (leguminose, tuberose ed erbe), con un patrimonio genetico sempre più omogeneo e uniforme. In regioni geografiche lontanissime tra loro si finisce oggi con l'utilizzare praticamente sempre le stesse sementi.

La biodiversità di varietà tradizionali, la presenza cioè di piante con caratteri ereditari diversi, è un patrimonio di caratteristiche genetiche preziose (fertilità, robustezza, resistenza ai fattori ambientali ostili caldo, freddo, siccità, etc. o alle malattie) ad oggi quasi sconosciuto o fortemente trascurato, che potrebbe essere utile in un prossimo futuro anche dal punto di vista economico. L'uomo ha sfruttato la variabilità genetica nel corso dei millenni, prima, inconsapevolmente, e poi applicando i principi della selezione,

costatando a quali drammi possa condurre l'impoverimento della biodiversità vegetale. Basta ricordare la tremenda carestia che colpì l'Irlanda nel 1845, causando la morte di 2 milioni di persone, dovuta all'utilizzo di patate moltiplicate da un "pool" genico ristretto, completamente distrutte da un'infestazione di *Phytophthora infestans*, un fungo contro cui non si avevano a disposizione mezzi di difesa. Sempre nel XIX secolo, la mancanza di variabilità genetica causò la totale distruzione delle piante di caffè sull'isola di Ceylon o, per restare più vicini a noi nel tempo e nello spazio, lo sviluppo della peronospora del tabacco negli anni '60 in Italia, o gli attacchi sul mais negli anni '70 in USA causarono danni altrettanto gravi e drammatici (Ricciardi L. and Filippetti A. 2003). In sostanza, l'erosione genetica influenza fortemente la capacità di adattamento delle specie alle mutevoli condizioni ambientali, e inibisce la possibilità di migliorare le specie agricole rispetto alle loro utili caratteristiche.

E' un merito in particolare dell'agricoltura biologica di molti paesi occidentali, l'aver recentemente riscoperto l'uso di antiche varietà agricole (dai cereali, ai legumi, alla frutta e agli ortaggi) assai poco utilizzate nel recente passato perché assai poco produttive. Attualmente queste varietà sono coltivate dai cosiddetti agricoltori "custode" nelle zone più marginali; è il *World Conservation Monitoring Centre* (WCMC) a indicarci quali siano state e siano tuttora le cause principali che conducono a questa erosione:

- la perdita e la modifica degli habitat, spesso associate alla loro frammentazione;
- il sovrasfruttamento delle risorse per motivi commerciali o di sussistenza;
- l'introduzione di specie esotiche, molto spesso in competizione con le specie indigene locali;
- il disturbo, l'estirpazione, lo sradicamento delle risorse genetiche e il loro accidentale o intenzionale prelievo;
- la limitata distribuzione delle risorse genetiche;
- le numerose fitopatie che colpiscono le piante.

Non esiste altro posto sulla Terra che presenti come l'Europa, ma, in particolare, il bacino del Mediterraneo, un'interazione continua tra biodiversità e azione dell'uomo. Sulle coste mediterranee, attraverso millenni di pratica agricola e allevamento tradizionale, si è creata una straordinaria varietà di ecotipi e pool genici di piante coltivate e di animali addomesticati. La maggiore ricchezza di biodiversità vegetale dell'area si riscontra in particolare sulle Alpi, nei Balcani, nell'Italia meridionale, nella penisola iberica, in Grecia, Cipro e sul Mar Nero. Ma è in assoluto l'Italia a essere la nazione più ricca di specie vegetali, con 5.650, seguita dai paesi balcanici con 5.350 specie, dalla Spagna (5.050), dalla Grecia (5.000) e dalla Francia (4.650) (Ricciardi L. and Filippetti A. 2003).

D'altra parte è proprio nel bioma del mediterraneo che si sono originate e diversificate numerose specie agrarie, per esempio: frumento, segale e avena per i cereali; fava, cece, pisello, lupino per le leguminose; erba medica, trifoglio, festuca e loglio per le foraggere. Il bioma del Mediterraneo è la patria anche di numerose specie orticole (carciofo, carota, cavolo, lattuga, cicoria, finocchio, asparago, etc.), oleaginose (olivo e cartamo), fruttifere (tra cui vite, fico e i generi *Malus*, *Pyrus* e *Prunus*), senza dimenticare la canapa, il lino, e le specie aromatiche, quali rosmarino, salvia, alloro e menta. Preservare la varietà genetica legata alle antiche tecniche di agricoltura riveste, in Europa e nell'area mediterranea più che altrove, un'importanza decisiva visto che gli agricoltori tradizionali, specie sulle sponde meridionali, stanno ulteriormente abbandonando varietà locali. E' perciò quanto mai urgente identificare e salvare i pool genici selvatici e ancestrali tipici ancora presenti, prima che essi si perdano completamente e siano, come ormai purtroppo è già accaduto per molti paesi a nord del bacino, rimpiazzati da nuove varietà anche all'interno delle loro aree di diffusione tradizionali.

Come asserito in precedenza, l'Italia è il paese, nel bacino del Mediterraneo, più ricco di specie vegetali; una ricchezza che si riflette ampiamente anche nelle specie vegetali agrarie, grazie alla sua lunga storia rurale, privilegiata sia climaticamente dalla sua conformazione geografica, che dal punto di vista storico-culturale dalle continue immissioni di popolazioni diverse e dalle lunghe vicissitudini storiche che risalgono a prima dell'impero romano.

Vavilov (2009) indica che la nostra penisola è il centro genetico del Mediterraneo, da cui si sono originate molte delle specie adatte alla coltivazione, analisi poi confermata anche dagli studi e dalle ricerche dell'istituto del Germoplasma di Bari. Nell'Italia centro-settentrionale è stata verificata la presenza di 551 varietà, escludendo le piante ornamentali. Attraverso il miglioramento genetico, molti ecotipi oggi scomparsi sarebbero certamente preziosi per giungere alla costituzione di interessanti e nuove varietà con particolari caratteristiche morfo-genetiche. Per quanto riguarda gli ortaggi, il grosso dell'erosione genetica in Italia si è verificato a partire dagli anni '50, ma ha raggiunto il culmine negli anni '60 (Grassi et al. 2006).

Fu proprio durante gli anni del cosiddetto "boom economico" che molte antiche varietà furono definitivamente perse e sostituite con altre definite più "moderne e produttive", ma sempre più uniformi dal punto di vista genetico, generate da sementi ibride massicciamente importate dal Nord Europa e dagli USA, che furono causa della diffusione di numerose fitopatie per il loro scarso adattamento ai climi meridionali. Per quanto riguarda le specie frutticole (Pero, Melo, Susino e Ciliegio) erano tutte specie che presentavano un numero elevato di varietà, adattate al clima, all'altitudine ma anche alle condizioni pedologiche dei suoli. Tra i frutti è necessario annoverare poi le castagne prodotte dai castagneti diffusi e tradizionalmente "coltivati" un po' su tutto il territorio italiano, da nord a sud. Le noci e le nocciole erano e sono diffuse ovunque in Italia; mandorle e pinoli solo nell'Italia centro-meridionale, mentre i pistacchi sono tuttora presenti solo in alcune zone dell'Italia meridionale.

Anche olivo e vite hanno modellato per secoli il paesaggio italiano presentando un insieme di varietà a seconda della loro presenza nelle diverse regioni. Secondo lo studio di Ricciardi e Filippetti (2003), l'uniformità colturale creata per la coltivazione del frumento induce a ritenere che in Italia, negli anni recenti, sia andato perso oltre il 95% delle antiche varietà di grano. In Puglia, così come in Sicilia, il 60% della superficie coltivata a grano è ormai costituita da una sola varietà. Erano oltre 250 le varietà di frumento coltivate negli anni '20 che nel 1971 erano comunque già date per scomparse; tale situazione è risultata aggravata anche dalla scomparsa dalle coltivazioni di specie affini (prezioso serbatoio di geni), in particolare le varietà di farro (oggi per fortuna in ripresa nelle regioni dell'Italia centrale). Una simile situazione di progressiva rarefazione è toccata alle varietà locali di segale, orzo e avena, cereali sempre meno utilizzati nel nostro Paese. Piselli, ceci, lenticchie, fave, fagioli, cicerchia, lupini, etc. sono le leguminose più diffuse in Italia, e costituivano le proteine per eccellenza del mondo rurale fino a non molti decenni fa. Tranne ceci, cicerchia (Italia centrale) e lenticchie (Italia centro-meridionale), i legumi sono coltivati ovunque lungo tutta la penisola, ma con considerevoli riduzioni nelle loro ricche popolazioni locali ed ecotipi. Si trattava infatti di colture utilizzate nelle rotazioni con i cereali per il loro elevato apporto di azoto, messe da parte con l'introduzione delle monoculture, ma anche per la scarsa attenzione scientifica ricevuta negli anni addietro (Ricciardi L. and Filippetti A. 2003).

1.5 Biodiversità forestale

Lo studio e la tutela della biodiversità vegetale rappresenta un presupposto indispensabile per la salvaguardia della vita animale e dunque dell'uomo stesso.

Le piante sono componenti centrali di tutti gli ecosistemi, responsabili della produzione primaria di biomassa e ossigeno, forniscono cibo e riparo alle componenti biotiche, regolano i cicli trofici degli elementi e dell'acqua, contribuiscono a formare e valorizzare il patrimonio paesaggistico e socio-culturale di paesi, città e nazioni. Basti pensare che in media il continente europeo, e in particolare l'Italia, possiedono circa il 35% del territorio coperto da foreste e boschi che ospitano una porzione elevata di biodiversità, in termini di specie, genotipi e processi ecologici (Fig. 4) (Armenise L. 2010). L'Italia meridionale, grazie alle caratteristiche del bioma Mediterraneo è stata un'importante area rifugio per numerose specie arboree che durante l'ultima glaciazione hanno potuto trovare, in siti specifici, condizioni ecologiche ottimali per la loro conservazione, e particolarmente idonee per iniziare il successivo processo di ricolonizzazione post-glaciale (Grassi et al. 2006). Una traccia di queste vicende è in alcuni casi ancora individuabile, ne sono esempio i 25 esemplari di abete dei Nebrodi delle Madonie (*Abies nebrodensis* Lojac. Mattei), la *Zelkova sicula* di Pasquale, Garfi&Quézel, e i popolamenti di agrifoglio dei Nebrodi (*Ilex aquifolium* L.), la Foresta Umbra del Gargano dove il faggio (*Fagus sylvatica* L.) è presente a quote insolitamente basse, le formazioni pure e miste di abete bianco (*Abies alba* Mill.) dell'Appennino centro-meridionale e tanti altri (Fineschi et al. 2002). Sebbene l'Italia sia una nazione con una superficie non molto estesa, per le sue caratteristiche geografiche e climatiche è caratterizzata da elevata variabilità ambientale; questa variabilità è mantenuta, in scala ridotta, anche in molte regioni italiane. Si pensi ad esempio alla Campania, questa regione racchiude in circa 13 Km² vari ecosistemi: quello marino, quello collinare e quello appenninico, facendo di questa regione una tra le più ricche di biodiversità vegetale.

La tutela degli ecosistemi forestali, in Italia e nel mondo, diventa, dunque, fondamentale per assicurare sia la conservazione delle specie vegetali e animali fortemente minacciate, che la gestione sostenibile della diversità biologica in essi contenuta. Gli ecosistemi forestali comprendono quelli terrestri, più dotati di biomasse e di maggiore complessità strutturale e funzionale, in grado di costituire habitat primari per un elevato numero di specie e di contribuire al ciclo vitale di molte altre specie grazie all'interazione funzionale con altri ecosistemi. A questo scopo dall'inizio degli anni '90 sono stati adottati numerosi strumenti giuridici che nel corso degli anni hanno posto l'accento sul ruolo delle foreste nella conservazione della biodiversità. Nell'Allegato I della Direttiva "Habitat", su un totale di 198 habitat prioritari elencati, 58 sono forestali.

Su scala nazionale, invece è stata la legge quadro sulle aree protette del 6 Dicembre 1991, n.394 che ha gettato le basi per una politica ambientale volta alla conservazione delle aree caratterizzate dalla presenza di ecosistemi intatti e di formazioni fisiche, geologiche, geomorfologiche, biologiche di elevato valore naturalistico. Un'efficace tutela della biodiversità, infatti, non può che essere compiuta a due diversi livelli tra loro strettamente connessi: la protezione delle specie e quella dei loro habitat. Al fine di assicurare una relativa capacità di auto-conservazione degli ecosistemi, che contribuiscono maggiormente alla diversità biologica, è necessario tutelare non solo le aree ove popolazioni e comunità sono presenti in modo peculiare, ma garantire una protezione sufficiente anche alle aree interposte fra queste, i cosiddetti corridoi biologici.

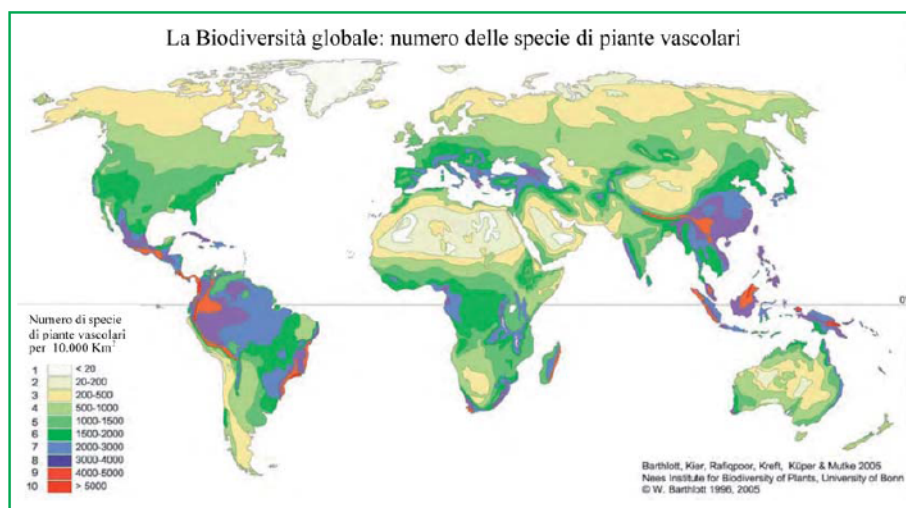


Fig.4. Rappresentazione grafica di aree omogenee per numero di specie di piante vascolari

Le Conferenze interministeriali sulla Protezione delle Foreste in Europa (MCPFE), hanno segnato, nel corso degli anni '90, le tappe del Processo Paneuropeo con il quale i Ministri, responsabili per la gestione delle foreste, si sono ufficialmente impegnati a portare avanti una serie di attività comuni volte alla salvaguardia delle nostre foreste. Durante le Conferenze di Strasburgo (1990), Helsinki (1993), Lisbona (1998) e Vienna (2003) sono state, infatti, redatte dichiarazioni generali e risoluzioni dalle quali discendono la definizione del concetto di "gestione forestale sostenibile" e l'individuazione dei relativi criteri e indicatori, nonché la definizione del concetto di "programma forestale nazionale". In seguito alla Risoluzione della Conferenza Ministeriale di Strasburgo (1990) per la Conservazione delle Foreste in Europa, è stato istituito il programma EUFORGEN (*European Forest Genetics*), che opera attraverso reti internazionali di specialisti forestali e genetisti per lo scambio di informazioni scientifiche e la collaborazione attiva, finalizzati allo sviluppo di metodi di gestione e conservazione delle risorse genetiche forestali. Il concetto di gestione forestale si è dunque evoluto nel tempo, passando da una concezione di tipo produttivistico, che valutava i sistemi e le tecniche colturali e i metodi di pianificazione in base alla misura della produzione legnosa, a quella attuale "sostenibile", che tiene conto anche delle variabili ecologiche e sociali.

Dopo la Conferenza di Rio (1992) e quella di Helsinki (1993) alla nozione di gestione sostenibile si è associata quella di biodiversità, ponendo i due principi in uno stato di interrelazione. La diversità, infatti, è un segno distintivo della natura e rappresenta la base della stabilità ecologica: le molteplici interazioni che la diversità biologica è in grado di determinare riescono, entro certi limiti, ad avere un effetto tampone rispetto alle alterazioni che gli eventi antropici producono all'interno degli ecosistemi. Con il concetto di tutela della biodiversità, dunque, non si identifica solo il problema della salvaguardia delle specie vegetali e animali rare o in via di estinzione, ma si pone l'accento sugli ecosistemi e il loro funzionamento includendo i processi coevolutivi tra i componenti che li costituiscono. Ecosistemi diversi danno luogo a forme di vita, culture e habitat diversi, la coevoluzione dei quali determina la conservazione della biodiversità (<http://www.ricercaforestale.it>).

L'articolo 7 della Convenzione sulla Diversità Biologica (Rio de Janeiro) richiede che i Paesi contraenti "identifichino le componenti della biodiversità importanti per la sua conservazione e il suo uso sostenibile e ne effettuino il monitoraggio, attraverso campionamenti o altre tecniche. La conoscenza della biodiversità forestale e il

monitoraggio sono, dunque, alla base della conservazione e della gestione del patrimonio biologico nella prospettiva di una gestione forestale sostenibile.

La conservazione della biodiversità forestale non è certamente solo la conservazione in ambiti ristretti di habitat e di specie rare e minacciate di estinzione, misura talora necessaria ma in fondo residuale e che, se fosse l'unica, certificherebbe la non adeguata gestione complessiva del fenomeno. Essa è, invece, un complesso di azioni destinate a conservare la diversità biologica ai diversi livelli di riferimento indicati dall'articolo 2 della Convenzione internazionale di Rio de Janeiro, cioè ecosistemi, specie e patrimonio genetico delle specie.

La biodiversità forestale, pur di primaria importanza, è solo una parte della più ampia diversità biologica. Per questo motivo la conservazione deve tenere conto anche della connettività e delle interazioni funzionali tra ecosistemi forestali e non forestali e, nel quadro più generale di conservazione della biodiversità. Il controllo dei dinamismi territoriali delle biocenosi forestali e il mantenimento di fenomeni di disturbo al loro insediamento ed evoluzione sono spesso un presupposto necessario per la conservazione di habitat non forestali e delle specie da questi sostenute.

1.6 Salicaceae

Il Servizio Forestale del Dipartimento dell'Agricoltura degli Stati Uniti ha recentemente richiamato l'attenzione su alcuni problemi relativi all'impiego delle Salicaceae (*Salix spp.* e *Populus spp.*) nel ripristino ambientale (Landis et al. 2003). Le Salicaceae, infatti, si caratterizzano e si differenziano da altre piante per alcuni aspetti particolarmente significativi:

- il sesso dei singoli individui è femminile o maschile (trattandosi di piante dioiche);
- sebbene la riproduzione per seme sia un'importante via naturale per la diffusione della specie nel territorio, in natura si riproducono spesso per via vegetativa e la stessa propagazione artificiale si basa prevalentemente sulla produzione di talee;
- nelle diverse regioni biogeografiche di Italia sono rappresentate da moltissime specie, che si concentrano soprattutto nelle aree ripariali. Il diffondersi delle attività di ripristino ambientale ha favorito l'impiego sempre più massiccio di talee di salice e di pioppo, in modo particolare negli interventi di rinaturalizzazione dei corsi d'acqua (aree golenali, meandri, etc.), di consolidamento spondale e nella costituzione di ecosistemi filtro. Partendo dalla constatazione che le talee di Salicaceae già radicate attecchiscono più facilmente e si sviluppano meglio rispetto a quelle non radicate, ai fini del restauro ambientale, negli USA si è rapidamente diffusa, per numerose specie, la produzione in vivaio di talee sia a radice nuda sia in contenitore (Hoag and Landis 2001).

L'utilizzo preminente delle tecniche di propagazione vegetativa, rispetto a quelle per via gamica, comporta una riduzione drastica degli scambi genici e determina una progressiva perdita di variabilità genetica dei popolamenti. Nel caso delle Salicaceae, come si è detto, ciò è aggravato dal fatto che la progenie di un individuo mantiene lo stesso sesso della pianta madre (Landis et al. 2003).

Un'altra causa di perdita di biodiversità è legata ai fenomeni di isolamento e speciazione. Le barriere geografiche come fiumi, ghiacciai, deserti o montagne e in certi casi la pressione antropica, sono le cause alla base della speciazione mediata dalla separazione fisica tra le popolazioni. Questo fenomeno è di maggiore rilevanza se il caso analizzato è un'isola (Grassi et al. 2006). Bush (1975) descrive due possibili casi di speciazione

allopatrica: il primo dove l'isolamento può avvenire a causa di barriere fisiche, che separano la popolazione in due sotto-popolazioni, abbattendo di fatto il flusso genico tra le piante. In questo caso ciascuna sottopopolazione evolve in maniera indipendente adattandosi al proprio ambiente. In questo caso le sottopopolazioni avranno la maggior parte della diversità genetica della popolazione di origine. Nel secondo caso ipotizzato da Bush (1975), la separazione avviene sempre a causa di una barriera fisica, ma in questo caso la sottopopolazione avrà dimensioni ridotte e formerà una nuova colonia che avrà solo una piccola parte della variabilità genetica della popolazione originaria. Le piante che andranno incontro a incroci tra parentali ridurranno la loro eterozigosità determinando la fissazione di numerosi caratteri. Esiste anche un altro tipo di speciazione detta simpatica, in cui una nuova specie ha origine all'interno dei confini geografici della sua forma parentale. In certi casi è possibile che si formino degli autoploidoidi fertili. Gli ibridi che si formano hanno un assetto cromosomico sbilanciato e quindi mostrano un certo livello di infertilità (Grassi et al. 2006). Il formarsi di un isolamento riproduttivo è chiaramente cruciale nella genesi di una nuova specie. Una volta che è avvenuta la speciazione, le due specie, a questo punto diverse, possono vivere nello stesso habitat senza incrociarsi.

1.7 Conservazione della biodiversità

Le classiche strategie d'intervento per la salvaguardia delle risorse genetiche, come definite dal Piano Nazionale sulla Biodiversità (D.M. 97/568 del 15 Maggio 1997) prevedono essenzialmente due approcci:

- la conservazione in situ;
- la conservazione ex situ.

L'obiettivo della conservazione in situ consiste nella reale possibilità di mantenere in collezione il maggior numero di alleli possibili e la massima diversità genetica permettendo, nel contempo, l'evoluzione del materiale genetico (Brush 1995). Quest'ultimo punto è di estrema importanza in quanto favorisce la ricombinazione genica grazie alla formazione di genotipi con caratteristiche adattative di risposta ai cambiamenti ambientali. La conservazione in situ, dunque, ha come fine il mantenimento della varietà di risorse genetiche, forme e processi adattativi delle specie e degli ecosistemi nello stesso ambiente in cui hanno avuto origine. L'insieme delle misure che concorrono alla conservazione in situ della diversità, oltre alle iniziative dirette di tutela e recupero del patrimonio naturale, devono essere finalizzate alla diminuzione della pressione sulla biodiversità da parte delle attività economiche, quali :

- contenimento dei fattori di impatto negativo sulla biodiversità da parte delle politiche settoriali;
- controllo del rischio di immissione negli ecosistemi di specie "aliene".

E' inoltre necessario creare le condizioni di consenso e coinvolgimento attivo delle comunità locali nella conservazione del patrimonio biologico presente sul territorio, attraverso:

- tutela delle "culture locali" che utilizzano in modo sostenibile il patrimonio naturale;
- sviluppo di attività economiche eco-sostenibili;
- educazione e sensibilizzazione all'importanza della conservazione della biodiversità.

La conservazione *ex situ*, invece, è rivolta a conservare e mantenere risorse genetiche fuori dal normale habitat di crescita, per esempio in una collezione di semi, o in una coltura di tessuto effettuata "in vitro" (Engelmann 1979), o anche in collezioni "in vivo" allestite in giardini e orti botanici, etc. Essa è una strategia fondamentale di conservazione della biodiversità quando questa è gravemente minacciata, oppure quando il numero degli individui di una specie è fortemente ridotto. Oltre a mantenere le risorse genetiche esistenti, tale conservazione è funzionale anche ad altri importanti obiettivi, quali: fornire popolazioni di riserva o stock da utilizzare per consentire la sopravvivenza delle specie durante le fasi di reintroduzione e ripopolamento; sviluppare nuove varietà, razze e ceppi durante i programmi di miglioramento genetico; fornire materiale per l'industria, per l'agricoltura, per la formazione e la ricerca; assicurare, attraverso lo stoccaggio a lunga scadenza, materiale per bisogni futuri; fornire materiale per la formazione di una coscienza ambientale. La conservazione *ex situ* presenta, tuttavia, lo svantaggio di mantenere solo una parte della variabilità genetica totale dei taxa, che pertanto può essere soggetta a deriva genetica. Da tutto ciò si deduce che un programma di conservazione della biodiversità deve prevedere un equilibrato bilanciamento dei due tipi di conservazione attuando una strategia volta a promuovere la biodiversità *ex situ* come complemento della conservazione della biodiversità *in situ*. In tempi recenti è maturata la consapevolezza che la piena integrazione delle due strategie, con l'ausilio di una serie di metodologie, tecniche e strumenti informatici, consente una più attenta e mirata valutazione delle risorse genetiche vegetali. Infatti, la disponibilità di nuovi e potenti strumenti informatici rende fattibile l'allestimento di banche di dati genetici e fenotipici, su base geografica. Tali strumenti di consultazione possono essere utilizzati nell'ottica di produrre una "mappa di reperibilità" di tali risorse genetiche in ambiente naturale e in collezioni, di contraddistinguere le località di reperimento e raccolta, di disporre di una prima caratterizzazione di ogni singola provenienza, di attribuire un "valore di conservazione" ad ogni popolazione/area/regione suscettibile di interventi di conservazione. In questo contesto, risulta innovativa la costituzione di banche del DNA, ovvero istituzioni normalmente legate a giardini botanici, Università, enti di ricerca o aree protette che si occupano dell'identificazione di specie caratteristiche o minacciate di una particolare area geografica, del loro campionamento e dell'estrazione del DNA, finalizzata sia alla conservazione dell'informazione genetica che al suo scambio.

Oggi, infatti, la biologia molecolare consente l'individuazione e la caratterizzazione della diversità genetica attraverso l'utilizzo di marcatori molecolari, che come descritto in precedenza (par. 1.2) permettono di analizzare l'intero genoma fornendo un sistema di analisi più potente e affidabile dei marcatori morfo-fisiologici e biochimici ampiamente impiegati nel passato. Attraverso l'uso di questi *tools* molecolari è possibile indagare e approfondire la variabilità tra le specie e addirittura tra gli individui permettendo una maggiore conoscenza delle risorse genetiche e la loro salvaguardia.

A livello internazionale aumentano ogni anno le attività rivolte alla salvaguardia e alla tutela della biodiversità così come richiede l'impegno sottoscritto dalle Nazioni Unite, in un momento in cui la perdita di biodiversità sul pianeta ha subito una formidabile accelerazione. In questo settore l'Italia ha percorso i tempi, infatti già nel 1970 ha inaugurato la prima Banca del Germoplasma delle varietà cereali. Diversi Paesi si sono impegnati nella realizzazione di banche del DNA per la conservazione del patrimonio genetico. Il caso più noto, anche grazie all'interesse che ha riscosso presso i media, è stato quello della Norvegia che, con un finanziamento governativo di 8 milioni di dollari donati dalla Fondazione Gates, ha realizzato nell'isola di Spitsbergen (arcipelago delle isole Svalbard situato in prossimità del circolo polare artico), un gigantesco deposito sotterraneo, la *Nordic Gene Bank*, dove custodire i semi di tutte le specie alimentari conosciute.

Ciò che rende innovativo l'approccio delle banche del DNA nel caratterizzare la biodiversità, risiede nel sistema d'integrazione delle informazioni morfologiche, ecologiche e genetiche delle risorse collezionate. Tra gli obiettivi che si propongono le banche del germoplasma ricordiamo quali sono i principali :

- completare l'inventario delle specie grazie all'ausilio di tecniche di sequenziamento del DNA;
- studiare le relazioni filogeografiche, prevedendo lo sviluppo dinamico e l'evoluzione degli ecosistemi in relazione ai cambiamenti climatici;
- produrre informazioni circa le collezioni di DNA, rendendole globalmente accessibili in rete nell'ambito di iniziative internazionali, quali "Global Biodiversity Information Facility" (<http://www.gbif.net>), "Species2000" (<http://www.species2000.org>) o di carattere regionale come "National Biodiversity Network" (<http://www.ukbiodiversity.net>).

Tra le diverse Banche di interesse internazionale del DNA vegetale si segnalano:

- *RGB Kew DNA Bank*, con oltre 22.000 campioni rappresentativi di una parte della flora mondiale;
- *Australian DNA Bank di Lismore* (NSW) con l'obiettivo di raccogliere campioni di DNA di tutta la flora australiana;
- *DNA Bank del Botanic Garden and Botanical Museum* (BGBM) di Berlino, inaugurata nell'aprile 2007, dedicata alle diverse specie vegetali coltivate e selvatiche di tutto il mondo;
- *Brazilian DNA Bank del Botanical Garden* di Rio de Janeiro, che si prefigge la raccolta delle oltre 1.000 specie minacciate della flora brasiliana;
- *DNA Bank di Kirtenbosch*, dedicata alla conservazione delle 2.200 piante rare o minacciate del Sud Africa;
- la Banca del DNA del CNR di Bari, inaugurata nel 2005 ed espressamente dedicata alle erbacee di interesse agronomico ad areale Mediterraneo.

1.7.1 Conservazione e mantenimento della biodiversità agraria

Esistono numerosi esempi di conservazione ex situ di germoplasma vegetale delle specie agrarie. Le più importanti collezioni mondiali sono mantenute dai Centri Internazionali di Ricerca Agricola (IARC) del Gruppo Consultivo di Ricerca Agricola Internazionale (CGIAR), che conservano complessivamente più di 680.000 campioni. La conservazione ex situ comporta anche attività di rigenerazione del materiale custodito nelle banche, la caratterizzazione e valutazione delle collezioni con l'impiego delle biotecnologie, la documentazione e la distribuzione agli utilizzatori.

Negli agrosistemi, mediante gli scambi di materiale genetico e l'interazione tra le forme di vita esistenti, continua l'evoluzione naturale e la Convenzione sulla Diversità Biologica, nell'articolo 8, indica la necessità della conservazione in situ e richiama a "...stabilire un sistema di aree protette, ovvero di aree dove sia necessario intraprendere provvedimenti speciali per conservare la diversità biologica..." e "...rispettare, salvaguardare e mantenere le conoscenze, le innovazioni e i metodi delle comunità indigene e locali, rafforzando quegli stili di vita tradizionali rilevanti per la conservazione e l'uso sostenibile della diversità biologica".

L'indicazione data è duplice: da una parte proteggere, con la creazione di parchi e riserve, gli ecosistemi naturali dove sopravvivono i progenitori selvatici di specie agrarie fondamentali per l'alimentazione umana, e, nelle aree coltivate, le varietà più rare;

dall'altra difendere la molteplicità colturale locale, grazie all'opera di tutti quegli agricoltori che, nel mondo, coltivano, conservano e trasmettono specie e varietà locali alle generazioni successive, tramandando con esse anche un grande patrimonio di conoscenze e tradizioni, definiti "agricoltori custode".

La biodiversità agraria rientra a pieno titolo nella voce cultura come "complesso di cognizioni, tradizioni, procedimenti tecnici, tipi di comportamento, trasmessi e usati sistematicamente, caratteristico di un popolo" (dal Vocabolario della lingua italiana, Nicola Zingarelli, X edizione).

La valorizzazione del patrimonio enogastronomico, che dimostra lo stretto rapporto con il territorio, il clima, le tradizioni popolari, può contribuire a conservare la biodiversità attraverso i prodotti tipici della tradizione agroalimentare. Il turismo enogastronomico, se progettato e gestito in maniera sostenibile e attenta, permette di sostenere le comunità locali e di salvaguardare la biodiversità agraria, valorizzando le tipicità e promuovendo l'integrazione tra le produzioni alimentari, il viaggio e il territorio (Howden et al. 2007).

La Regione Campania, grazie alla sua ricchezza colturale e alla presenza di ambienti geografici molto diversificati, è una delle regioni italiane provviste di un ampio e variegato germoplasma autoctono e, in particolare, di varietà ed ecotipi locali di specie ortive, frutticole e di vitigni. In Campania, nell'ultimo decennio, ad opera del settore della ricerca e degli enti e istituzioni locali, sono stati messi in atto numerosi programmi di recupero, caratterizzazione e valorizzazione di varietà ed ecotipi regionali sia di specie erbacee che arboree.

Il progetto AGRIGENET, nel quale rientra una parte di questo progetto di dottorato, prevede la realizzazione di una rete di ricerca per la salvaguardia e la gestione delle risorse genetiche vegetali campane, prendendo in considerazione varietà ed ecotipi di specie ortive, fruttifere e viticole. Le istituzioni scientifiche coinvolte, operanti nella Regione Campania sono:

- il Consiglio per la Ricerca e la Sperimentazione in Agricoltura (CRA, nelle sedi operative campane CRA-FRC, CRA-ORT e CRA-CAT);
- l'Istituto Nazionale di Ricerca per gli Alimenti e la Nutrizione (INRAN ex ENSE),
- l'Università degli Studi di Salerno (UNISA),
- la Seconda Università degli Studi di Napoli (SUN),
- l'Università degli Studi del Sannio (UNISANNIO).

Il progetto nel suo complesso mira al recupero della biodiversità vegetale, e in particolare di quella agraria, minacciata da erosione genetica. Si è così proceduto in primo luogo, alla individuazione, caratterizzazione e catalogazione delle risorse genetiche.

Gli ecotipi sono descritti mediante una scheda gestibile e aggiornabile anche a livello informatico in modo da costituire un inventario e una banca dati generale delle risorse genetiche vegetali campane.

Il progetto AGRIGENET ha previsto prove agronomiche, tipizzazione genetica tramite marcatori molecolari, determinazione di componenti ad attività nutraceutica e salutistica, caratterizzazione chimico-fisica e analisi sensoriale mediante "consumer test" e "panel test". Sono stati previsti e realizzati campi-catalogo di varietà ed ecotipi di specie orticole, collezioni di varietà ed ecotipi di fruttiferi e della vite per la conservazione in situ del germoplasma, definendo contestualmente anche i protocolli per la loro gestione nel lungo periodo e procedendo alla formazione di "agricoltori custodi".

In questo modo sono state integrate le due strategie di conservazione della biodiversità, in situ attraverso la formazione degli agricoltori custodi, cui siano affidati gli ecotipi individuati; ex situ attraverso la creazione di punti di raccolta e stoccaggio del DNA caratterizzato

molecolarmente, presso gli enti di ricerca della regione Campania, in modo da creare una rete di conservazione e tutela degli ecotipi orticoli campani.

1.7.2 Conservazione e mantenimento della biodiversità forestale

L'Italia è stata la prima nazione al mondo a istituire a Cittaducale (RI), nel Febbraio 2007, una banca del germoplasma di DNA dedicata alle sole specie boschive, oltre a essere stata anche tra i primi paesi a istituire una banca del DNA di specie cerealicole. La "banca centrale del DNA Forestale" rappresenta un progetto di salvaguardia moderno ed efficace per la conoscenza delle risorse genetiche di specie forestali, della loro distribuzione spaziale e dell'individuazione delle rarità e delle emergenze, nonché della loro descrizione biologica, ecologica e corologica.

Il fine ultimo, infatti, risiede nella possibilità di disporre per ogni specie forestale di una collezione di DNA il più possibile rappresentativa delle risorse genetiche presenti all'interno dei singoli popolamenti, fornendo altresì dati di tipo bioclimatico, edafico, ecologico, atti a descrivere l'ambiente ove le caratteristiche genetiche del popolamento si sono evolute.

La banca del DNA Forestale prevede la raccolta e la conservazione del patrimonio genetico (liofilizzato tissutale, DNA nucleare e organellare,) delle specie arboree presenti in una o più regioni (singoli individui e popolamenti), integrata con informazioni di tipo biologico, ecologico e molecolare e la messa a disposizione della comunità scientifica internazionale di dati e aliquote del materiale stesso. I criteri della rarità, della vulnerabilità e dell'endemicità sono sicuramente i primi a essere considerati nella scelta del materiale da salvaguardare e rigenerare, ma non i soli. La "Banca centrale del DNA Forestale" agisce anche da importante centro di studio per quanto concerne i taxa rappresentativi di habitat e per quelle entità ritenute fondamentali nella ricostituzione di aree degradate o fortemente compromesse.

Capitolo II

Materiali e Metodi

2.1 Campionamenti

2.1.1 Raccolta del materiale vegetale di ecotipi orticoli

L'aglio (*Allium sativum* L.) è una pianta coltivata bulbosa, assegnata tradizionalmente alla famiglia delle Liliaceae, ma che la recente classificazione APG III attribuisce alle Amaryllidaceae (sottofamiglia Allioideae) (Bremer et al. 2009). È una pianta erbacea perenne alta fino a un metro; da adulta presenta 40-60 radici cordiformi e superficiali, che interessano i primi 30 centimetri di terreno. I fiori sono sterili, e così anche i semi. L'unica via di propagazione si ha attraverso la divisione e piantumazione dei singoli bulbi; si ha quindi una riproduzione clonale e vegetativa (Keller and Senula 2013). Ogni bulbo contiene da 6 a 14 bulbilli o spicchi stretti fra loro e ricoperti da scaglie membranose. Lo spicchio, che rappresenta l'organo di moltiplicazione è attaccato direttamente al fusto. I bulbilli appena raccolti non sono in grado di germinare trovandosi infatti in uno stato di dormienza (Block 2009). Per germinare hanno necessità di passare attraverso una serie di stadi fisiologici legati alle condizioni climatiche. Le foglie sono basali, avvolgenti il fusto e, al contrario di ciò che avviene nella cipolla, non funzionano da organi di riserva. Le foglie si formano avvolgendosi l'una all'altra per un lungo tratto tanto che spesso sono scambiate come un fusto cilindrico. La parte della foglia non avvolgente ha la forma lineare e sono larghe fino a 3 cm terminanti con una punta acuminata e possono raggiungere una lunghezza totale di 80 cm.

La cipolla (*Allium cepa* L.) appartiene, come l'aglio, alla famiglia delle alle Amaryllidaceae secondo la classificazione APG III (Block 2009). È una pianta erbacea a ciclo biennale che diventa annuale in coltivazione, con radici superficiali e con delle foglie che nella parte basale si ingrossano dando la parte commestibile. Forma un lungo stelo florale che porta un'infiorescenza a ombrella con fiori di colore bianco-giallastro. Il frutto è una capsula. Possiede radici superficiali che giungono ad una profondità di non oltre i 40 cm, sviluppandosi in orizzontale per 20 o 25 cm dal fusto; il bulbo è commestibile, largo 1 o 2 cm e si trova sotto il livello del terreno. Mostra foglie alterne e opposte di 180°, esse sono costituite da due parti: una parte basale detta guaina e una lamina di consistenza carnosa, provvista di cera. La pianta produce fiori di colore bianco, giallo o verdognolo e roseo, mentre il frutto è una capsula trilobata che porta dai 2 ai 3 semi per lobo, questi hanno una forma irregolare, abbastanza piatti, di colore nero. I fiori sono di colore bianco, giallo-verdognolo o roseo e la fecondazione avviene principalmente ad opera degli insetti, soprattutto le api in quanto il polline (maschile) viene emesso molto prima che lo stilo (femminile) sia ricettivo (questo fenomeno si chiama proterandria) (Chrispeels M. and Sadava D. 1996).

Il mais (*Zea mays* L.) è una pianta erbacea annuale della famiglia delle Poaceae. È uno dei più importanti cereali largamente coltivato sia nelle regioni tropicali sia in quelle temperate, in quest'ultimo caso a ciclo autunno-primaverile. Il mais rappresenta la base alimentare tradizionale nelle popolazioni dell'America latina e, localmente, in alcune regioni

dell'Europa e del Nord America. Nelle regioni temperate è principalmente destinato all'alimentazione degli animali domestici, sotto forma di granella, farine o altri mangimi. Il mais è pianta monoica diclina: cioè i fiori maschili e femminili sono sulla stessa pianta portati da infiorescenze separate. L'infiorescenza femminile, che porta le cariossidi, si chiama correttamente *spiga* ma viene più spesso impropriamente chiamata "pannocchia", mentre l'infiorescenza maschile, posta sulla cima del fusto (*stocco*) della pianta è un panicolo terminale, costituito da numerose ramificazioni sulle quali si trovano le spighe; ogni spigetta consta di due fiori con tre stami ciascuno. Il mais è specie «proterandra» ossia la fioritura inizia con la deiscenza del polline dei fiori maschili del pennacchio, seguita poi dopo 2-3 giorni dall'emissione degli stigmi nelle infiorescenze femminili. L'emissione dei pennacchi non è contemporanea in un campo, ma si protrae per più giorni; anche la deiscenza del polline in un'infiorescenza dura qualche giorno. Gli stigmi, appena compaiono, sono suscettibili di essere fecondati e restano recettivi per il polline per parecchio tempo. Però, dato che l'antesi delle antere precede la comparsa degli stigmi, può darsi che gli ovuli della punta della spiga, gli ultimi a maturare, non arrivino ad essere fecondati per mancanza di polline. Nel mais la fecondazione incrociata è la regola: in condizioni normali si calcola che solo l'1% dei fiori si fecondino in autogamia (Kingsley R. Stern et al. 2009).

L'indivia (*Cichorium endivia* L.) è una pianta commestibile appartenente alla famiglia delle Asteraceae (o Composite). Si presenta fondamentalmente in due varietà: la varietà classica detta *crispum* e la varietà *latifolium*, comunemente chiamata scarola. Allo stato spontaneo è pianta biennale, che forma la rosetta fogliare (fase vegetativa) nel periodo autunno invernale ed emette lo scapo fiorale in primavera; come pianta coltivata si deve considerare annuale. La radice si approfondisce nel terreno con ramificazioni parallele, quasi perpendicolari che possono superare i 100 cm di profondità; tuttavia l'80% delle radici rimane nello strato di terreno fino a 20 cm, il 15% si trova tra i 20 ed i 50 cm e solo il 5% raggiunge le profondità più elevate. La rosetta fogliare è formata da 50-70 foglie a portamento più o meno eretto che presentano una nervatura centrale molto estesa di colore bianco ed un lembo fogliare tanto più involucente quanto più è sviluppato in larghezza; le foglie del cuore tendono ad autoeziolarsi soprattutto nelle cultivar più pregiate. I singoli fiori in numero di 15-25 per capolino sono ermafroditi, tutti ligulati, con corolla zigomorfa di colore variabile dall'azzurro pallido all'azzurro violaceo. Deve ritenersi specie autogama con possibilità di fecondazione incrociata; un accentuato eteromorfismo stilare è da ricollegarsi appunto al sistema di impollinazione entomofila. La fertilità, ovvero la capacità di produrre semi integri e germinabili, può risultare ridotta specie nelle piante isolate, per la formazione di frutti partenocarpici e di semi quiescenti; nelle prove di isolamento delle singole piante non si rintracciano piante sterili e mancano quindi meccanismi di auto incompatibilità (Kingsley R. Stern et al. 2009).

La zuccina o zucchini (*Cucurbita pepo* L.) è una specie della famiglia delle Cucurbitaceae. La pianta produce fiori monossuati. I fiori maschili sono sterili, ma necessari per l'impollinazione dei fiori femminili, a partire dai quali si sviluppa il frutto. L'impollinazione avviene per azione di insetti o del vento (vista la vicinanza sulla stessa pianta dei fiori maschili e femminili). Tale pianta annuale possiede un fusto di natura erbacea, particolarmente flessibile, con un portamento rampicante oppure strisciante. Le infiorescenze della zuccina sono monossuati: i fiori maschili, infatti, sono di natura sterili, ma ugualmente importanti per compiere il processo di impollinazione dei fiori femminili. Tale processo avviene, principalmente, mediante l'azione degli insetti oppure grazie al vento. Il fiore della zuccina costituisce, anch'esso, un alimento molto utilizzato in cucina: esso porta anche il nome di fiorillo. Il suo colore giallo-arancione, è spesso fritto dopo esser stato immerso nella pastella. Il frutto della pianta di zuccine si consuma sia in forma cruda che cotta ed è costituito da un cilindro color verde scuro, con leggere striature

biancastre in alcune varietà.

Nell'ambito del progetto AGRIGENET ci siamo occupati della caratterizzazione molecolare di ecotipi orticoli campani appartenenti a cinque diverse specie (Tab. 1):

Specie	Ecotipo/varietà
Aglione	Schiacciato
	Tondo di torella
Cipolla	Febbrarese
	Marzatica
	Ramata di Montoro
	Vatolla
Mais	Bianco di Acerra
	Spiga Bianca
	Spiga napoletana bianca (Montano)
	Spiga napoletana rossa (Montano)
	Spiga rossa
	Spiga bianca Napoletana
Scarola	Riccia schiana
Zucchini	Cilentano
	San Pasquale

Tab. 1. Elenco degli ecotipi orticoli campani caratterizzati attraverso metodiche molecolari.

Il materiale vegetale degli ecotipi sopra riportati è stato raccolto presso i campi sperimentali Eureco (Acerra, Na) e presso l'ENSE (Ente Nazionale Sementi Elette) di Battipaglia (Tab. 2).

Le foglie dei semenzali di ogni ecotipo è stato raccolto in quantità sufficiente e in modo tale da avere una quadro sufficientemente informativo della variabilità genetica sia intra- che inter-varietale. All'uopo per ogni varietà è stato prelevato materiale vegetale da circa 10 piante e poi essiccato in gel di silice così da mantenere tale materiale nelle condizioni ottimali per la successiva estrazione degli acidi nucleici.

Inoltre sono stati richiesti agli organi preposti alla conservazione *ex-situ* dei diversi ecotipi anche semi e/o bulbilli, al contrario nel caso degli ecotipi di aglio, mais e scarola, si è provveduto alla semina in camera climatica con temperatura e fotoperiodo controllati e disponibile presso il Dipartimento di Chimica e Biologia dell'Università di Salerno (Tab. 2).

Ecotipo	Luogo di raccolta		Metodo di conservazione	Condizioni di crescita	Estrazione DNA	Marcatore molecolare
Aglio "Tondo di torella"	Centro Cirio (2011)	ricerca Ex (EURECO)	gel di silice	Fotoperiodo controllato 12 ore di luce; T 25°C	Kit Qiagen C-TAB	SSR
Aglio "Schiacciato"	Centro Cirio (2011)	ricerca Ex (EURECO)	gel di silice	Fotoperiodo controllato 12 ore di luce; T 25°C	Kit Qiagen C-TAB	SSR
Cipolla "Febbrarese"	Centro BATTIPAGLIA (2012)	ricerca ENSE-	gel di silice		Kit SIGMA C-TAB	SSR
Cipolla "Marzatica"	Centro BATTIPAGLIA	ricerca ENSE-	gel di silice		Kit SIGMA C-TAB	SSR
Cipolla "Ramata di Montoro"	Centro BATTIPAGLIA	ricerca ENSE-	gel di silice		Kit SIGMA C-TAB	SSR
Cipolla "Vatolla"	Centro BATTIPAGLIA	ricerca ENSE-	gel di silice		Kit SIGMA C-TAB	SSR
Mais "Bianco di Acerra"	Centro Cirio (2011)	ricerca Ex (EURECO)	gel di silice	Fotoperiodo controllato 12 ore di luce; T 25°C	Kit Qiagen Kit Sigma C-TAB	SSR
Mais "Spiga Bianca"	Centro Cirio (2011)	ricerca Ex (EURECO)	gel di silice	Fotoperiodo controllato 12 ore di luce; T 25°C	Kit Qiagen Kit Sigma C-TAB	SSR
Mais "Spiga napoletana bianca"	Centro EXCirio (2011)	ricerca (EURECO)	gel di silice	Fotoperiodo controllato 12 ore di luce; ; T 25°C	Kit Qiagen Kit Sigma C-TAB	SSR
Mais "Spiga napoletana rossa"	Centro Cirio (2011)	ricerca Ex (EURECO)	gel di silice	Fotoperiodo controllato 12 ore di luce; T 25°C	Kit Qiagen C-TAB	SSR
Mais "Spiga rossa"	Centro Cirio (2011)	ricerca Ex (EURECO)	gel di silice	Fotoperiodo controllato 12 ore di luce; ; T 25°C	Kit Qiagen Kit Sigma C-TAB	SSR
Scarola "Riccia Schiana"	Centro Cirio (2011)	ricerca Ex (EURECO)	gel di silice	Fotoperiodo controllato 12 ore di luce; T 25°C	Kit Qiagen Kit Sigma C-TAB	SSR
Zucchini "Cilentano"	Centro Cirio (2011)	ricerca Ex (EURECO)	gel di silice		C-TAB Kit Sigma	SSR
Zucchini "San Pasquale"	Centro Cirio (2011)	ricerca Ex (EURECO)	gel di silice		C-TAB Kit Sigma	SSR

Tab. 2. La tabella riporta le informazioni riguardanti la conservazione dei semi e/o bulbi, le condizioni di crescita in camera climatica, le metodologie di estrazione del DNA e la metodica molecolare utilizzata per lo studio della biodiversità.

2.1.2 Raccolta del materiale vegetale di pioppo bianco

Populus alba L.

Comunemente chiamato pioppo bianco, è un albero caducifoglie della famiglia delle Salicaceae, alto più di 30 metri, con un'ampia chioma; è una specie centro europea meridionale, nativa della Spagna e del Marocco, il cui areale arriva fino all'Africa settentrionale e all'Asia centrale. Il suo habitat naturale è rappresentato da suoli incoerenti, sciolti limosi-argillosi, che rimangono umidi tutto l'anno ma senza subire regolari inondazioni, dove si associa a specie arboree, quali l'ontano, il frassino, l'olmo e il *Salix alba* L. In Italia si trova dalla pianura fino a circa 1500 m s.l.m. È abbastanza resistente alla salsedine. È più termofilo di altre specie del genere il tronco è eretto e sinuoso con corteccia bianco-grigia, liscia a solchi larghi neri, in giovane età, e rugosa in età avanzata; può raggiungere un diametro di 120 cm. Le foglie, sorrette da un picciolo depresso lateralmente lungo fino a 5 cm, hanno una forma ovale o rotondeggiante irregolarmente lobata (4-8 cm.). La pagina fogliare superiore è lucida, di colore verde scuro, mentre quella inferiore è ricoperta da una fitta peluria biancastra, da cui il nome comune della pianta (Camarda 1985). Sono distinte in turionali, le foglie con dimensioni maggiori che crescono sui rami più lunghi, e brachiblastali, quelle che crescono sui rami più corti, e quindi di piccole dimensioni. Il pioppo è una pianta dioica e per questo vi sono alberi di sesso femminile, i cui fiori sono provvisti di soli pistilli, e altri di sesso maschile, con soli stami. La fecondazione è anemofila (ad opera del vento).

La pianta dispone di un apparato radicale molto esteso che può diramarsi dalla pianta madre per oltre 20 metri. In aggiunta, la specie presenta una rapida crescita e grande biomassa, una buona resistenza agli inquinanti ed è accompagnata nel suo sviluppo da rapporti di simbiosi a livello radicale con particolari funghi, chiamati micorrize, che magnificano l'assorbimento di acqua e sali minerali dal suolo (Aikio S., Ruotsalainen A., 2002). Per queste formazioni simbiotiche il comune denominatore non è tanto il reciproco vantaggio, quanto la possibilità per un organismo di acquisire nuove capacità metaboliche dal proprio partner.

I campioni di *Populus alba* L. sono stati raccolti in tutta la regione Sardegna, in parte seguendo le coordinate geografiche riportate in Brundu et al. (2008), e in parte campionando nuove popolazioni di pioppo bianco mai analizzate in precedenza. Da ogni sito di campionamento sono state raccolte le foglie da almeno tre piante; per ogni pianta sono state prelevate sei foglie giovani della stessa età e dimensione. Per conservare memoria della morfologia e delle dimensioni, le foglie raccolte sono state poste su carta millimetrata in modo da procedere all'acquisizione delle immagini (Fig. 5). Inoltre per ogni sito di campionamento sono state registrate le coordinate geografiche e alcune note identificative dell'area di raccolta (Tab. 3).



Fig 5. La foto mostra le foglie di *P. alba* campionate e poste su carta millimetrata insieme ad un righello e al cartellino identificativo.

Le coordinate geografiche sono state importate nel software QGIS (v. 1.5) allo scopo di evidenziare i punti di campionamento su una mappa georeferenziata della Sardegna (Fig. 6).

Sigla Sito	Cod. cartellino foto	GPS Data	F	Zona	UTME (WGS84)	UTMN (WGS84)	Quota (m)	Sex
PER	PER-α1-Pos50	SC 09/04/11	32	T	491614	4520272	37,19	M(?)
PER	PER-α2-Pos51	SC 09/04/11	32	T	491648	4520290	39,01	M(?)
VID-MUNT-SP90	VID-MUNT-SP90-α3-Pos52	SC 09/04/11	32	T	488763	4532421	14,33	
VID-MUNT-SP90	VID-MUNT-SP90-α4-Pos53	SC 09/04/11	32	T	488758	4532419	14,02	
VID-MUNT-SP90	VID-MUNT-SP90-α5-Pos54	SC 09/04/11	32	T	488861	4532362	11,28	
VID-MUNT-SP90	VID-MUNT-SP90-α6-Pos55	SC 09/04/11	32	T	488874	4532372	8,53	
VIDD	VIDD-α7-Pos56	SC 09/04/11	32	T	491208	4528397	9,45	
VIDD	VIDD-α8-Pos57	SC 09/04/11	32	T	491231	4528417	-3,05	
FREG	FREG-α9-Pos58	SC 09/04/11	32	T	494757	4524140	60,05	

Tab. 3. La tabella riporta un esempio delle informazioni riguardanti i siti di campionamento.



Fig 6. La figura riporta una foto satellitare della regione Sardegna ottenuta tramite Google Earth, georeferenziata per poter aggiungere il *layer* dei punti di campionamento con il software QGIS.

2.2 Estrazione del DNA

Il DNA degli ecotipi orticoli campani è stato estratto attraverso tre metodiche. Inizialmente, per tutti i campioni è stata eseguita un'estrazione rapida del DNA mediante il kit commerciale *"Extract-N-Amp™ Plant PCR Kit"* (Sigma-Milano, IT) allo scopo di usarlo per la messa a punto delle condizioni di amplificazione in PCR. Quindi si è proceduto all'estrazione del DNA genomico sia attraverso il kit *"DNeasy Plant Mini Kit"* QIAGEN (Milano-Italia) che mediante il protocollo CTAB (Doyle and Doyle 1990). Tale protocollo prevede l'estrazione ripetuta del DNA mediante una miscela di cloroformio e isopropanolo, utile a denaturarlo e purificarlo dalle proteine ad esso associate.

L'estrazione e la purificazione del DNA genomico delle foglie di pioppo campionate in Sardegna è stata ottenuta mediante l'uso di un kit commerciale: *"DNeasy Plant Mini Kit"* (QIAGEN - Milano, IT) seguendo le istruzioni riportate dal manuale fornito dalla medesima

ditta fornitrice. Il DNA estratto e purificato è poi stato sottoposto a corsa elettroforetica su gel di agarosio 1,0%, e per stimarne la concentrazione il DNA da me purificato è stato confrontato con concentrazioni note (25 ng, 50 ng e 75 ng) di DNA, di *Fago lambda* (Fermentas – Milano, IT). In questo modo è stato possibile anche valutare l'integrità del DNA genomico estratto dai diversi ecotipi oggetto dell'indagine. Di seguito è riportata una foto della corsa elettroforetica a titolo esemplificativo (Fig. 7).

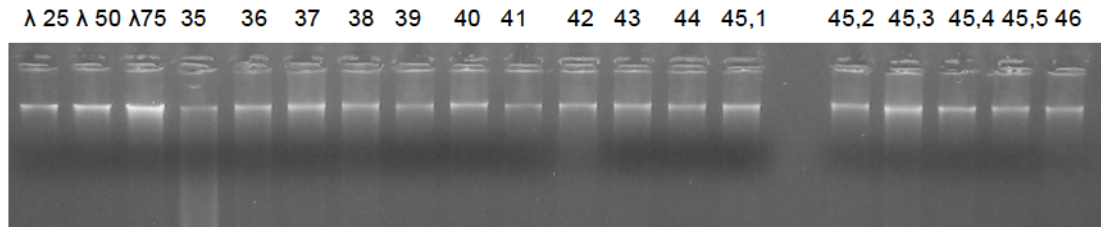


Fig. 7. L'immagine soprastante riporta, da sinistra a destra il DNA di fago λ a concentrazione nota (25 ng, 50 ng e 75 ng) e il DNA estratto da foglie di *P. alba* campionate in Sardegna.

Inoltre, la concentrazione e il grado di purezza dei DNA estratti sono stati valutati attraverso lo spettrofotometro UV/visibile "Picopet01" (Picodrop - UK). Il grado di purezza stimato come il rapporto tra l'assorbanza a 260 e 280 nanometri è risultato in tutti i casi ottimale essendo di circa 1,8.

2.3 ANALISI MOLECORARI

2.3.1 Analisi molecolari sugli ecotipi orticoli campani

Il DNA estratto secondo le metodiche indicate nel paragrafo 2.2 è stato amplificato usando come marcatori molecolari gli SSR. Grazie ad una attenta ricerca bibliografica sono stati individuati le coppie di *primer* SSR specifici per ogni specie e riportati in tabella 4.

Specie	Ecotipo	Marker	Primer FW	Primer RV	T. annealing °C
Aglio	Schiacciato	SSR	TTGGACTGAATT CTGAATACCT	GGGTGTGTGGTT CAAGGA	60
		SSR	CACAGCAACATG CACCAT	TGCCGGAACCTCG ATATT	60
	Tondo di torella	SSR	ACAAGGTCGAC ATCGTTTG	GGGCTTCACCTG AACACA	61
		SSR	CTTGCCGGAAC CGATATT	CACAGCAACATG CACCAT	61
		SSR	CACGCGAATCTT TCTTGG	TGCAAAGCAATAT GGCAG	60
		SSR	TGTTCCAACCAG ATTTAATGC	AAGTGGCGGTTG TGTCTG	60
		SSR	AATCTCCCTCCA AAGTCCC	CCTGTATTTTGTG TAAAGCATCA	60
		SSR	GGTCTCCTCATC CACCGT	GTGTGGGGCATG ATTGAC	60
		Cipolla	Febbrarese	SSR	TAACCCTAGGAT GAGTTGAG
Marzatica	SSR		GCCACGATGTT GAGATTTTCG	CCCGAATATCCCA CCAGTTC	52
Ramata di Montoro	SSR		ACCTTTAAATT GACGATATTCC	CTGCACTATTCTG TGATGTATTTC	57
Vatolla	SSR		ATGGGGCCTGG TAAGTTTT	TGCACACCGTTTC CATTTTA	59
	SSR		ACGGTTTGATGC ACAAGATG	CCAACCAACAGTT GATACTGC	54
SSR	TTTCAGCAACAT AGTATTGCGTC		TCTTCGGGATTG GTATGGAG	58	
SSR	CCACATGGATGA AAAACACAA		CGCTGGTAGCTG AAGCAAAT	60	
SSR	ACACACACAAGA GGGAAGGG		CACACACCCACA CACATCAA	55	
SSR	CCTTTGCTAACC AAATCCGA		CTTGTGAGAAG GAGGACGC	57	
SSR	CACCGTTCCTCA GCTCACTT		AGAGGGACGAAA TGAAAGCA	58	
SSR	TCTCATTCAAC TTTCTACCTATC C		CTGACATTTGCTC GACTGGA	60	
SSR	AAAACGAAGCAA CAAACAAA		CGACGAAGGTCA TAAGTAGGC	60	
SSR	CCCATTCTTCTT CATTTTCTCA		TGCTGTTGCTGTT GTTGTTG	60	
SSR	GCAGTTCTCCCT TTGTAAAATCA		GTGATGGATGAG TGGATGGA	60	
Mais	Bianco di Acerra		SSR	CTTCCAGCCGC AACCTC	CCAACAACGCGG ACGTGA
	Spiga Bianca	SSR	CAGATATCACAG CATCAGAAGATC A	AAAATGTAAGAAC TTGTTTGGGATT	60
	Spiga napoletana bianca (Montano)	SSR	GCACATGAAGAT CCTGCTGA	TGTGGATGACGG TGATGC	57
	Spiga napoletana rossa (Montano)	SSR	AATCAAACACAC ACCTTGCG	GCAAGGGAATAA GGTGACGA	57

Specie	Ecotipo	Marker	Primer FW	Primer RV	T. annealing °C
	Spiga rossa	SSR	CATGCAATCAATA	TTAGCGTAACC	65
			ACGATGGCGAGT	CTTCTCCAGTC AGC	
	Spogna Bianca	SSR	AGGCGGCGTGCT	CGCTTCATCTC	63
			GAACACCT	CCGTGACAATG	
		SSR	TGTTGCTCGGTCA	GCACACACACA	62
			CCATACC	GGACGACAGT	
SSR	AGATGACCAGGG	CCAGCTTCACC	68		
	CCGTCAACGAC	AGCTTGCTCTT CGTG			
SSR	ACTGTTCCACCAA	AGTAGGGGTTG	66		
	ACCAAGCCGAGA	GGGATCTCCTC C			
	TTGTCTTTCTTCCT	ATTTCCAGTTG			
SSR	CCACAAGCAGCGA	CCACCGACGAA	65		
	A	GAACT			
Scarola	Riccia schiana	SSR	CCAACGGATACCA	AACCGCACGG	58 (55)
			AGGTGTTT	GTTCTATG	
	SSR	CCGGCAGAATTTT	CAGGTCATAGG	56	
		TAGGG	TCCATGTGAAA		
	SSR	GTGCCGGTCTTCA	CGCCTACCGAT	54 (60)	
		GGTTACA	TACGATTGA		
	SSR	GGAGCAGGTAGA	CGTTTGAAAAT	58 (60)	
		GTCCCATC	TTATACCAAAA TG		
	SSR	TTCGAGTCTTGCC	CAGACGACCTT	59	
		TTAATTGTT	ACGGCAACT		
	SSR	CTGCTATGGACAG	CAATTCAGTTG	57(60)	
		TTCCAGT	TGATAGACGC		
	SSR	GCCATTCTTTCA	AACCCAAAACC	61(60)	
		AGAGCAG	GCAACAATA		
	SSR	GGCATCGGGATA	TCAATGCCTCA	52	
GAAAAACA		ACAGAAATCC			
SSR	AAAGTCACACATC	GTAGCAGCAG	55(60)		
	GCATTTCT	CAGCCATCTT			
SSR	AGCACGACTCTGC	CGAGCCATGTT	59(60)		
	TGTCTTTT	AGGGTTTGT			
SSR	CTCAACGAATGCT	CCTCGCGGTA	57		
	TTGGACA	GCTTATTGTT			
SSR	TATTGCATTGTTGT	TGTATTTAGAA	56		
	TCCTTG	GAGGGAAATAG ATG			
SSR	ATGTCGGAGCAAA	CATGTTCCCGC	57		
	ATCGTTC	TCATGAATA			
SSR	GGACACCGAGCT	TTCCACTTTCG	53(60)		
	GGAGAA	GGAGTTACC			
SSR	TGTAGACACACAA	ACCGGTTGAAA	55(60)		
	AATGCACA	ACATGAAAT			
SSR	TCGTCATCAGAAA	CAAAGAAGGCA	56		
	CAAAGCAA	CTCTTGTCG			
SSR	CCTACACTCGGCC	TCGACGGTATA	57		
	ACCTACT	ACAACACCTG			
SSR	TGTACAAACGTTG	ACAACACTGTCT	59		
	GATTGAA	CACTTCTCAC			
Zucchino	Cilentano	SSR	GGAGGAGGGATA	ACGATTTCTTC	63
	GAGAGAGAAACA		AGCCCTCAAAG		
San Pasquale	SSR	GTGGGCTAAGTTC	GAACCCAATCT	58	
		AAATCGTTC	TCTCATTTCCA		

Specie	Ecotipo	Marker	Primer FW	Primer RV	T. annealing °C
		SSR	CAAGCCTTTACCA AAACCAAAC	ACGATCTTCCTC CTCCTCTTCT	58
		SSR	GAACCTCGTTGCT GGTTCTTCT	TCTAGCATCTTA CGCACGCTTC	57
		SSR	TGCAAATTCGCT TCCATCATC	ACGACTGTGAG AACGGTGAAAA	57
		SSR	TGGAATCCAGAGA CTGATGAAG	CACGATTCCGA TAACAACAAGA	60
		SSR	TGGAGGTATCCTT CCGTATGTT	CATACGGGAGT TTCGTTTTTCT	60
		SSR	CTTCTCTACCGGA CCCTCCTAC	CTCAAGAGTTC CAACCTGAACC	60
		SSR	ATCCCAGGTCCCA ATTTTCTTC	CCATACCTGAG GGACCTGAAAC	57
		SSR	CGAATTTCTCCGA AAAGAAACC	CCTCGATCAAT CCTCCTACACC	58
		SSR	CGACTACAGCTCA ACAAAGACC	GAAGTCTATTA CGCCCAATTCC	58
		SSR	CGAGTACGTTTAC AACCATCCA	CAACTCTCAAC AAGAACCCAAG	57
		SSR	TCTTCTCGGAGTT TGATGTCCA	TTTACCCTCAT CGGAACCTCAT	58
		SSR	GACTACTACGAGC AAAATGACC	CGTGTGGAGCT TTCTCACAAC	57
		SSR	GCAGCACACTCGA TTCCTCTTT	TAGGAAAAATG GCGGTGAAGA A	64
		SSR	TCGCATTTGTAAA TTCCCTCT	ATGCTTTTCGAG GGATAAATGGA	60
		SSR	AGACTTTGGGTTT GTAGGAAGG	GCTAAGCATTG TTAGGGCTTCT	58
		SSR	CACGAGGAAGGA AGAATCGAAG	GAAATGCAAGA CCTCAAGCAGA	62
		SSR	CAAAACTGCCCAT GACCACTAC	CTTTCTACCCC AACCCACAT	57
		SSR	GACGAATACTTTC GCAATCAAG	CATGTGGAAGT TGTTGTTGTTG	62
		SSR	CTAATCAGGCTGG CAAAGAAGA	CGACGTTACCG ACAAGATTTC	60
		SSR	AGCACTGGGTAA CAGAGGAAA	ATGATACAGTA ATCGGCGTCCT	58
		SSR	CCTCTTCACCTTC GTCGTCTGT	GCTGCAATTTT CCCCTTCTTCT	58
		SSR	CCCATTCTCTCCA GCACTCTTT	AGGGAGGGAG ATGTGAGAGTT G	58
		SSR	TGTGATTTGTTGG GCTCTCTGT	ACATGGAATC CACCTCTTCGT	58
		SSR	TCCACACCAAGTGG TTCTTTCGT	TTTCACCATCC GGACAATCATC	58
		SSR	GCCGGTGATTCTG GATGAAGTA	CAGGTAGAAG GTGGGGAGAT GG	58
		SSR	TTGGACAATCTGA GGAAGTTGG	AAGGTTTGTCC TTCGACAAGGT	55
		SSR	TATGGTGGGTGG GTCTTTCAGT	GGAGTGTGCTT GATTGGGTTTC	56
		SSR	CCCTCGGTAGCAA TTTTGTAGT	TTTGGTGGGAG TGATACTGATG	54

Tabella 4. Dettaglio delle sequenze oligonucleotidiche disegnate e delle temperature di rinaturazione ottimale.

La ricerca bibliografica ha permesso di selezionare circa 10-15 coppie di *primer* altamente polimorfici. I *primer*, sintetizzati dalla ditta Primm (Milano-IT), sono stati saggiati su un limitato numero di campioni di DNA di ciascuna specie vegetale, oggetto di indagine molecolare, in modo da mettere a punto le migliori condizioni di amplificazione in PCR dei singoli loci SSR. Una volta raggiunto l'obiettivo si è proceduto all'amplificazione dei vari loci su tutti i campioni di DNA estratto in precedenza (Par. 2.2).

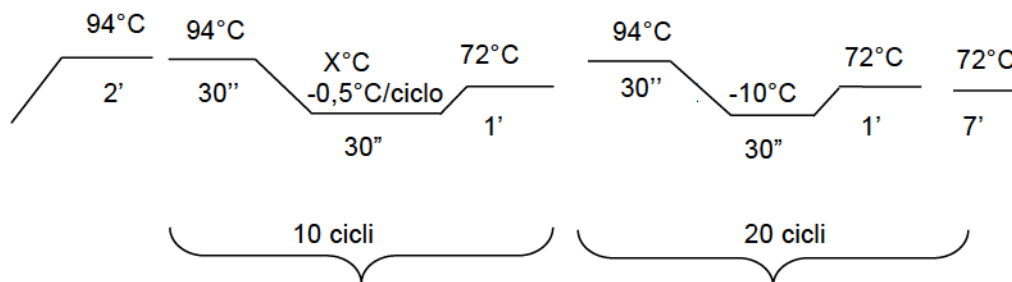
La reazione di PCR è stata allestita in un volume finale di 10 µl, preparando una miscela contenente i seguenti reagenti alle concentrazioni finali di:

- 0,2 µM *Primer Forward*
- 0,2 µM *Primer Reverse*
- 200 mM dNTPs
- 1X Buffer
- 1,0U Taq polimerase
- 10 ng DNA stampo

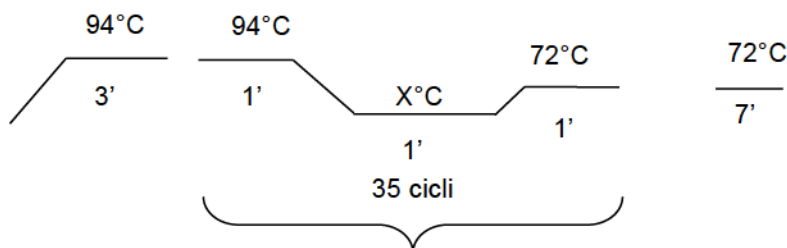
Le temperature di rinaturazione dei *primer* risultate ottimali per l'amplificazione sono riportate nella tabella precedente (Tab. 4), mentre i programmi di PCR utilizzati sono descritti di seguito (la temperatura di rinaturazione è indicata con X).

Programma PCR

1 -



2 -



I prodotti di PCR sono stati separati elettroforeticamente mediante sequenziatore automatico così da poter individuare i pesi molecolari degli alleli SSR amplificati. I dati

sono stati poi inseriti in una matrice e sottoposti ad analisi bioinformatica utilizzando specifici software.

2.3.2 MSAP

Gli MSAP sui DNA di pioppo sono stati ottenuti adattando il protocollo proposto da Reyna-Lopez et al. (1997). Esso si basa sull'amplificazione selettiva di frammenti di restrizione del DNA mediante PCR, ottenuti digerendo il DNA genomico, purificato secondo le metodiche sopradescritte, con due enzimi di restrizione, che riconoscono la stessa sequenza di taglio 5'-CCGG-3', ma sono diversamente sensibili alla metilazione delle citosine presenti nel sito di restrizione. I frammenti ottenuti dalla digestione sono poi "ligati" mediante l'enzima DNA Ligasi a una coppia di adattatori che forniscono i siti di legame per i *primer* utilizzati nella successiva amplificazione in PCR definita amplificazione preselettiva. Il prodotto dell'amplificazione preselettiva è infine utilizzato come substrato per un'ulteriore amplificazione in PCR definita amplificazione selettiva.

- Digestione

Per ogni campione di DNA sono state allestite due reazioni di digestione indipendenti e incubate per 6 h a 37°C come illustrato nella seguente tabella:

Digestione I	Digestione II
200 ng DNA	200 ng DNA
5 unità <i>MspI</i> (Fermentas)	5 unità <i>HpaII</i> (Fermentas)
5 unità <i>EcoRI</i> (Fermentas)	5 unità <i>EcoRI</i> (Fermentas)
Buffer 10X (Fermentas)	Buffer 10X (Fermentas)
H ₂ O q.b. per portare il volume a 40 µl	H ₂ O q.b. per portare il volume a 40 µl

Tab. 5. Tabella riassuntiva delle digestioni enzimatiche eseguite su DNA di *P. alba* per la metodica MSAP.

- Ligasi degli adattatori

Ai frammenti digeriti come da Tab. 5 sono poi ligati gli adattatori riportati in Tab. 6:

Nome primer	Sequenza di Riconoscimento
<i>EcoRI</i> adapter A	5'-CTCGTAGACTGCGTACC-3'
<i>EcoRI</i> adapter B	3'-CATCTGACGCATGGTTAA-5'
<i>MspI</i> - <i>HpaII</i> adapter A	5'-GATCATGAGTCCTGCT-3'
<i>MspI</i> - <i>HpaII</i> adapter B	3'-AGTACTCAGGACGAGC-5'

Tab. 6. Lista riassuntiva degli adattatori impiegati per la ligasi nella metodica MSAP.

Dopo aver preparato le soluzioni contenenti i due adattatori *EcoRI* (0,5 pmol/µl) e i due adattatori *MspI*-*HpaII* (5 pmol/µl), si è proceduto alla loro denaturazione e successiva rinaturazione mediante il seguente programma di PCR:

94°C per 3'; 70°C per 5", 60°C per 5", 50°C per 5", 40°C per 5", 25°C per 5", 14°C per 1', 4°C ∞

La reazione di ligazione prevede l'aggiunta ai 40 µl della miscela di digestione di:

0,5 pmol *EcoRI* adapters
 5,0 pmol *MspI*-*HpaII* adapters
 Buffer T4 10X (Fermentas)
 1,0 U ligasi T4 (Fermentas)
 H₂O q.b. per portare il volume a 50 µl

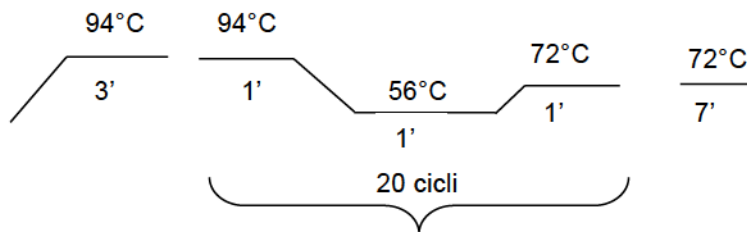
I campioni sono poi stati incubati a 37° C per le successive 3 h.
 Le reazioni di digestione e ligazione sono inattivate riscaldando le miscele a 65°C per 10'.

- **Amplificazione preselettiva**

Per ogni reazione di PCR occorre preparare una mix contenente:

2,5 µl digestione-ligasi
 dNTPs 200µM
 Buffer 1X
 MgCl₂ 1,0mM
 Primer Eco+1 37,5ng
 Primer MH+0 37,5ng
 Taq 1U
 H₂O q.b. per portare il volume a 50 µl

Il programma di amplificazione in PCR utilizzato è di seguito schematizzato:



I prodotti dell'amplificazione sono stati diluiti 1:20 in H₂O sterile e conservati a -20°C sino al loro impiego in PCR.

I *primer* usati per la reazione di amplificazione PCR preselettiva sono riportate in Tab. 7.

Nome primer	Sequenza di Riconoscimento
Eco+1	5'-GACTGCGTACCAATTC-3'
MH+0	5'-ATCATGAGTCCTGCTCGG-3'

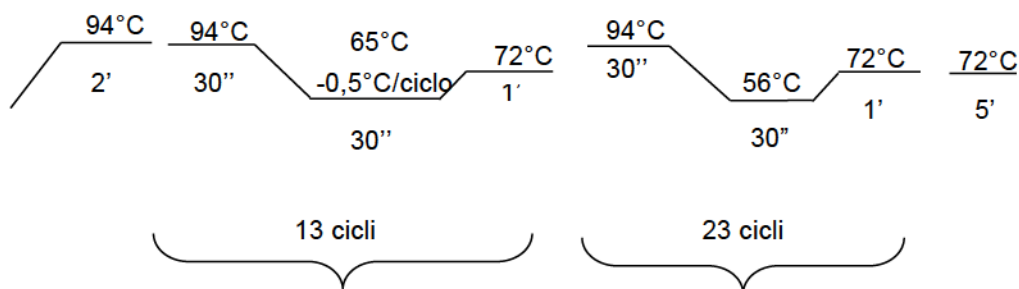
Tab. 7. Lista dei primer usati per l'amplificazione preselettiva dei DNA digeriti come sopradescritto.

- **Amplificazione selettiva**

Per ogni reazione di PCR è stata allestita una miscela di reazione contenente:

- 2,5µl prodotto di PCR (diluito in H₂O in rapporto 1 : 20)
- dNTPs 200µM
- Buffer 1X
- MgCl₂ 1mM
- Primer *Eco*RI fluorocromato 3ng
- Primer *Msp*I-*Hpa*II +3 o +4 15ng
- Taq 0,5 U
- H₂O q.b. per portare il volume a 10 µl

Il programma di amplificazione in PCR è qui di seguito riportato:



I primer *Eco*RI e *Msp*I-*Hpa*II utilizzati nella reazione di pre-amplificazione hanno un'estensione in posizione 3' di 1 o 2 basi puriniche o pirimidiniche rispetto agli adattatori, mentre quelli usati nell'amplificazione selettiva (Tab. 8) hanno un'estensione di 3 o 4 basi rispetto ai primer usati nella PCR preselettiva; ciò permette di avere un'ulteriore riduzione del numero di frammenti di DNA amplificabili. Utilizzando combinazioni di primer con diverse estensioni è stato possibile ottenere un numero notevole di frammenti MSAP che mi hanno permesso di indagare una frazione sufficientemente rappresentativa del genoma. Dopo aver effettuato uno screening tra le possibili combinazioni di primer utilizzabili per l'amplificazione selettiva, sono state impiegate le seguenti coppie di primer:

Nome primer	Sequenza di Riconoscimento
<i>EcoRI</i> + 3 AGC (FAM)	5'GACTGCGTACCAATTCAAGC3'
<i>EcoRI</i> + 3 ACC (HEX)	5'GACTGCGTACCAATTCAACC3'
<i>MspI</i> – <i>HpaII</i> AT	5'ATCATGAGTCCTGCTCGGAT3'
<i>MspI</i> – <i>HpaII</i> TGA	5'ATCATGAGTCCTGCTCGGTGA3'

Tab. 8. Lista delle sequenze di DNA dei *primer* utilizzati per la seconda amplificazione (amplificazione selettiva).

Un microlitro di ogni reazione di PCR selettiva è stato aggiunto a 10µl di formammide e denaturato a 94° per 3'. Grazie all'impiego dei *primer EcoRI* marcati con fluorocromi, i frammenti di DNA ottenuti dalla reazione di amplificazione selettiva presentano fluorescenza in posizione 3', che ne consente l'identificazione in base al suo peso molecolare. Ciò è possibile grazie all'impiego di un sequenziatore automatico ad elettroforesi capillare (Avant 310, Applied Biosystems-USA). La dimensione dei frammenti è determinata con l'ausilio di uno standard interno (Gene Scan 500 Rox Size Standard, Applied Biosystems).

- Corsa elettroforetica sul gel di poliacrilammide

Mediante la corsa su sequenziatore automatico sono stato in grado di individuare e annotare le bande amplificate ottenute con le digestioni *EcoRI* – *MspI* o *EcoRI* – *HpaII* ma ovviamente non è stato possibile ottenere informazioni sull'identità della sequenza o sulla loro posizione nel genoma di ogni singola specie vegetale analizzata. A questo scopo sono state allestite delle corse su gel di poliacrilammide. Obiettivo principale, sia delle corse elettroforetiche capillari che di quelle su gel di poliacrilammide, è stato l'individuazione di bande di DNA polimorfiche tra medesimi cloni di pioppo, ovvero individui che presentano genotipi identici, cresciuti in condizioni ambientali differenti e a chilometri di distanza. Il vantaggio offerto dal gel di poliacrilammide consiste nella possibilità di prelevare, purificare, clonare e sequenziare le bande risultate polimorfiche così da poter valutare se e quali geni, o sequenze regolative ad essi associati, siano modulati dai meccanismi di metilazione del DNA.

2.4 Analisi bioinformatica

I frammenti di DNA polimorfici da me ottenuti, sia che essi fossero alleli SSR, o frammenti MSAP e AFLP, sono state codificati per presenza-assenza in una matrice rappresentativa del profilo genetico di ogni singolo individuo analizzato.

2.4.1 Analisi dei marcatori SSR e MSAP

L'analisi dei dati SSR e MSAP è stata realizzata tramite il software Gene Mapper v. 3.7 (Applied Biosystems - USA) che permette di determinare, per confronto con un marcatore di peso molecolare noto, l'esatta dimensione dei prodotti dell'amplificazione.

I pesi molecolari dei frammenti SSR amplificati sono stati annotati in un foglio di calcolo, mentre per i frammenti ottenuti dalla metodica MSAP per i pioppi bianchi si è proceduto alla costruzione di una matrice binaria di presenza/assenza dei frammenti amplificati. Entrambe le matrici sono state usate per le elaborazioni biostatistiche e per il calcolo degli indici di similarità genetica. In particolare sui dati elaborati sulla base delle amplificazioni in PCR dei campioni digeriti con gli enzimi *EcoRI* e *MspI* è stato possibile ricavare informazioni sulla biodiversità genetica delle popolazioni di pioppi, e quindi

valutare la similarità tra i diversi genotipi raccolti e le differenti popolazioni. I dati elaborati sono stati rappresentati attraverso la costruzione di dendrogrammi. Il metodo più comune per l'elaborazione grafica del livello di similarità tra individui, ovvero i dendrogrammi, si basa sulle distanze genetiche di tipo probabilistico ed è noto come UPGMA (*Unweighted Pair-Group Method with Arithmetic averaging*). L'UPGMA unisce gli elementi a coppie usando medie aritmetiche non pesate, analizzando la matrice e raggruppando elementi simili (o dissimili tra loro), al fine di suddividere la popolazione in gruppi sufficientemente omogenei (*cluster*). A questo scopo dai dati della distanza genetica dei campioni è stata elaborata una matrice quadrata di similarità utilizzando il coefficiente di Jaccard (Jaccard 1901), disponibile con il software NTSYSpc version 2.1 (Exeter software-NY-USA).

Percentuale di polimorfismo

Per ogni sito di campionamento e per ogni genotipo individuato è stata calcolata la percentuale di polimorfismo che varia nell'intervallo 0-100%:

$$P\% = (n/T) \times 100$$

n è il numero di loci polimorfici e T il numero di loci analizzati.

Grazie a queste elaborazioni statistiche abbiamo stimato il livello di biodiversità genetica delle popolazioni di pioppo bianco campionate in Sardegna.

Sull'insieme dei dati molecolari ottenuti sono applicate tecniche statistiche multivariate. In particolare sia per le elaborazioni dei dati molecolari ottenuti dalle analisi genetiche delle popolazioni di pioppo bianco che su quelli derivanti dalle indagini sugli ecotipi orticoli campani sono stati calcolati gli indici riportati di seguito utilizzando software *ad hoc* quali: *Genalex* v. 6 (Peakall and Smouse 2006) *Power Marker* (Liu and Muse 2005.) e *Structure* (Pritchard et al. 2000). Di seguito sono riportati le descrizioni degli indici calcolati.

Stima del livello di clonalità (intervallo 0 – 1)

Attraverso l'equazione di seguito riportata è stata stimata la probabilità che in una popolazione naturale, sottoposta a riproduzione sessuata e casuale, si possa ritrovare un identico genotipo:

$$P_{gen} = (\prod p_i) 2^h$$

dove p_i è la frequenza di ogni singolo allele per locus e h il numero di loci eterozigoti.

Polymorphic Index Content (PIC)

Il calcolo del PIC (Botstein et al. 1980) è stato realizzato tramite il software *PowerMarker*. Questo indice è una misura della diversità genetica basata sul livello di polimorfismo dei loci stessi, calcolato attraverso la seguente equazione matematica:

$$\widehat{PIC}_t = 1 - \sum_{u=1}^k \tilde{p}_{iu}^2 - \sum_{u=1}^{k-1} \sum_{v=u+1}^k 2\tilde{p}_{iu}^2 \tilde{p}_{iv}^2$$

dove p_{lu} e p_{lv} sono rispettivamente la frequenza dell' u -esimo e del v -esimo allele per il locus l .

Inferenza del numero di popolazioni

Per stimare il numero di popolazioni costituenti il pool di campioni analizzati è stato usato il software *Structure*. Il software fornisce l'inferenza del numero di popolazioni distinte e l'assegnazione di ogni singolo individuo oggetto di indagine a una specifica popolazione. Questo programma prevede l'utilizzo del metodo statistico Monte Carlo basato sulla catena di Markov (MCMC) (Diaconis 2009) per il campionamento di distribuzioni di probabilità in cui lo stato della catena, dopo un grande numero di iterazioni, è usato come un campione della distribuzione desiderata. Il numero di popolazioni è stato calcolato sulla base di questa distribuzione di probabilità e con un'inferenza di 20 iterazioni.

I parametri da me impiegati per l'analisi *Structure* prevedono 10 corse indipendenti con K compreso tra 1-10, utilizzando 10^5 come lunghezza della MCMC e 10^3 come periodo di non campionamento ("burn-in").

I modelli matematici usati si fondono su e assumono le seguenti proprietà:

1. "admixture" (mescolamento), consente di attribuire un singolo individuo a più di una popolazione tra quelle individuate;
2. modello per le frequenze alleliche correlate: assume che per alcune popolazioni, derivate dalla popolazione originaria, le frequenze alleliche di ogni popolazione individuata siano correlate con quelle della popolazione ancestrale;
3. non sono state fornite informazioni di partenza sulle popolazioni, ovvero sono stati considerati tutti gli individui come appartenenti a un unico pool genico, senza popolazioni definite *a priori*.

La distribuzione di probabilità è stata calcolata dopo un *burn-in period* (periodo dopo il quale il software campiona i dati riducendo gli effetti della configurazione iniziale) pari a 1000 rimescolamenti e su un numero di simulazioni basate su MCMC pari a 100'000 (considerato adeguato per ottenere parametri accurati). Il numero di popolazioni è stato inferito utilizzando il metodo proposto da Evanno et al. (2005) basato sul calcolo del logaritmo medio di vero somiglianza dei dati calcolati per ogni iterazione (10).

2.4.2 Stima della differenziazione delle popolazioni (Intervallo da 0 - 1):

Questo indice può essere calcolato attraverso vari algoritmi. Tra i più conosciuti e usati va annoverato l' F_{st} (Wright 1946). Tuttavia sono stati sviluppati ulteriori indici dello stato di differenziazione delle popolazioni e degli individui tra cui il G_{st} (Nei and Chesser 1983; Meirmans and Hedrick 2011) e il D_{est} via G_{st} (Jost 2008). Dato che in questo progetto è stato affrontato lo studio della biodiversità sia di popolazioni naturali di pioppo bianco che di ecotipi orticoli campani, ottenuti mediante ancestrale selezione artificiale, e considerate le proprietà dei campioni (es. numero dei campioni) e le differenti metodiche molecolari usate (SSR e MSAP), sono stati impiegati due indici differenti per la stima della differenziazione delle popolazioni o degli ecotipi. Per quanto riguarda gli ecotipi orticoli campani l'indice di differenziazione scelto e che distorce meno il risultato ottenuto dall'analisi molecolare dei campioni (dato il ridotto numero di campioni e la frammentazione delle popolazioni) è il D_{est} descritto dalla seguente equazione:

$$D_{est} = \left(\frac{k}{k-1} \right) \left(\frac{cH_T - cH_S}{1 - cH_S} \right)$$

H_T = eterozigosità totale stimata delle sottopopolazioni corretta

H_s = eterozigosità stimata all'interno delle sottopopolazioni;
 K = numero di popolazioni.

Mentre per l'elaborazione dei dati molecolari dominanti ottenuti dallo studio della biodiversità del pioppo bianco è stato utilizzato l' F_{st} . Tale scelta è stata determinata dal numero di popolazioni indagate e dal basso livello di biodiversità genetica atteso. Di seguito è riportata un'immagine (Fig. 8) che descrive i comportamenti dei diversi indici al variare del numero di popolazioni (Whitlock 2011).

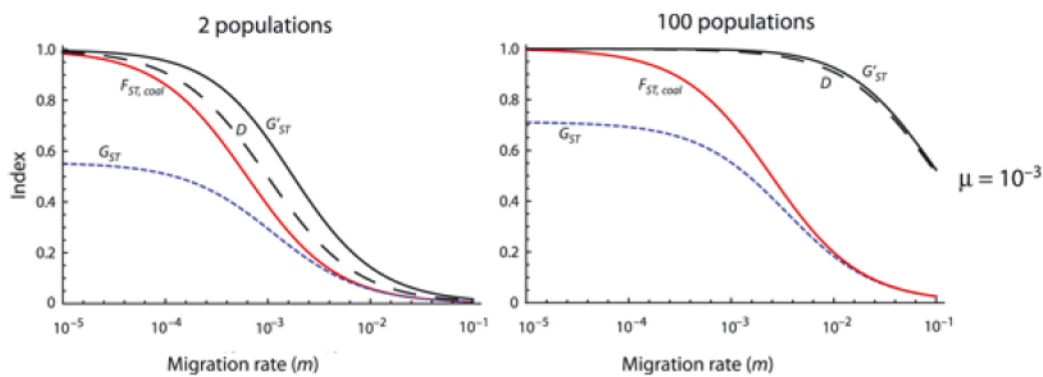


Fig.

8. L'immagine descrive il comportamento dei diversi indici presi in considerazione al variare del numero di popolazione e del numero di migranti.

2.4.3 Analisi della Varianza Molecolare (AMOVA)

Le matrici dei dati genetici sono state analizzate attraverso la metodica statistica AMOVA allo scopo di valutare il grado di diversità genetica intra- e inter- popolazione (Excoffier et al. 1992; Peakall et al. 1995; Michalakis and Excoffier 1996). Questa analisi crea una matrice di distanza genetica tra i campioni, e in base a tale distanza il software calcola le distanze genetiche intra popolazione e tra le popolazioni. Per limitare l'effetto del campionamento casuale effettuato e che potrebbe alterare il risultato finale, sono state imposte una serie di permutazioni dei dati che ricalcolano la varianza tra le popolazioni, senza rispettare la struttura originale dei dati per un numero prestabilito di volte (in questo caso 9999); nel caso in cui i valori generati in seguito alle permutazioni comprendono il valore reale, allora si può asserire che si è di fronte a una popolazione; ma se invece i valori ottenuti non comprendono quello reale, allora le popolazioni si possono considerare tra loro distinte. Il test eseguito con il software *Genalex* restituisce come risultato il valore $P(rand \geq data)$. Questo valore va inteso come la probabilità che un valore casuale generato dalle permutazioni sia maggiore o uguale a quello osservato. Nel caso ciò accada per un numero di volte maggiore o uguale al 95%, l'ipotesi nulla che esista una sola popolazione è rigettata.

2.4.5 Analisi marcatori microsatelliti degli ecotipi orticoli campani

La biodiversità genetica-molecolare degli ecotipi orticoli campani, essendo stata analizzata mediante marcatori codominanti (SSR) è stata studiata anche attraverso la stima del numero di alleli per singolo locus, alla frequenza allelica e alla percentuale di eterozigosi (H), sia osservata (H_o) che attesa (H_e). In particolare, il valore dell'eterozigosi attesa (H_e), fondata sulla legge di Hardy-Weinberg, ma comunemente impiegato anche per la stima

della biodiversità genetica di varietà, ed ecotipi ottenuti per selezione artificiale, è una misura che valuta tale variabilità in una popolazione in base alla probabilità che, estraendo casualmente due alleli dalla popolazione, questi siano diversi.

Le equazioni utilizzate per queste ulteriori elaborazioni bioinformatiche sono state realizzate tramite il software *Genalex* v. 6,5 e sono qui di seguito riportate:

Eterozigosità attesa o Diversità genica (che varia nell'intervallo 0-1):

$$H_e = 1 - \sum p_i^2$$

p_i è la frequenza dell'*i*-esimo allele a ciascun locus analizzato;

Eterozigosità attesa corretta (intervallo 0-1):

$${}^uH_E = \frac{2n}{2n-1} \left(1 - \sum p_i^2 \right)$$

dove: n è il numero di campioni e p_i è la frequenza dell'*i*-esimo allele.

Eterozigosità osservata (che varia nell'intervallo 0-1):

$$H_o = H/N$$

H è il numero di eterozigoti e N è il numero di campioni nella popolazione;

Information Index, equivalente all'indice Shannon-Weaver usato in ecologia:

$$I = \sum p_i \ln p_i$$

dove \ln è il logaritmo naturale e p_i è la frequenza dell'*i*-esimo allele.

Indice di fissazione (intervallo -1;1)

$$F = \frac{H_E - H_o}{H_E}$$

dove H_E è l'eterozigosità attesa e H_o è l'eterozigosità osservata.

2.4.6 Analisi bioinformatica per popolazioni naturali di *P. alba* *Landscape genetics*

Lo studio della biodiversità del pioppo bianco è stata elaborata anche in funzione della posizione geografica attraverso il software Geneland (Guillot et al. 2011). Questo software è stato sviluppato grazie a una routine creata in R che permette l'elaborazione dei dati genetici associati a quelli geografici per delineare la struttura della popolazione e la sua

organizzazione spaziale e paesaggistica. L'obiettivo principale della *Landscape genetics* è quello di individuare la struttura della popolazione in forma di variazione sistematica della frequenza degli alleli partendo dall'assunzione che le popolazioni rispettino il principio di Hardy-Weinberg. Anche in questo caso, come in *Structure*, l'algoritmo è basato sul rimescolamento casuale dei campioni secondo il modello MCMC. Grazie alle coordinate geografiche è possibile visualizzare su una *mappa a posteriori* la probabilità di appartenenza dell'individuo indagato alla popolazione, importando il risultato statistico direttamente in *Google Maps* o in software GIS dedicati come *Grass*. E' stata inoltre calcolata una matrice di distanza genetica in base ai profili genetici e a quelli di metilazione allo scopo disegnare degli "alberi delle connessioni ottimali" (*minimum spanning tree*) (Gower and Ross 1969) che riflettano le connessioni genetiche tra le varie popolazioni in funzione delle coordinate spaziali. Anche in questo caso questa rete di connessioni ottenuta tra le popolazioni è stata sovra-imposta alla mappa della regione Sardegna.

MSAP

Il programma *msap* (Pérez-Figueroa 2013) fornisce un'analisi di variazione epigenetica partendo da una matrice di dati binari contenente i profili ottenuti con la digestione dei due enzimi di restrizione isoschizomeri *HpaII* e *MspI*, con sensibilità differenziale alla metilazione della citosina. Dopo aver confrontato i frammenti di restrizione, il programma determina se ogni frammento è suscettibile di metilazione e determina anche la variabilità genetica elaborando la combinazione *EcoRI* – *MspI*. Il pacchetto prevede, in una interfaccia a riga di comando *user-friendly*, una *pipeline* di diverse analisi della variazione (genetica ed epigenetica) tra i gruppi definiti dall'utente, nonché la classificazione degli eventi di metilazione di questi gruppi. La presenza dei frammenti in entrambe le digestioni amplificate (*EcoRI* – *HpaII* e *EcoRI* – *MspI*) denota uno stato non metilato; presenza prodotti dei in solo *EcoRI* – *HpaII* (1 / 0) o solo *EcoRI* – *MspI* (0 / 1) rappresenta stati metilati (emimetilazione o metilazione della C interna), mentre assenza sia in *EcoRI* – *HpaII* e *EcoRI* – *MspI* (0 / 0) è considerato come uno stato non informativo.

Capitolo III

Risultati

3.1 Ecotipi orticoli Campani – Simple Sequence Repeat (SSR)

Gli ecotipi orticoli campani sono stati indagati mediante metodiche molecolari allo scopo di valutarne la biodiversità e tracciarne un profilo genetico; in particolare sono stati usati dei marcatori codominanti SSR quali gli SSR (Simple Sequence Repeat). L'analisi molecolare con gli SSR ha riguardato gli ecotipi orticoli appartenenti alle seguenti specie: aglio (*Allium sativum* L.), cipolla (*Allium cepa* L.), mais (*Zea mays* L.), scarola (*Cichorium endivia* L.) e zucchino (*Cucurbita pepo* L.). Il DNA estratto è stato amplificato in PCR con diverse coppie di oligonucleotidi, utilizzando, per ciascuna coppia, le condizioni di amplificazione ottimali. Dall'analisi condotta sui profili SSR ottenuti dalla corsa elettroforetica su sequenziatore automatico capillare è stato possibile ricavare le dimensioni di ciascun allele e distinguere tra la condizione di omozigosi e quella di eterozigosi per ciascun locus SSR indagato (un esempio è dato in Fig. 9).

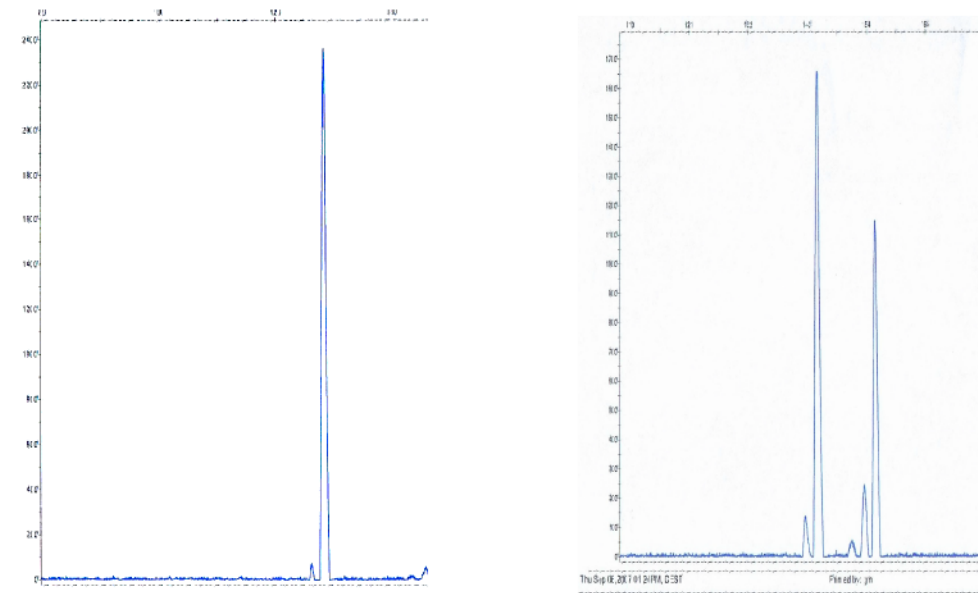


Fig. 9. L'immagine riporta l'esempio di un'amplificazione di un locus SSR: a sinistra profilo omozigote e a destra quello eterozigote.

3.1.1 Aglio

Per i due ecotipi di aglio (*"Schiacciato"* e *"Tondo di torella"*) sono state utilizzate otto coppie di *primer* utilizzando il programma di amplificazione ottimale. Di seguito (Fig. 10), è riportata un'immagine esemplificativa della corsa elettroforetica su gel di agarosio del DNA estratto dai campioni di aglio.

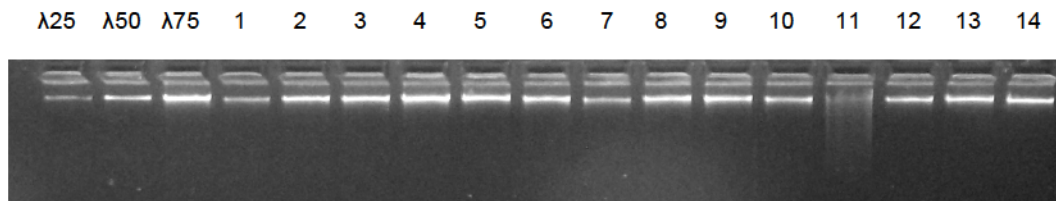


Fig. 10. Corsa elettroforetica su gel di agarosio del DNA genomico estratto da aglio. λ 25, λ 50, λ 75 corrispondono al DNA di fago lambda rispettivamente di 25, 50 e 75 nanogrammi; i pozzetti da 1 a 7 corrispondono ai primi 7 campioni di DNA estratto da aglio "Schiacciato" mentre i pozzetti 8 al 14 corrispondono ai primi 7 campioni di DNA estratto da aglio "Tondo di Torella".

Il DNA genomico di aglio è stato amplificato in PCR con i *primer* riportati nella tabella 10 di Materiale e Metodi (Par. 2.3.1) e i prodotti dell'amplificazione sono stati sottoposti a corsa elettroforetica su gel di agarosio per verificare il risultato dell'amplificazione (Fig. 11).

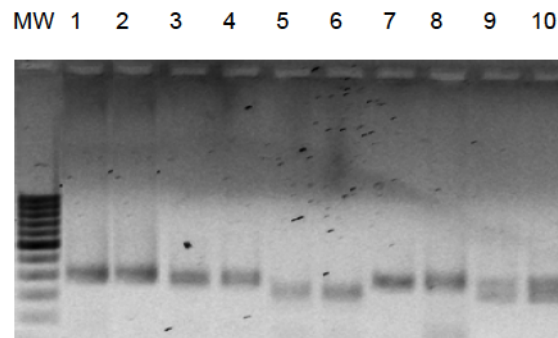


Fig.11. Immagine della corsa elettroforetica su gel di agarosio dei prodotti di PCR ottenuti dalle amplificazioni delle regioni SSR del DNA di aglio: MW marcatore di peso molecolare; da 1 a 5 campione aglio "Schiacciato", da 6 a 10 campione aglio "Tondo di Torella".

Accertata l'amplificazione, i campioni sono stati analizzati mediante corsa elettroforetica capillare. La capacità discriminante degli SSR, come pure di altri marcatori molecolari, è legata alle variazioni di dimensioni degli alleli o dei frammenti di DNA analizzati. I pesi molecolari di ciascun allele per ciascun locus SSR sono stati riportati in un foglio di calcolo e sottoposti a elaborazioni bioinformatiche. I risultati di tali elaborazioni (Par. 2.4.5) sono qui di seguito riportati (Tab. 9):

Pop	Locus	N	Na	Ne	I	Ho	He	UHe	F	GeneDiversity	PIC
"Schiacciato"	Locus1	10	2	1,98	0,69	0,90	0,50	0,52	-0,82	0,5	0,375
	Locus2	10	3	2,20	0,86	1,00	0,55	0,57	-0,83	0,5	0,375
	Locus3	10	2	2,00	0,69	1,00	0,50	0,53	-1,00	0,5	0,375
	Locus4	10	2	2,00	0,69	1,00	0,50	0,53	-1,00	0,5	0,375
	Locus5	10	2	2,00	0,69	1,00	0,50	0,53	-1,00	0,5	0,375
	Locus6	10	2	2,00	0,69	1,00	0,50	0,53	-1,00	0,5	0,375
	Locus7	10	2	2,00	0,69	1,00	0,50	0,53	-1,00	0,5	0,375
	Locus8	10	2	2,00	0,69	1,00	0,50	0,53	-1,00	0,5	0,375
"Tondo di torella"	Locus1	5	2	1,92	0,67	0,80	0,48	0,53	-0,67	0,5	0,375
	Locus2	5	2	2,00	0,69	1,00	0,50	0,56	-1,00	0,5	0,375
	Locus3	5	2	2,00	0,69	1,00	0,50	0,56	-1,00	0,5	0,375
	Locus4	5	2	2,00	0,69	1,00	0,50	0,56	-1,00	0,5	0,375
	Locus5	5	2	2,00	0,69	1,00	0,50	0,56	-1,00	0,5	0,375
	Locus6	5	2	2,00	0,69	1,00	0,50	0,56	-1,00	0,5	0,375
	Locus7	5	2	2,00	0,69	1,00	0,50	0,56	-1,00	0,5	0,375
	Locus8	5	2	2,00	0,69	1,00	0,50	0,56	-1,00	0,5	0,375

Tab. 9. I valori riportati in tabella corrispondono a: N=Numero di campioni, Na=numero di alleli, Ne= numero di alleli effettivi per locus, I=indice di Shannon, H_o = eterozigosità osservata, H_e = eterozigosità attesa, UHe = eterozigosità attesa corretta, F= indice di fissazione, PIC = Polymorphic Index Content.

I valori dell'eterozigosità osservata e dell'indice di fissazione suggeriscono che per tutti gli otto loci vi sia un eccesso di eterozigosità dovuta al fatto che gli ecotipi di aglio studiato potrebbero essere degli ibridi ottenuti con tecniche di selezione basate su eterosi. Nonostante l'elevata eterozigosità il valore di D_{est} risulta pari a 0, indice di assenza di diversità nella frequenza allelica delle due popolazioni.

Sulla matrice dei pesi molecolari è stato eseguito un test statistico allo scopo di verificare la separazione dei campioni in due popolazioni differenti. Il risultato ha evidenziato come i due ecotipi di aglio non siano discriminabili mediante questi otto loci SSR in quanto le singole piante sono attribuibili all'ecotipo "Schiacciato" o "Tondo di Torella" con una percentuale del 50%; infatti anche l'indice di diversità genetica, che indica quale è la probabilità che due alleli campionati in una popolazione siano differenti, risulta essere pari a 0,5.

La difficoltà di separare i due ecotipi in base ai loro profili SSR emerge dal dendrogramma costruito in base alla matrice di similarità, calcolata secondo l'indice di Jaccard (Fig. 12). Dalle analisi molecolari effettuate è emerso che i due ecotipi sono identici per questi otto loci analizzati.

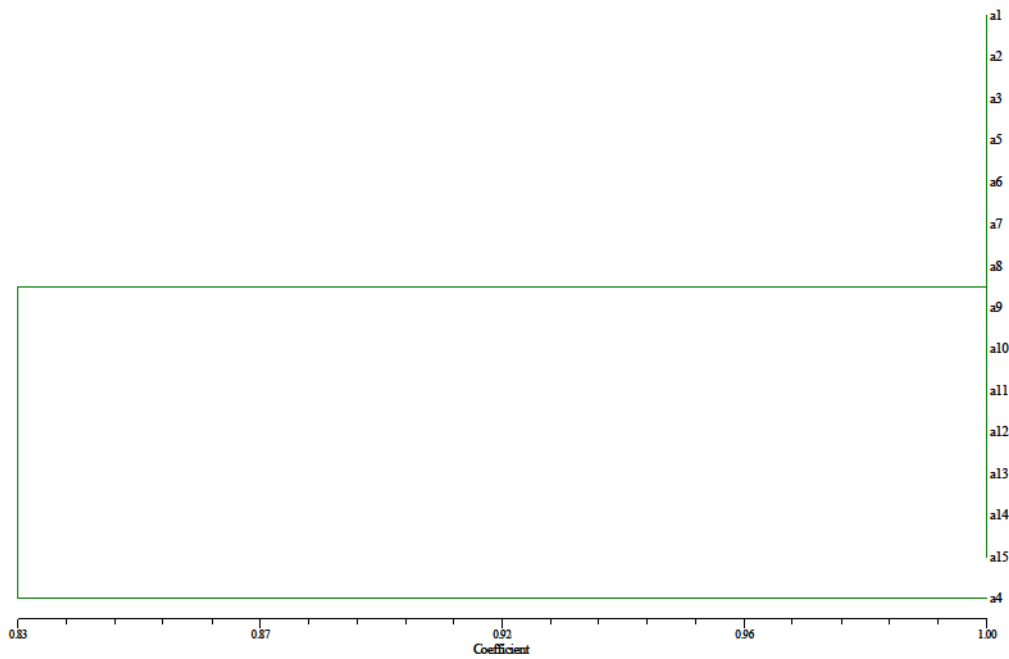


Fig. 12. Albero di similarità genetica ottenuto dalla matrice di similarità calcolata con l'indice di Jaccard. Sulle ascisse il coefficiente di similarità di Jaccard. Sulle ordinate campioni da a1 ad a10 aglio "Schiacciato", da a11 ad a15 "Tondo di Torella".

L'albero di similarità genetica mostrato in figura 12 evidenzia come le piante di aglio campionate siano tutte uguali ad eccezione del campione a4 che si differenzia per due alleli. I dati molecolari sono stati sottoposti al test statistico dell'AMOVA (Par. 2.4.3) per stimare la varianza molecolare inter- e intra-popolazione e per valutare se esistano differenze statisticamente significative tra i due ecotipi.

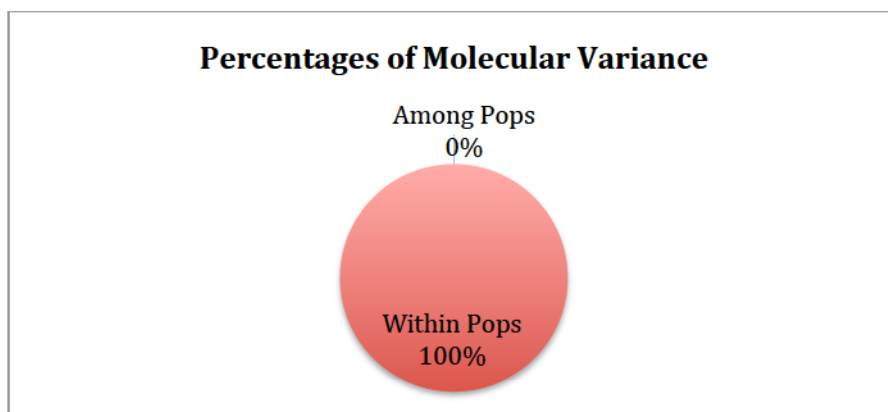


Fig. 13. Grafico a torta delle percentuali relative di diversità genetica tra le popolazioni, tra gli individui e nell'individuo.

Il grafico a torta mostrato in figura 13, ottenuto tramite analisi AMOVA, conferma come tra i due ecotipi con ci sia nessuna differenza. Tuttavia la diversità all'interno della popolazione risulta essere del 100% a causa del diverso peso molecolare di due alleli.

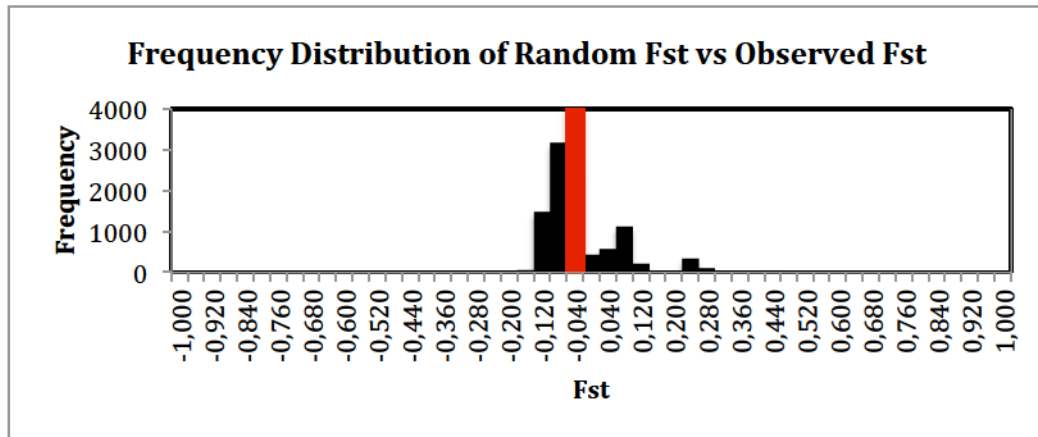


Fig.14. Grafico della distribuzione degli Fst simulati (barre nere) e degli Fst reali.

Il grafico a barre (Fig. 14) mostra il risultato del test di significatività statistica effettuato confrontando i valori reali di Fst a quelli simulati ottenuti permutando i dati originali. Dato che i valori reali (barra rossa) si posizionano nell'intervallo generato dalla distribuzione simulata, l'ipotesi nulla che non ci sia nessuna differenza statistica è accettata con α (livello di significatività) pari a 0,005. I dati sottoposti a elaborazioni statistiche confermano che non è possibile discriminare tra i due ecotipi di aglio mediante gli otto loci SSR scelti per l'analisi.

3.1.2 Cipolla

Gli ecotipi di cipolla analizzati erano quattro, Cipolla "Febbrarese", Cipolla "Marzatica", Cipolla "Ramata di Montoro" e Cipolla "Vatolla", e sono stati indagati sette loci SSR. Questi quattro ecotipi di cipolla sono stati confrontati con due varietà testimone "Test 1" e "Test 2", allo scopo di valutare i profili genetici degli ecotipi campani di interesse rispetto a varietà non inserite nel progetto. L'immagine seguente (Fig. 15) mostra una corsa elettroforetica di DNA estratto dalle foglie di cipolla.

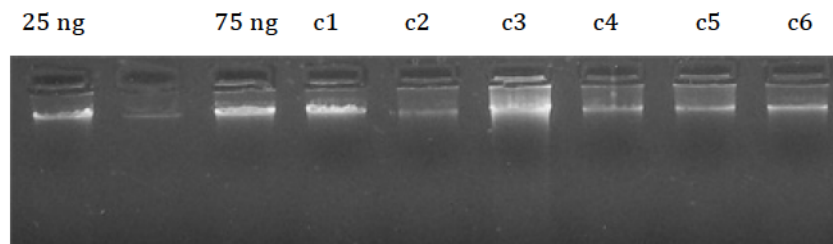


Fig. 15. Immagine della corsa elettroforetica su gel di agarosio del DNA di fago λ di 25 e 75 ng e dei campioni di DNA estratti dalle foglie di cipolla così ordinati: c1 = "Febbrarese", c2 = "Marzatica", c3 = "Ramata di Montoro", c4 = "Vatolla", c5= test1 e c6 = test2.

Come si evidenzia dalla figura 16, il DNA è di buona qualità e di concentrazione sufficiente da permettere le indagini molecolari. Tale osservazione è stata confermata dalle misure effettuate allo spettrofotometro (dati non mostrati). Il DNA genomico estratto è stato analizzato tramite i marcatori SSR amplificando i sette loci (Par. 2.3.1). La figura 16 illustra la corsa elettroforetica su gel di agarosio dei loci SSR amplificati mediante PCR.

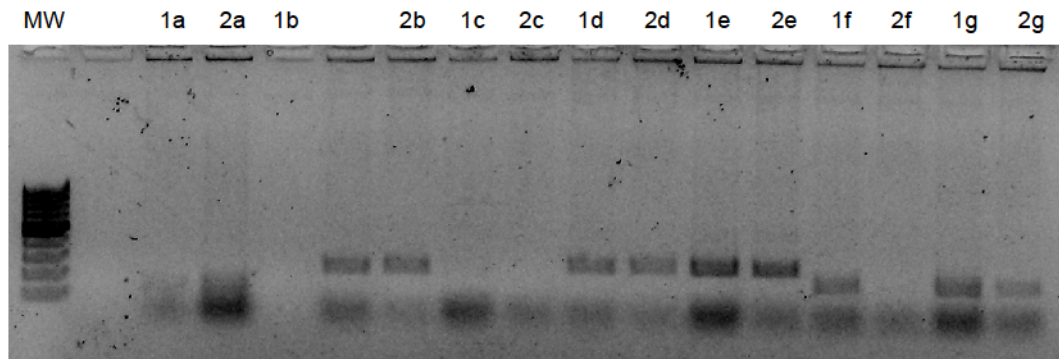


Fig. 16. L'immagine riporta una corsa elettroforetica su gel di agarosio di alcuni prodotti di PCR ottenuti da DNA estratto da foglia di cipolla e amplificato con i *primer* fiancheggianti le regioni SSR. MW = marcatore di peso molecolare, campione 1 = "Febbrarese" campione 2 = "ramata di Montoro"; a = locus 1, b = locus 2, c = locus =3, d = locus =4, e = locus 5, f = locus 6, g = locus 7.

I pesi molecolari dei frammenti amplificati sono stati annotati in un foglio di calcolo in modo da consentire successive elaborazioni biostatistiche. Nella tabella seguente (Tab. 10) sono riportati i valori degli indici descritti in Materiali e Metodi (Par. 2.4.5).

Pop	Locus	N	Na	Ne	I	Ho	He	uHe	F	Gene Diversity	PIC
"Febbrarese"	Locus1	8	6	3,76	1,54	0,75	0,73	0,78	-0,02	0,73	0,7
	Locus2	9	1	1	0	0	0	0	#N/D	0	0
	Locus3	7	2	1,32	0,41	0,29	0,24	0,26	-0,17	0,24	0,21
	Locus4	7	7	5,44	1,81	0,57	0,82	0,88	0,3	0,82	0,79
	Locus5	8	2	1,88	0,66	0,5	0,47	0,5	-0,07	0,47	0,36
	Locus6	8	3	2,51	0,98	0,88	0,6	0,64	-0,45	0,6	0,52
	Locus7	8	4	3,46	1,31	0,88	0,71	0,76	-0,23	0,71	0,66
	Media	8	4	2,77	0,96	0,55	0,51	0,55	#N/D	0,51	0,46
"Marzatica"	Locus1	4	5	3,2	1,39	0,75	0,69	0,79	-0,09	0,69	0,65
	Locus2	6	1	1	0	0	0	0	#N/D	0	0
	Locus3	6	2	1,8	0,64	0,67	0,44	0,48	-0,5	0,44	0,35
	Locus4	6	4	3	1,24	0,33	0,67	0,73	0,5	0,67	0,62
	Locus5	5	2	1,47	0,5	0	0,32	0,36	1	0,32	0,27
	Locus6	5	2	2	0,69	1	0,5	0,56	-1	0,5	0,38
	Locus7	5	2	1,92	0,67	0,8	0,48	0,53	-0,67	0,48	0,36
	Media	5	3	2,06	0,73	0,51	0,44	0,49	#N/D	0,44	0,38
"Ramata di Montoro"	Locus1	11	6	3,51	1,45	0,82	0,71	0,75	-0,14	0,75	0,71
	Locus2	11	1	1	0	0	0	0	#N/D	0,1	0,09
	Locus3	10	3	1,5	0,61	0,4	0,34	0,35	-0,19	0,44	0,41
	Locus4	9	6	4,05	1,57	0,67	0,75	0,8	0,11	0,78	0,75
	Locus5	11	7	2,95	1,46	0,45	0,66	0,69	0,31	0,75	0,72
	Locus6	11	3	2,07	0,82	0,82	0,52	0,54	-0,58	0,57	0,49
	Locus7	10	5	3,39	1,35	0,6	0,71	0,74	0,15	0,72	0,67
	Media	10	4	2,64	1,04	0,54	0,53	0,55	#N/D	0,59	0,55
"Vatolla"	Locus1	9	3	2,22	0,87	0,89	0,55	0,58	-0,62	0,55	0,45
	Locus2	10	1	1	0	0	0	0	#N/D	0	0
	Locus3	9	2	1,91	0,67	0,78	0,48	0,5	-0,64	0,44	0,35
	Locus4	9	5	3,68	1,43	0,67	0,73	0,77	0,08	0,77	0,73
	Locus5	9	5	3,24	1,34	0,56	0,69	0,73	0,2	0,66	0,61
	Locus6	10	2	1,92	0,67	0,8	0,48	0,51	-0,67	0,48	0,36
	Locus7	10	3	2,74	1,05	0,5	0,64	0,67	0,21	0,65	0,58
	Media	9	3	2,39	0,86	0,60	0,51	0,54	#N/D	0,51	0,44

Pop	Locus	N	Na	Ne	I	Ho	He	uHe	F	Gene Diversity	PIC
Test1	Locus1	9	4	3,52	1,31	0,78	0,72	0,76	-0,09	0,68	0,62
	Locus2	9	2	1,12	0,21	0,11	0,1	0,11	-0,06	0,12	0,11
	Locus3	9	2	1,38	0,45	0,33	0,28	0,29	-0,2	0,38	0,3
	Locus4	8	8	6,4	1,96	0,5	0,84	0,9	0,41	0,86	0,84
	Locus5	9	3	1,74	0,73	0,33	0,43	0,45	0,22	0,36	0,33
	Locus6	9	3	2,16	0,85	0,44	0,54	0,57	0,17	0,55	0,46
	Locus7	9	2	1,25	0,35	0	0,2	0,21	1	0,32	0,29
	Media	9	3	2,51	0,84	0,36	0,44	0,47	0,21	0,47	0,42
Test2	Locus1	4	2	1,6	0,56	0,5	0,38	0,43	-0,33	0,54	0,46
	Locus2	5	1	1	0	0	0	0	#N/D	0	0
	Locus3	4	1	1	0	0	0	0	#N/D	0	0
	Locus4	4	3	2,67	1,04	0	0,63	0,71	1	0,79	0,76
	Locus5	5	2	1,22	0,33	0,2	0,18	0,2	-0,11	0,52	0,46
	Locus6	5	2	1,47	0,5	0,4	0,32	0,36	-0,25	0,41	0,32
	Locus7	4	2	1,6	0,56	0,5	0,38	0,43	-0,33	0,28	0,24
	Media	4	2	1,51	0,43	0,23	0,27	0,30	#N/D	0,36	0,32

Tab.10. I valori riportati in tabella corrispondono a: N=Numero di campioni, Na=numero di alleli, Ne= numero di alleli effettivi per locus, I=indice di Shannon, H_o = eterozigosità osservata, H_e = eterozigosità attesa, U_{He} = eterozigosità attesa corretta, F= indice di fissazione, PIC = Polymorphic Index Content.

Dai valori dell'eterozigosità attesa, dell'indice di fissazione e dal *PIC* emerge che gli ecotipi di cipolla studiati non sono geneticamente stabilizzati. Infatti i valori di F positivi prossimi ad 1 indicano fenomeni di incroci o alleli nulli, mentre valori negativi tendenti a -1 indicano un eccesso di eterozigosità dovuta a selezione ottenuta con tecniche basate sull'eterosi. Esaminando i quattro ecotipi di cipolla inseriti nel progetto AGRIGENET appare evidente come, per la maggior parte dei loci esaminati, lo scambio genetico sia compatibile con fenomeni di flusso genico naturale e casuale (per esempio i loci 1, 3 e 5 della cipolla "Febbraiese"). A supporto di questo dato sono stati calcolati i valore del D_{est} che variano dallo 0 del locus 2 allo 0.55 del locus 4; i differenti loci mostrano valori dissimili tra loro nella frequenza allelica tra le sei popolazioni analizzate, evidenziando come tra gli ecotipi indagati non vi sia un'elevata biodiversità genetica, ad eccezione del locus 2, che, come precedentemente detto, mostra un D_{est} pari zero. I risultati ottenuti per la cipolla "Marzatica" e per alcuni loci degli altri ecotipi evidenziano fenomeni di flusso genico tra le popolazioni, compatibili con tecniche di selezione basate sull'eterosi (locus 7 pop 2 e locus 3 pop 4). Le varietà testimone a loro volta non risultano varietà stabilizzate, ma sembrano sottoposte a fenomeni di incroci casuali, come si evince dai valori di F e di PIC; infatti anche la probabilità che due alleli presi a caso nella popolazione (*Gene Diversity*) siano differenti è molto elevata per quasi tutti i loci delle varietà testimone.

Sui dati molecolari ottenuti sono state svolte ulteriori analisi biostatistiche allo scopo di stimare il livello di biodiversità intra- e inter-popolazione mediante la funzione AMOVA (Fig. 17), e per determinare il numero di popolazioni possibili in riferimento ai dati molecolari in nostro possesso, grazie al software *Structure* (Fig. 19).

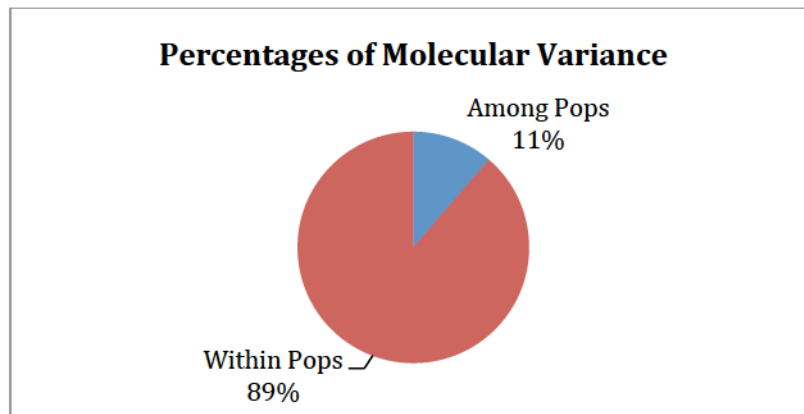


Fig. 17. Il grafico mostra la percentuale di varianza molecolare tra le popolazioni e all'interno della popolazione.

Il grafico in figura 17 riporta la percentuale di biodiversità intra-popolazione rispetto alla percentuale di biodiversità inter-popolazione. È evidente da un'attenta osservazione del grafico che la diversità genetica all'interno delle popolazioni (89%) è molto maggiore rispetto alla diversità tra le popolazioni (11%). Anche questo dato conferma la grande variabilità genetica riscontrata all'interno di ciascun ecotipo di cipolla.

Il grafico riportato nella seguente figura (Fig. 18) è stato generato tramite la funzione AMOVA fornita dal software GenAlex, allo scopo di verificare se, dopo un numero elevato di permutazioni, i profili genetici di cipolla fossero compresi all'interno della distribuzione simulata.

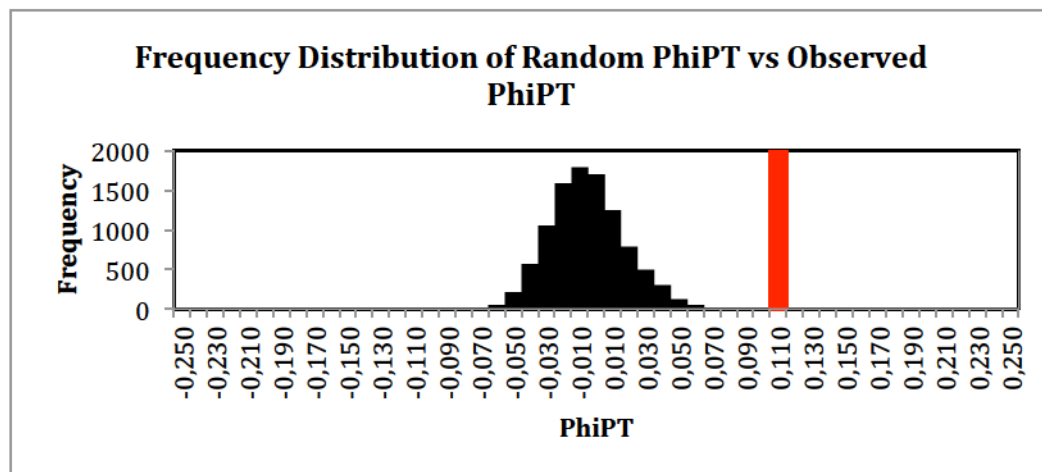


Fig. 18. Grafico della distribuzione degli Fst simulati (barre nere) e degli Fst reali.

Come si evince dal grafico di figura 18 i valori di PhiPT (un analogo di F_{st}) ottenuti dalle analisi molecolari sugli ecotipi di cipolla (banda rossa) si posiziona in maniera netta oltre l'intervallo della distribuzione di valori di PhiPT simulata. Questo dato implica il rifiuto dell'ipotesi nulla, e cioè il rifiuto dell'ipotesi che tra gli ecotipi di cipolla non vi siano

differenze statisticamente significative ($\alpha < 0.005$). Il numero di popolazioni possibili è stato definito mediante il software *Structure* (Par. 2.4.1). Il risultato è stato elaborato con il metodo di (Evanno et al. 2005) che ha suggerito la distribuzione dei campioni in otto possibili popolazioni (dati non mostrati). Come i genomi riconosciuti da queste otto ipotetiche popolazioni si ridistribuiscono in ciascun campione analizzato è riportato in figura 19.

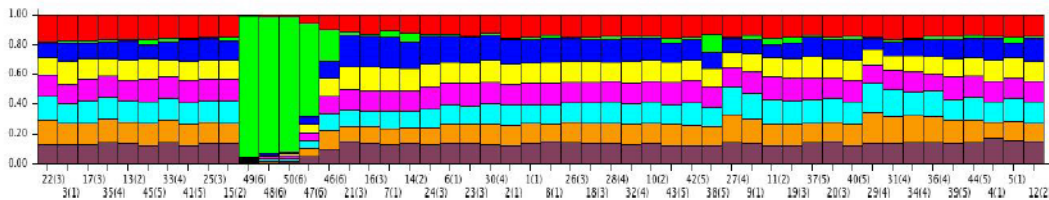


Fig. 19. Il grafico mostra la distribuzione dei diversi genomi per ciascun campione. Campioni da 1 a 9 = "Febbrarese", da 10 a 15 = "Marzatica", da 16 a 26 = "Ramata di Montoro", da 27 a 36 = "Vatolla", da 37 a 45 = "Test1", da 46 a 50 = "Test2".

Come mostrato dal grafico in figura 19 le piante di cipolla indagate risultano costituite dall'apporto di più genomi, a conferma dell'enorme biodiversità presente in queste accessioni. Tale dato è confermato anche dal dendrogramma riportato in figura 20.

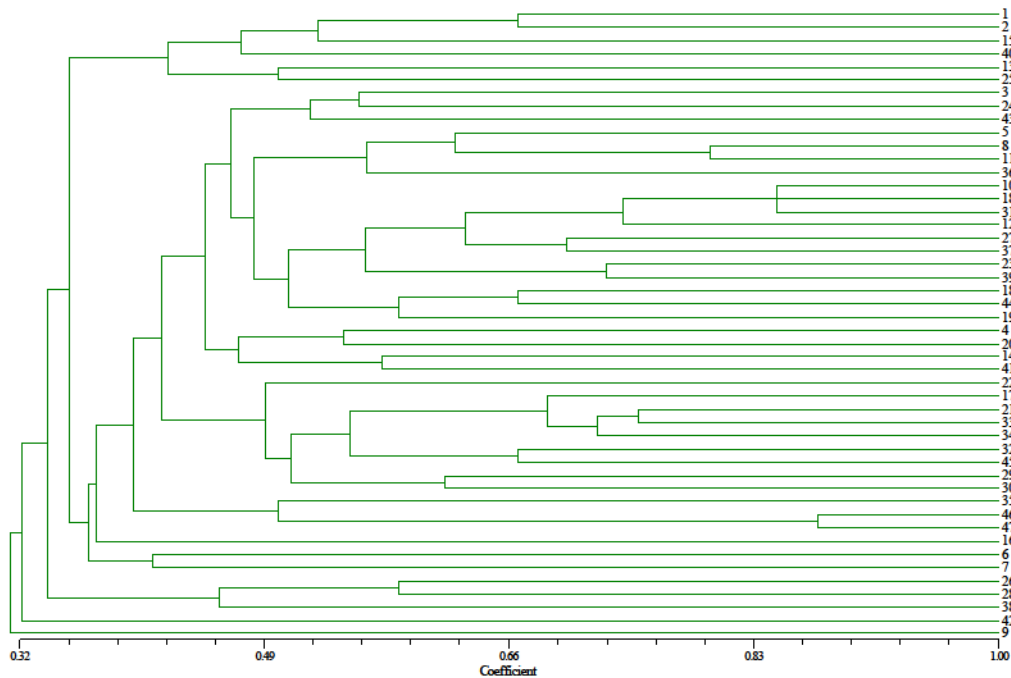


Fig. 20. Albero di similarità genetica ottenuto dalla matrice di similarità calcolata con l'indice di Jaccard. Sulle ascisse il coefficiente di similarità di Jaccard. Campioni da 1 a 9 = "Febbrarese", da 10 a 15 = "Marzatica", da 16 a 26 = "Ramata di Montoro", da 27 a 45 = "Test1", da 46 a 50 = "Test2".

Il dendrogramma mostrato in figura 20 è stato costruito sulla matrice di similarità calcolata secondo l'indice di *Jaccard* (Par. 2.4.1). Anche tale elaborazione conferma l'elevata biodiversità nel campione indagato mediante gli otto loci SSR. Appare chiaro, quindi, che si tratta di accessioni non stabilizzate dotate di un'enorme biodiversità genetica.

3.1.3 Mais

Gli ecotipi di mais analizzati sono stati: Mais "Bianco di Acerra", Mais "Spiga napoletana rossa", Mais "Spiga rossa", Mais "Spiga Bianca", Mais "Spiga napoletana bianca", Mais "Spogna bianca".

I *primer* utilizzati e descritti in Materiale e Metodi (Par. 2.3.1, Tab. 4) amplificano sette loci SSR. Di seguito sono riportati due esempi di corse elettroforetiche su gel di agarosio del DNA purificato dalle foglie di mais (Fig. 21) e dei prodotti delle amplificazioni in PCR (Fig. 22).

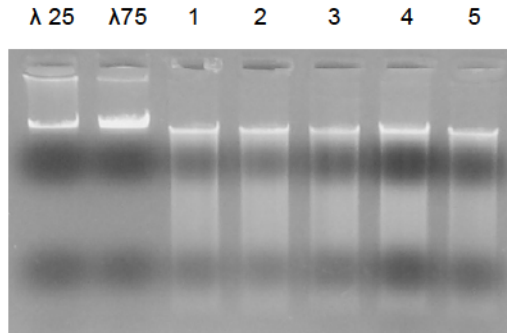


Fig. 21. L'immagine riporta una corsa elettroforetica su gel di agarosio di cinque campioni di DNA di mais confrontati con DNA di fago λ di peso noto: λ25 = 25ng e λ75 = 75ng. Campione 1= Mais "Bianco di Acerra", 2= Mais "Spiga napoletana rossa", 3 = Mais "Spiga rossa", 4 = Mais "Spiga Bianca", 5= Mais "Spiga napoletana bianca".

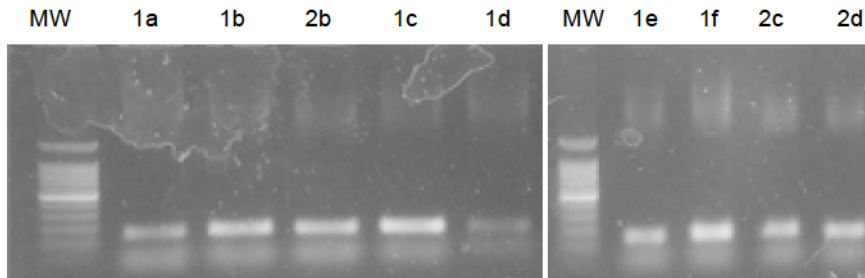


Fig. 22. L'immagine riporta una corsa elettroforetica su gel di agarosio di alcuni prodotti di PCR ottenuti da DNA estratto da foglia di mais e amplificato con i *primer* fiancheggianti le regioni SSR. Campione 1 = Mais "Bianco di Acerra", campione 2 = Mais "Spiga rossa", a = locus 1, b = locus 2, c = locus 3, d = locus 4, e= locus 5, f = locus 7.

Conclusa la fase di ottimizzazione delle condizioni di amplificazione per ciascuna coppia di *primer*, si è proceduto all'analisi di tutti i campioni e alla loro corsa elettroforetica capillare su sequenziatore automatico. Il risultato dell'elettroforesi capillare è stato analizzato tramite il software Gene mapper e i pesi molecolari degli alleli amplificati sono stati annotati in un foglio di calcolo, in modo da stimare gli indici decripti in Materiale e Metodi e di seguito riportati (Tab. 11).

Pop	Locus	N	Na	Ne	I	Ho	He	UHe	F	Gene Diversity	PIC
"Bianco di Acerra"	Locus1	10	4	2,86	1,16	0,60	0,65	0,68	0,08	0,67	0,60
	Locus2	10	3	2,25	0,94	0,30	0,56	0,58	0,46	0,56	0,49
	Locus3	9	2	1,12	0,21	0,11	0,10	0,11	-0,06	0,10	0,10
	Locus4	10	2	1,47	0,50	0,20	0,32	0,34	0,38	0,32	0,27
	Locus5	10	4	2,15	1,03	0,50	0,54	0,56	0,07	0,54	0,50
	Locus6	10	4	2,11	0,98	0,20	0,53	0,55	0,62	0,53	0,48
	Locus7	5	4	3,33	1,28	0,80	0,70	0,78	-0,14	0,70	0,65
	Media	9	3	2,18	0,87	0,39	0,49	0,51	0,20	0,49	0,44
"Spiga Napoletana Rossa"	Locus1	5	3	2,27	0,95	0,40	0,56	0,62	0,29	0,56	0,50
	Locus2	10	2	2,00	0,69	0,00	0,50	0,53	1,00	0,50	0,38
	Locus3	8	1	1,00	0,00	0,00	0,00	0,00	#N/D	0,00	0,00
	Locus4	10	4	2,60	1,09	0,60	0,62	0,65	0,02	0,62	0,54
	Locus5	10	1	1,00	0,00	0,00	0,00	0,00	#N/D	0,00	0,00
	Locus6	8	4	2,72	1,14	1,00	0,63	0,68	-0,58	0,63	0,57
	Locus7	7	2	1,69	0,60	0,57	0,41	0,44	-0,40	0,41	0,32
	Media	8	2	1,90	0,64	0,37	0,39	0,42	#N/D	0,39	0,33
"Spiga rossa"	Locus1	8	1	1,00	0,00	0,00	0,00	0,00	#N/D	0,00	0,00
	Locus2	10	3	2,06	0,82	0,60	0,52	0,54	-0,17	0,52	0,42
	Locus3	8	2	2,00	0,69	0,00	0,50	0,53	1,00	0,50	0,38
	Locus4	10	3	2,99	1,10	0,70	0,67	0,70	-0,05	0,67	0,59
	Locus5	8	2	1,28	0,38	0,00	0,22	0,23	1,00	0,22	0,19
	Locus6	9	3	2,22	0,87	0,56	0,55	0,58	-0,01	0,55	0,45
	Locus7	8	2	1,44	0,48	0,38	0,30	0,33	-0,23	0,30	0,26
	Media	8	2	1,86	0,62	0,32	0,39	0,42	#N/D	0,39	0,33
"Spiga bianca"	Locus1	9	4	2,57	1,12	0,67	0,61	0,65	-0,09	0,61	0,56
	Locus2	10	4	2,74	1,14	0,40	0,64	0,67	0,37	0,64	0,57
	Locus3	7	3	2,33	0,96	0,57	0,57	0,62	0,00	0,57	0,50
	Locus4	10	3	2,90	1,08	0,60	0,66	0,69	0,08	0,66	0,58
	Locus5	8	3	2,67	1,04	0,63	0,63	0,67	0,00	0,63	0,55
	Locus6	10	3	1,80	0,75	0,60	0,45	0,47	-0,35	0,45	0,38
	Locus7	10	2	1,10	0,20	0,10	0,10	0,10	-0,05	0,10	0,09
	Media	9	3	2,30	0,90	0,51	0,52	0,55	-0,01	0,52	0,46

Pop	Locus	N	Na	Ne	I	Ho	He	UHe	F	Gene Diversity	PIC
"Spiga Napoletana bianca"	Locus1	9	4	3,06	1,21	0,89	0,67	0,71	-0,32	0,67	0,61
	Locus2	9	4	2,79	1,16	0,89	0,64	0,68	-0,38	0,64	0,58
	Locus3	9	2	1,53	0,53	0,44	0,35	0,37	-0,29	0,35	0,29
	Locus4	9	2	1,25	0,35	0,22	0,20	0,21	-0,13	0,20	0,18
	Locus5	9	3	2,05	0,83	0,11	0,51	0,54	0,78	0,51	0,43
	Locus6	9	3	2,16	0,90	0,56	0,54	0,57	-0,03	0,54	0,47
	Locus7	8	3	2,84	1,07	0,50	0,65	0,69	0,23	0,67	0,61
	Media	9	3	2,24	0,86	0,52	0,51	0,54	-0,02	0,51	0,45
"Spogna Bianca"	Locus1	5	4	2,94	1,22	0,80	0,66	0,73	-0,21	0,66	0,61
	Locus2	8	4	3,12	1,25	0,88	0,68	0,73	-0,29	0,68	0,62
	Locus3	8	2	1,13	0,23	0,13	0,12	0,13	-0,07	0,12	0,11
	Locus4	8	2	1,97	0,69	0,63	0,49	0,53	-0,27	0,49	0,37
	Locus5	7	2	1,15	0,26	0,14	0,13	0,14	-0,08	0,26	0,24
	Locus6	7	2	1,15	0,26	0,14	0,13	0,14	-0,08	0,13	0,12
	Locus7	8	3	2,72	1,04	0,88	0,63	0,68	-0,38	0,63	0,56
	Media	7	3	2,03	0,71	0,51	0,41	0,44	-0,20	0,42	0,38

Tab. 11. I valori riportati in tabella corrispondono a: N=Numero di campioni, Na=numero di alleli, Ne= numero di alleli effettivi per locus, I=indice di Shannon, H_o = eterozigosità osservata, H_e = eterozigosità attesa, UH_e = eterozigosità attesa corretta, F= indice di fissazione, PIC = Polymorphic Index Content.

Come si evince dalla tabella 11, i valori di eterozigosità osservata sono confrontabili con quella di eterozigosità attesa per la maggior parte dei locus genici testati; inoltre l'indice di fissazione (F), nella maggior parte dei casi, è prossimo allo 0 suggerendo una riproduzione casuale. I valori positivi, invece, indicano sostanziali consanguineità o alleli nulli non rilevati. I valori medi di *Gene diversity* sono compresi tra 0.39 e 0.52; tuttavia la situazione non è descritta in maniera esauriente da tali valori in quanto questo indice è molto variabile in relazione al locus e al campione. E' possibile notare, infatti, come per tutti i loci i valori sono compresi tra un minimo di 0.10 e un massimo di 0.70. I valori medi del PIC confermano l'elevata biodiversità degli ecotipi di mais indagati (min 0.33, max 0.45); stesso risultato è evidenziato dall'indice di Shannon (I) che presenta dei valori medi compresi tra 0,64 e 0,90, raggiungendo dei massimi di 1,28 e 1,21, rispettivamente per i loci 7 e 1, per le popolazioni 1 e 5. Mediante il software Genalex è stato eseguito anche il test statistico AMOVA. La stima della varianza molecolare inter- e intra-popolazione è riportata nella seguente figura (Fig. 23).

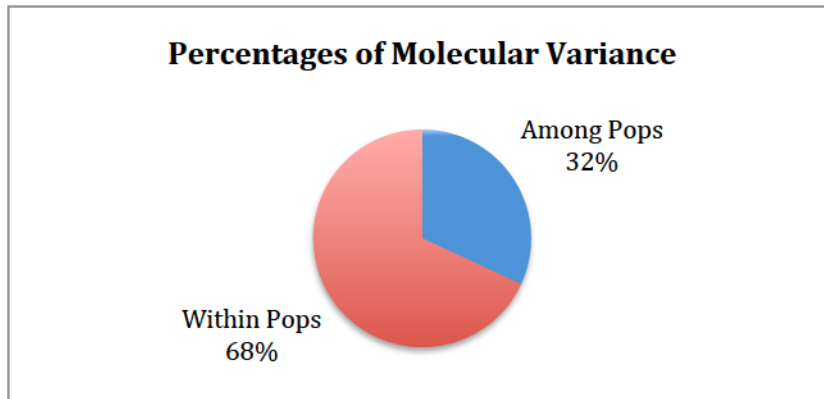


Fig. 23. Il grafico mostra la percentuale di varianza molecolare tra le popolazioni e nella popolazione.

Il grafico di figura 24 mostra la percentuale di varianza genetica-molecolare inter-popolazione rispetto a quella intra-popolazione. E' evidente come la varianza molecolare intra-popolazione sia maggiore (68%) di quella inter-popolazione (32%). Dal test di significatività statistica eseguito contestualmente all'AMOVA è stata rigettata l'ipotesi nulla (ovvero, nessuna differenza tra le popolazioni) con un livello di significatività pari a 0,001; infatti analizzando il grafico della distribuzione degli F_{st} simulata, rispetto ai valori reali, è palese come questi ultimi si posizionino fuori dalla distribuzione simulata, risultando statisticamente significativi (Fig. 24).

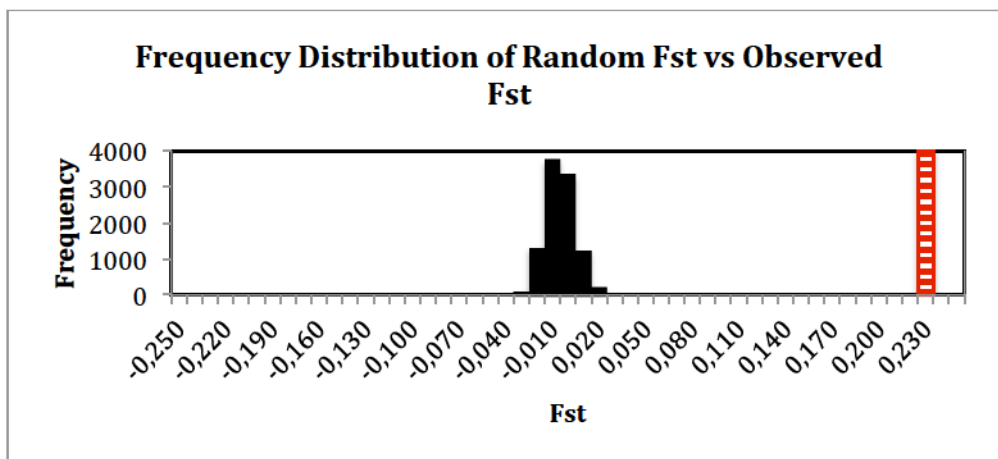


Fig. 24. Grafico della distribuzione degli F_{st} simulati (barre nere) e degli F_{st} reali.

I dati molecolari degli ecotipi di mais sono stati elaborati tramite il software *Structure* allo scopo di determinare il numero di popolazioni mediante il metodo di Evanno et al. (2005). Il numero di popolazioni stimato (K) è pari a due; la simulazione ricalcolata sulla base di $K=2$ è riportata in figura 25.

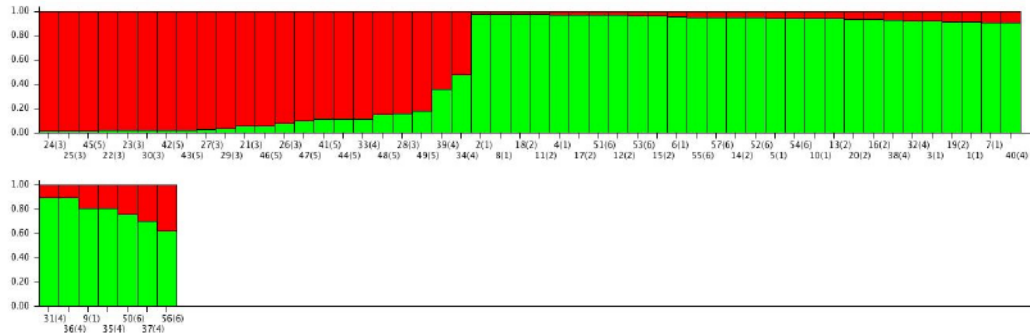


Fig. 25. Il grafico mostra la distribuzione dei diversi genomi all'interno di ogni campione per K = 2. Campioni da 1 a 9 = Mais "Bianco di Acerra", da 11 a 20 = Mais "Spiga napoletana rossa", da 21 a 30 = Mais "Spiga rossa", da 31 a 40 = Mais "Spiga Bianca", da 41 a 49 = Mais "Spiga napoletana bianca", da 50 a 57 = Mais "Spogna bianca".

Dal grafico a barre in figura 25 si può osservare che esistono due gruppi che comprendono i 6 ecotipi di mais indagati; ciò implica che i campioni analizzati potrebbero appartenere a due macro gruppi. In figura 26 è riportato un *triangle plot* che descrive in maniera efficace la distribuzione dei sei ecotipi, rappresentati dai diversi colori, tra le due popolazioni stimate (clusters).

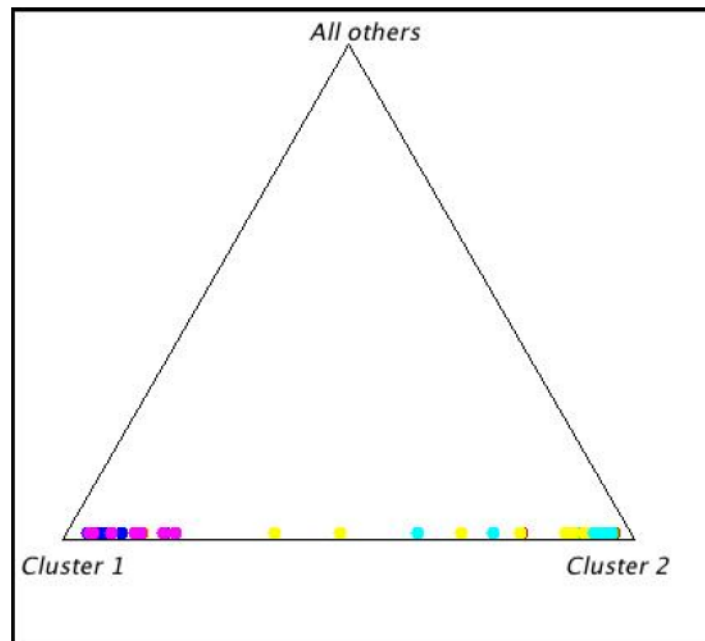


Fig. 26. Il *triangle plot* mostra la distribuzione dei sei ecotipi di mais. Ad ogni colore corrisponde un singolo ecotipo di mais: rosso "Bianco di Acerra", verde "Spiga napoletana rossa", blue "Spiga rossa", giallo "Spiga Bianca", viola "Spiga napoletana bianca", azzurro "Spogna bianca".

Gli ecotipi "Spiga rossa" e "Spiga napoletana bianca" formano il *cluster 1* e sono ben separati dagli altri che costituiscono il *cluster 2*, ovvero il mais "bianco di Acerra", la

“Spogna bianca”, la “Spiga napoletana rossa”. I campioni di “Spiga bianca” (in giallo) sono invece distribuiti tra i due *cluster*.

Anche il dendrogramma costruito sulla matrice di similarità ha evidenziato l'eterogeneità genetica delle piante di mais (Fig. 27).

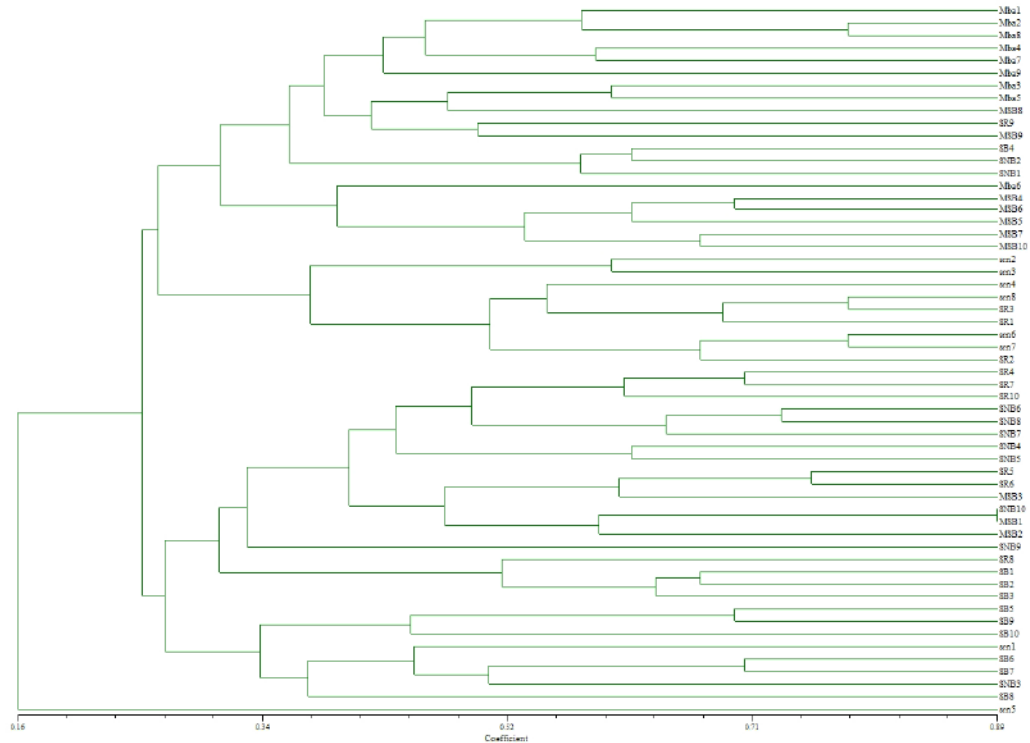


Fig. 27. Albero di similarità genetica ottenuto dalla matrice di similarità calcolata con l'indice di Jaccard. Sulle ascisse il coefficiente di similarità di Jaccard. Campioni da 1 a 9 = Mais “Bianco di Acerra”, da 11 a 20 = Mais “Spiga napoletana rossa”, da 21 a 30 = Mais “Spiga rossa”, da 31 a 40 = Mais “Spiga Bianca”, da 41 a 49 = Mais “Spiga napoletana bianca”, da 50 a 57 = Mais “Spogna bianca”.

Dal dendrogramma riportato in figura 27 emerge la presenza di due macro raggruppamenti a conferma del risultato di *Structure*. Inoltre si può osservare come tutte le piante mostrino tra di loro dei coefficienti di similarità (*Jaccard*) piuttosto bassi (max = 0,75). Fanno eccezione i campioni SNB10 e MB1, che sono gli unici a presentare un coefficiente di similarità pari a 0,89. Il dendrogramma, infatti, evidenzia come i sei ecotipi di mais analizzati non siano linee stabilizzate e diversificate; al contrario esse presentano molti alleli differenti, all'interno degli stessi ecotipi, tanto che la biodiversità intra-ecotipo è addirittura maggiore di quello inter-ecotipo. Questo dato è in accordo con i precedenti, e conferma la difficoltà di distinguere chiaramente i sei ecotipi di mais mediante i sette loci SSR utilizzati, tuttavia grazie a questi marcatori siamo stati in grado di stimare l'elevata biodiversità genetica esistente tra gli ecotipi.

3.1.4 Scarola

L'ecotipo di scarola studiato è chiamato “*Riccia schiana*”. I loci SSR analizzati sono otto, mentre i campioni di foglia da cui è stato estratto il DNA per le indagini molecolari sono 10. La figura 28 riporta la corsa elettroforetica del DNA estratto seguendo le metodiche riportate in Materiali e Metodi (Par. 2.2).

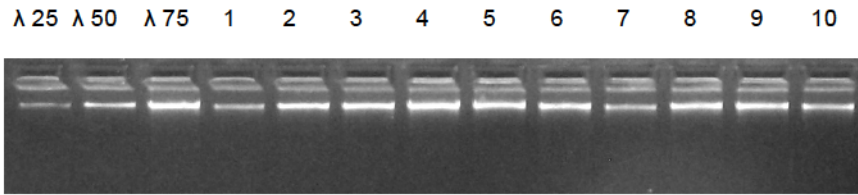


Fig 28. L'immagine riporta la corsa elettroforetica su gel di agarosio di 10 campioni di DNA di Scarola confrontati con DNA di fago λ di peso noto: λ25 = 25ng, λ50= 50ng e λ75 = 75ng. Campioni da 1 a 10 = scarola "Riccia Schiana".

Le indagini molecolari previste sono state condotte seguendo i protocolli e le condizioni di amplificazione riportate in Materiali e Metodi (Par. 2.3.1). La figura 29 mostra una corsa elettroforetica su gel di agarosio dei loci SSR ottenuti dall' amplificazione in PCR.

MW 1a 1b 1c 1d 1e 1f 1g 1h

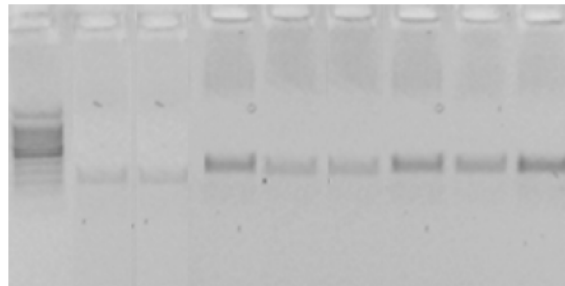


Fig. 29. L'immagine riporta la corsa elettroforetica su gel di agarosio dei prodotti di PCR ottenuti da DNA estratto da foglia di scarola e amplificato con i primer fiancheggiante le regioni SSR. Lettera a = locus 1, b = locus 2, c = locus 3, d = locus 4, e = locus 5, f = locus 6, g = locus 7, h = locus 8.

I pesi molecolari relativi agli alleli amplificati sono stati inseriti in un foglio di calcolo ed elaborati come descritto in Materiali e Metodi (Par. 2.4.5). Nella seguente tabella (Tab. 12) sono riportati i valori degli indici calcolati.

Pop	Locus	N	Na	Ne	I	Ho	He	uHe	F	Gene Diversity	PIC
"Riccia Schiana"	Locus1	10	2	1,47	0,50	0,00	0,32	0,34	1,00	0,32	0,27
	Locus2	8	1	1,00	0,00	0,00	0,00	0,00	#N/D	0,00	0,00
	Locus3	9	3	1,74	0,73	0,11	0,43	0,45	0,74	0,43	0,37
	Locus4	10	2	1,11	0,20	0,10	0,10	0,10	-0,05	0,10	0,09
	Locus5	9	3	1,59	0,68	0,22	0,37	0,39	0,40	0,37	0,34
	Locus6	10	1	1,00	0,00	0,00	0,00	0,00	#N/D	0,00	0,00
	Locus7	10	4	2,33	1,00	0,80	0,57	0,60	-0,40	0,57	0,49
	Locus8	10	2	1,98	0,69	0,90	0,50	0,52	-0,82	0,50	0,37
	Media	9,5	2,25	1,53,	0,47	0,27	0,28	0,30	#N/D	0,28	0,24

Tab.12. In tabella sono riportati il: Numero di campioni (N), =numero di alleli (Na), Ne= numero di alleli effettivi per locus, I=indice di Shannon, Ho = eterozigosità osservata, He= eterozigosità attesa, UHe= eterozigosità attesa corretta, F= indice di fissazione, PIC = Polymorphic Index Content.

Dai valori di eteroziguità osservata e dal valore dell'indice di fissazione (F), si può notare che per alcuni alleli esiste una certa stabilità genetica dovuta alla ripetuta omoziguità di quei campioni, come nel caso dei loci 1, 2 e 6. In quest'ultimo caso il valore di F risulta matematicamente non calcolabile in quanto sia l'eteroziguità osservata che quella attesa risultano pari a zero. Per i loci 1 e 3, il valore di F è molto elevato a causa dell'eteroziguità osservata prossima allo zero; valori di F pari a 1, infatti, indicano fenomeni di *inbreeding*. Valori di F vicini a 0 (come evidenziato per il locus 4) di solito si osservano in situazioni di riproduzione casuale; al contrario valori molto negativi (come osservato per i loci 7 e 8) sono caratteristici di fenomeni di eterozigosi eccessiva e probabilmente sono dovuti a fenomeni di flusso genico, volti a favorire l'eterosi. La probabilità di campionare due alleli differenti nella popolazione è bassa, infatti i valori maggiori si riscontrano per i loci 7 e 8 (rispettivamente 0,57 e 0,50). Esaminando l'indice di Shannon (I) è possibile notare che il valore medio (0,47) suggerisce una ridotta biodiversità.

I dati molecolari ottenuti sono stati elaborati con il software NTSys PC per calcolare una matrice di similarità basata sul coefficiente di Jaccard. Il risultato è mostrato in figura 30.

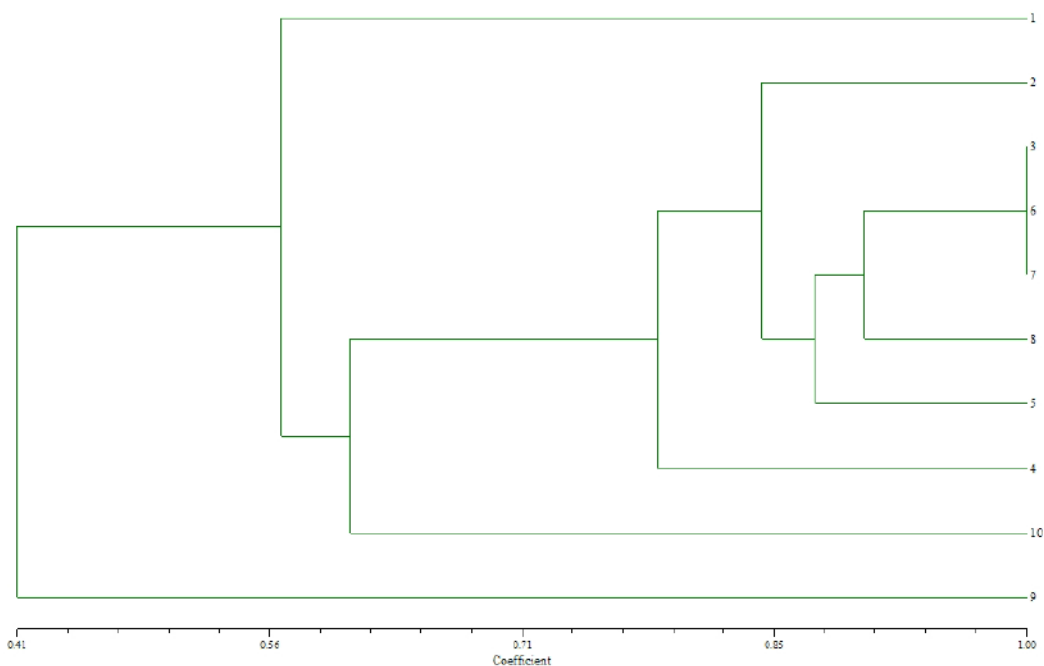


Fig. 30. Albero di similarità genetica ottenuto dalla matrice di similarità calcolata impiegando l'indice di Jaccard. Sulle ascisse il coefficiente di similarità di Jaccard. Campioni da 1 a 10 = scarola "Riccia Schiana".

Dal dendrogramma in figura 30 si può osservare come, ad eccezione dei campioni 1, 9 e 10, le piante di scarola indagate costituiscono un gruppo abbastanza omogeneo, con valori di similarità compresi tra 0,85 e l'unità, come nel caso dei campioni 3, 6 e 7.

Per sette campioni su 10 la percentuale di similarità è elevata (circa 85 %), tuttavia gli altri tre campioni dello stesso ecotipo sembrano essere maggiormente biodiversi.

3.1.5 Zucchino

Gli ecotipi di zucchino sottoposti ad analisi molecolare erano due: "Cilentano" (Pop1) e "San Pasquale" (Pop2). Dai campioni di foglia è stato estratto il DNA (Fig. 31) secondo le metodiche riportate in Materiali e Metodi (Par. 2.3.1).

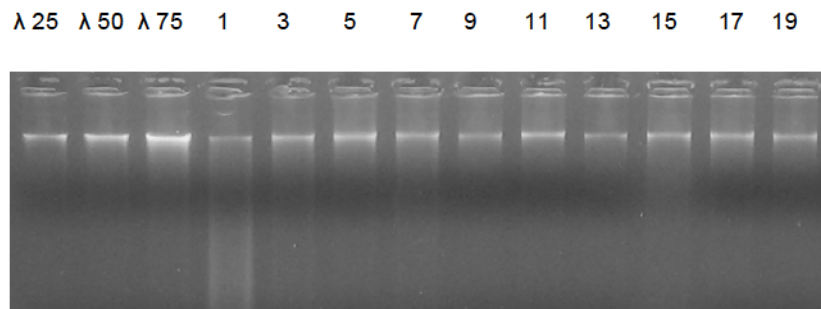


Fig. 31. L'immagine riporta una fotografia della corsa elettroforetica su gel di agarosio di 10 campioni di DNA di Zucchini confrontati con DNA di fago λ a concentrazione nota: λ 25 = 25ng, λ 50= 50ng e λ 75 = 75ng. Campioni da 1 a 9 = zucchini "Cilentano", da 11 a 19 = zucchini "San Pasquale".

Il DNA genomico è stato quindi sottoposto ad amplificazione dei loci SSR ottimizzando i programmi di amplificazione PCR (Fig. 32).

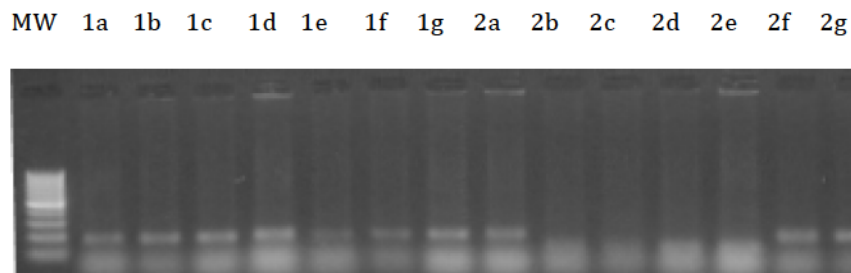


Fig. 32. L'immagine riporta la corsa elettroforetica su gel di agarosio dei prodotti di PCR ottenuti da DNA estratto da foglia di zucchini e amplificato con i *primer* fiancheggianti le regioni SSR. Campione 1 = zucchini "Cilentano", campione = 2 = zucchini "San Pasquale". Lettera a = locus 1, b = locus 2, c = locus 3, d = locus 4, e = locus 5, f = locus 6, g = locus 7.

I prodotti dell'amplificazione sono stati separati mediante elettroforesi capillare su sequenziatore automatico e i pesi molecolari degli alleli amplificati sono stati riportati in un foglio di calcolo ed elaborati mediante software *ad hoc*, come riportato nel capitolo Materiali e Metodi. Nella seguente tabella sono mostrati gli indici di biodiversità calcolati.

Pop	Locus	N	Na	Ne	I	Ho	He	UHe	F	Gene Diversity	PIC
"Cilentano"	Locus1	9	3	1,26	0,43	0,11	0,2	0,22	0,45		
										0,20	0,19
	Locus2	10	2	1,22	0,33	0,2	0,18	0,19	-0,11	0,18	0,16
	Locus3	10	1	1	0	0	0	0	#N/D	0,00	0,00
	Locus4	10	2	1,22	0,33	0	0,18	0,19	1	0,18	0,16
	Locus5	10	4	2,9	1,17	0,9	0,66	0,69	-0,37	0,66	0,59
	Locus6	6	1	1	0	0	0	0	#N/D	0,00	0,00
	Locus7	8	5	4,41	1,54	0,75	0,77	0,83	0,03	0,77	0,74
Media	9	3	1,86	0,54	0,28	0,28	0,30	#N/D	0,28	0,26	
"San Pasquale"	Locus1	10	4	3,45	1,30	1,00	0,71	0,75	-0,41	0,71	0,66
	Locus2	10	1	1,00	0,00	0,00	0,00	0,00	#N/D	0,00	0,00
	Locus3	10	1	1,00	0,00	0,00	0,00	0,00	#N/D	0,00	0,00
	Locus4	10	2	1,47	0,50	0,20	0,32	0,34	0,38	0,32	0,27
	Locus5	10	4	2,82	1,16	0,90	0,65	0,68	-0,40	0,65	0,59
	Locus6	9	1	1,00	0,00	0,00	0,00	0,00	#N/D	0,00	0,00
	Locus7	10	4	2,44	1,02	1,00	0,59	0,62	-0,69	0,59	0,50
	Media	10	2	1,88	0,57	0,44	0,32	0,34	#N/D	0,32	0,29

Tab.13. I valori riportati in tabella corrispondono a: N=Numero di campioni, Na=numero di alleli, Ne= numero di alleli effettivi per locus, I=indice di Shannon, Ho = eterozigosità osservata, He= eterozigosità attesa, UHe= eterozigosità attesa corretta, F= indice di fissazione, PIC = Polymorphic Index Content.

La tabella 13 riporta gli indici di biodiversità calcolati grazie ai software GenAlex e PowerMarker. Il valore medio dell'indice di Shannon, sia per l'ecotipo "Cilentano" che per il "San Pasquale", indica una biodiversità non elevata (rispettivamente 0,54 e 0,57); tuttavia in tabella si alternano campioni con maggiore eterozigosità osservata (Pop 1 loci 5 e 7, Pop 2 loci 1, 5 e 7), ad altri in cui l'omozigosi è totale (loci 3 e 5 di entrambe le popolazioni). Ovviamente anche il valore di F (calcolato sui valori dell'eterozigosità) varia da valori fortemente negativi (Pop 2, loci 6 e 7), che suggeriscono un eccesso di eterozigosità probabilmente causati da fenomeni di ibridazione e selezione artificiale tesi a implementare l'eterosi, a valori molto positivi (Pop 1 loci 1 e 4) causati da fenomeni di inbreeding. Tuttavia esistono altri loci il cui valore di F è compatibile con fenomeni di riproduzione avvenuti in maniera casuale (Pop 1, loci 2 e 7).

I dati molecolari sono stati elaborati, allo scopo di stimare il numero di popolazioni, mediante il software *Structure*; i risultati ottenuti sono stati elaborati secondo il metodo di Evanno et al. (2005). Questo metodo ha permesso di stimare il numero di popolazioni (K) pari a due. Ottenuta questa informazione è stata avviata la simulazione come descritto in Materiale e Metodi e il risultato è illustrato in figura 33 e 34.

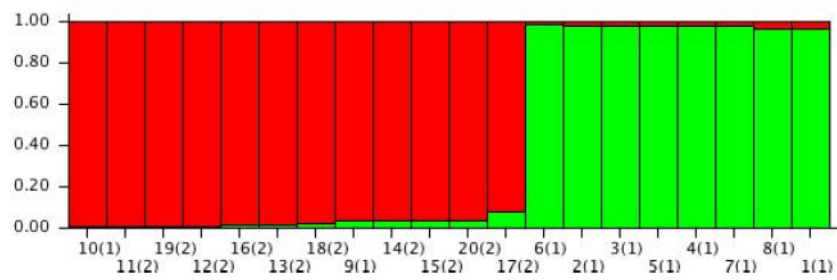


Fig. 33. Il grafico illustra la distribuzione dei diversi genomi all'interno di ogni campione, per il numero di popolazione per $K = 2$. Campioni da 1 a 10 = zucchini "Cilentano", campioni da 11 a 20 = zucchini "San Pasquale".

Il grafico a barre mostrato in figura 33 descrive chiaramente che i campioni analizzati si distribuiscono in due popolazioni, nettamente separate dal punto di vista genetico. I campioni dall'1 al 10 rappresentano l'ecotipo "Cilentano", mentre dall'11 al 20 quello "San Pasquale". In realtà dal grafico sopra riportato i campioni 9 e 10 sono riconducibili al *cluster* rosso, che rappresenta l'ecotipo di zucchini "San Pasquale". Nella figura seguente (Fig. 34) è rappresentato un *triangle plot* che descrive la distribuzione dei campioni studiati tra i due *cluster*.

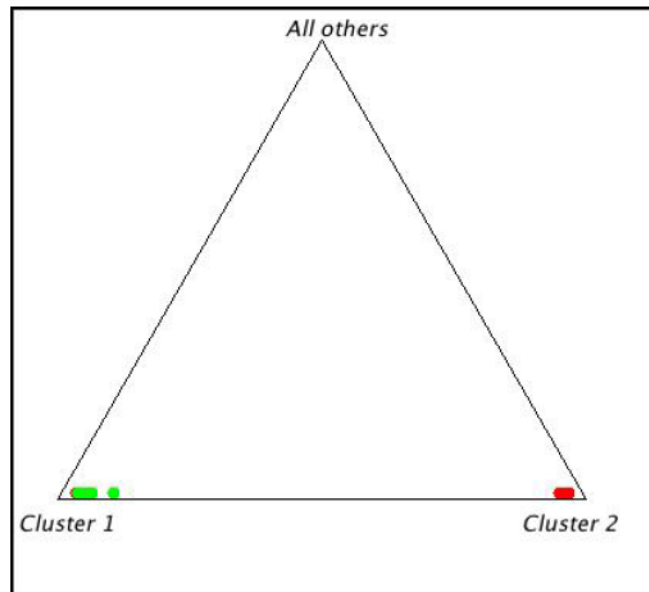


Fig. 34. Il grafico (*triangle plot*) illustra la distribuzione dei due ecotipi di zuccchino. Il colore rosso corrisponde all'ecotipo "San Pasquale", mentre il colore verde allo zuccchino "Cilentano".

Dal "*triangle plot*", così come dal precedente grafico, emerge una netta separazione tra i due ecotipi. Per stimare la percentuale di similarità dei campioni analizzati è stata calcolata una matrice di similarità, utilizzando l'indice di *Jaccard*. Il dendrogramma di figura 35 illustra graficamente i valori di similarità calcolati mediante il programma NTSys-PC.

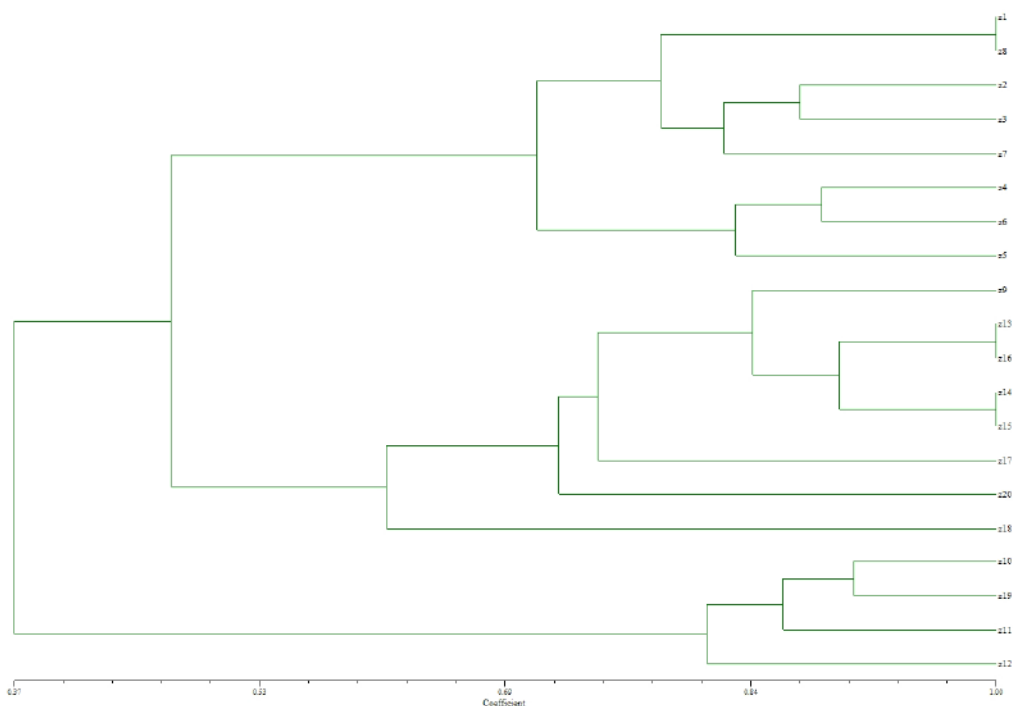


Fig. 35. Albero di similarità genetica ottenuto dalla matrice di similarità calcolata impiegando l'indice di Jaccard. Sulle ascisse è riportato il coefficiente di similarità. Campioni da 1 a 10 = zucchini "Cilentano", campioni da 11 a 20 = zucchini "San Pasquale".

L'albero di similarità mostra come i campioni da 1 a 8 si aggregano in un unico sottogruppo, con similarità prossima a 0,70. Tutti gli altri, ad eccezione dei campioni 10, 11, 12 e 19 formano un secondo sottogruppo con un valore di similarità pari a 0,55. Dal dendrogramma emerge la presenza di un terzo raggruppamento distinto dagli altri due, che ha valori di similarità rispetto agli altri due sottogruppi di circa 0,40.

3.2 Popolazioni naturali di *P. alba*

Una parte del progetto di ricerca di dottorato discusso in questa tesi ha avuto come principale obiettivo quello di implementare un lavoro pubblicato precedentemente dal gruppo di ricerca del Prof. Castiglione (Brundu et al. 2008). Tale studio era volto alla caratterizzazione della biodiversità genetica di popolazioni naturali di *P. alba* della Sardegna, della Sicilia, del Parco Regionale Lombardo del Ticino, della Macedonia e di altre regioni dell'Europa meridionale e, quindi confrontati con campioni di riferimento di *Populus tremula* L. provenienti dalla Scozia. Il lavoro dimostrava che le popolazioni sarde di pioppo bianco sono isolate geneticamente dalle rimanenti e manifestano una spiccata riproduzione di tipo vegetativo. Alla luce dei risultati pregressi, i dati molecolari (SSR) sono stati rielaborati, nell'ambito del progetto di dottorato, attraverso il software *Structure* allo scopo di analizzare la struttura dell'intero gruppo di pioppi campionati, senza utilizzare alcuna conoscenza a priori riguardo il numero di cluster (popolazioni) presenti nel data set e, in seguito, di stimare il numero di popolazioni ottimali (K) in cui sono suddivisi i campioni raccolti e ampiamente distribuiti nel bacino Mediterraneo. Il risultato delle elaborazioni condotte su 116 campioni ha stimato un K pari a 2. Di seguito è riportato il diagramma a barre che ricostruisce la struttura genetica delle popolazioni indagate (Fig. 36):

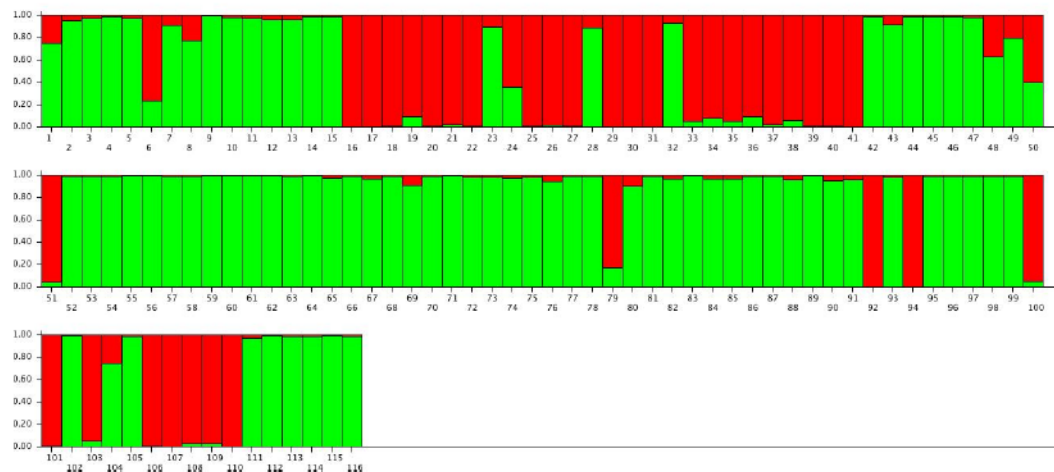


Fig. 36. Il grafico mostra la distribuzione dei diversi genomi all'interno di ogni campione per $K = 2$. Campioni da 1 a 15 = *P. tremula* Scozia (riferimento); campioni da 16 a 42 *P. alba* regine Sardegna; campioni da 43 a 90 *P. alba* parco del Ticino; campioni dal 91 al 118 bacino del mediterraneo (vedi tabella 14).

Il grafico a barre presentato in figura 36 mostra la struttura delle 2 popolazioni e la distribuzione dei genomi di pioppo bianco analizzati. In generale appare evidente che quasi tutti i pioppi campionati possiedono un genoma omogeneo, con elevata proporzione di appartenenza, e poche sono le eccezioni caratterizzate da "mescolamento genetico" (per es. campioni 7, 8 e 23). In particolare i campioni di *P. alba* provenienti dalla Sardegna (da 16 a 41), nonostante siano stati campionati in un'area molto vasta (decine di migliaia di chilometri quadrati), presentano dei genomi ben diversificati (bande rosse) dal resto dei campioni considerati e dai pioppi tremuli di riferimento (bande verdi). Inoltre, l'elaborazione statistica ha consentito di evidenziare come alcuni pioppi bianchi campionati nel bacino del mediterraneo (Tab. 14) possiedono dei genomi molto simili a quelli della popolazione sarda (per es. 92, 94, 101).

Id	Specie	Provenienza
89	<i>P.alba</i>	Camargue
90	<i>P.alba</i>	Rodi
91	<i>P.alba</i>	Rodi
92	<i>P.alba</i>	Rodi
93	<i>P.alba</i>	Rodi
94	<i>P.alba</i>	Lesvos
95	<i>P.alba</i>	Corsica
96	<i>P.alba</i>	Corsica
97	<i>P.alba</i>	Corsica
98	<i>P.alba</i>	Corsica
99	<i>P.alba</i>	Corsica
100	<i>P.alba</i>	Sicilia
101	<i>P.alba</i>	Sicilia
102	<i>P.alba</i>	Sicilia
103	<i>P.alba</i>	Sicilia
104	<i>P.alba</i>	Sicilia
105	<i>P.alba</i>	Macedoinia
106	<i>P.alba</i>	Macedoinia
107	<i>P.alba</i>	Macedoinia
108	<i>P.alba</i>	Macedoinia
109	<i>P.alba</i>	Macedoinia
110	<i>P.alba</i>	Macedoinia
111	<i>P.alba</i>	Macedoinia
112	<i>P.alba</i>	Camargue
113	<i>P.alba</i>	Camargue
114	<i>P.alba</i>	Camargue
115	<i>P.alba</i>	Camargue
116	<i>P.alba</i>	Camargue

Tab. 14. Elenco dei pioppi bianchi campionati nel bacino del Mediterraneo dal gruppo di ricerca del Prof. Castiglione.

La particolare struttura genetica delle popolazioni di *P. alba* della regione Sardegna, indagata tramite SSR, ha ispirato il progetto di dottorato illustrato in questa tesi. La natura clonale di queste popolazioni è idonea a valutare se e come differenti condizioni ambientali possano alterare lo stato di metilazione del DNA.

A questo scopo è stato eseguito un nuovo campionamento dei pioppi sardi che ha visto, non soltanto il recupero di materiale fogliare dalle popolazioni indagate nel precedente lavoro, ma anche da piante raccolte in nuovi siti (Par.2.1.2). I DNA estratti da queste foglie, come descritto in Materiale e Metodi (Par. 2.3.2), sono stati processati tramite la metodica MSAP (Par. 2.3.2). A differenza del precedente lavoro, questa tecnica molecolare

permette di stimare la biodiversità genetica, indagando tutto il genoma, ma anche lo stato di metilazione del DNA.

3.2.1 *Populus alba* – Methylation Sensitive Amplified Polymorphism

Le indagini effettuate sui campioni delle popolazioni naturali di pioppo bianco della Sardegna hanno visto l'impiego della metodica MSAP. Questa tecnica biomolecolare ha permesso di ottenere informazioni sulla biodiversità genetica delle popolazioni indagate, e di valutare lo stato di metilazione del DNA estratto dalle foglie di ciascuna pianta. Le foglie di *P. alba* sono state raccolte in modo da garantire un campionamento rappresentativo delle piante indagate così come riportato in Materiali e Metodi (Par. 2.1.2). Dalle foglie raccolte è stato estratto il DNA e la sua concentrazione e qualità è stata stimata sia mediante corsa elettroforetica su gel di agarosio (Fig. 37) che mediante spettrofotometria UV/visibile.

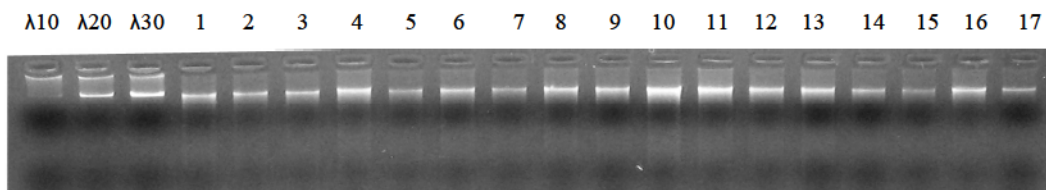


Fig. 37. Corsa elettroforetica su gel di agarosio dei campioni di DNA estratti dalle foglie di pioppo bianco. λ 10, λ 20, λ 30 corrispondono al DNA di fago λ rispettivamente di 10, 20 e 30 nanogrammi.

La figura 37 visualizza la corsa elettroforetica su gel di agarosio di alcuni campioni di DNA estratto dalle foglie di pioppo bianco. La stima della concentrazione del DNA è stata ottenuta per confronto dei DNA purificati con quella di fago λ , di cui è nota la quantità (espressa in nanogrammi: 10, 20 e 30). Grazie alla corsa elettroforetica è stato possibile anche apprezzare la qualità dei DNA estratti e l'assenza di contaminazioni da RNA e/o proteine. Il grado di purezza dei DNA è stato confermato anche dal rapporto dell'assorbanza alle lunghezze d'onda di 260 e 280 nm, risultato per tutti i campioni pari a circa 1,8. Il DNA ottenuto è stato quindi processato mediante metodica MSAP che, come descritto in M&M (Par. 2.3.2), prevede due digestioni parallele: una con la coppia di enzimi *EcoRI* – *MspI* (digestione 1), e l'altra con la coppia *EcoRI* – *HpaII* (digestione 2). I frammenti di DNA ottenuti dalle digestioni, in seguito ligati a specifici adattatori sono stati quindi amplificati attraverso due PCR consecutive. I prodotti delle due amplificazioni (PCR pre-selettiva e selettiva) sono stati visualizzati tramite corsa elettroforetica su gel di agarosio (Figg. 38-39).

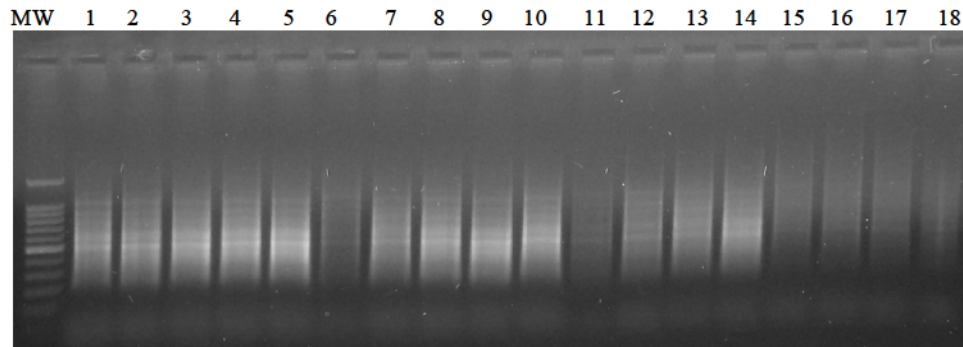


Fig.38. Corsa elettroforetica su gel di agarosio dei prodotti dell'amplificazione preselettiva. MW= marcatore di peso molecolare.

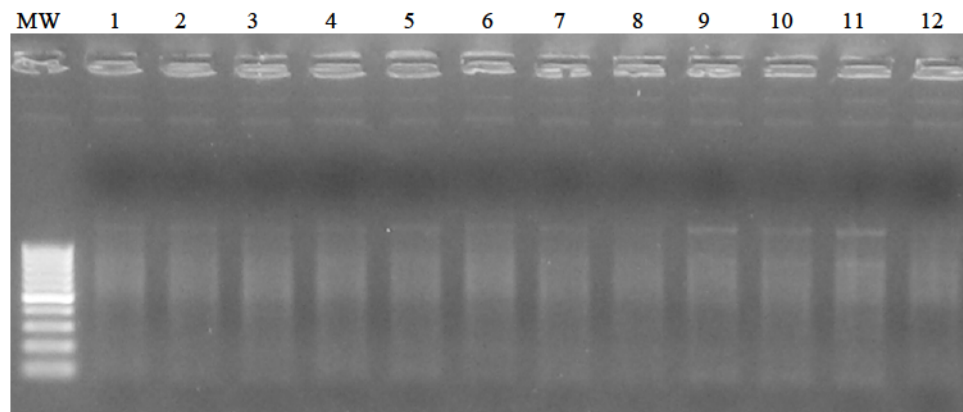


Fig. 39. Corsa elettroforetica su gel di agarosio dei prodotti dell'amplificazione selettiva. MW= marcatore di peso molecolare.

Le corse elettroforetiche su gel di agarosio dei prodotti dell'amplificazione pre-selettiva e selettiva hanno permesso di verificare l'avvenuta amplificazione. In seguito, i prodotti dell'amplificazione selettiva sono stati analizzati tramite corsa elettroforetica capillare su sequenziatore automatico e i risultati sono stati poi analizzati con il software Gene mapper. Le bande visualizzate sono state codificate in una matrice di presenza assenza elaborata come descritto in Materiali e Metodi (Par. 2.4.1). I dati ottenuti dalle analisi degli elettroferogrammi dei frammenti di DNA amplificati, prodotti con la digestione *EcoRI-HpaII*, sono stati elaborati indipendentemente da quelli ottenuti dalla digestione con *EcoRI-MspI*. L'analisi dei profili MSAP, ottenuta dalla digestione con la coppia di enzimi *EcoRI-MspI*, fornisce informazioni di tipo genetico (*MspI* è un enzima insensibile alla metilazione). I risultati delle elaborazioni statistiche dei dati genetici sono riportati di seguito.

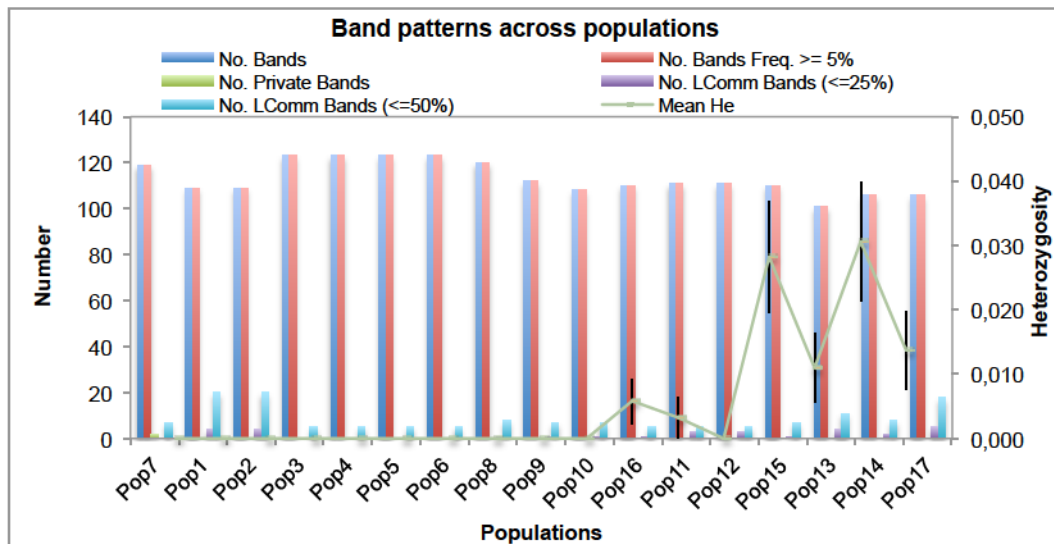


Fig. 40. I valori riportati in figura corrispondono a: No. Bands = Numero di bande, No. Bands Freq. $\geq 5\%$ = Numero di bande con una frequenza $\geq 5\%$; No. Private Bands = Numero di bande specifiche di una singola popolazione; No. LComm Bands ($\leq 25\%$) = Numero di bande comuni (Freq. $\geq 5\%$) trovate nel 25% o meno delle popolazioni; No. LComm Bands ($\leq 50\%$) = Numero di bande comuni (Freq. $\geq 5\%$) trovate nel 50% o meno delle popolazioni; Mean He = Eterozigosit  attesa media.

Il grafico riportato in figura 40 permette di visualizzare il numero delle bande di DNA delle diverse popolazioni indagate, il numero delle bande specifiche per ogni singola popolazione e quelle comuni, in base a percentuali variabili di condivisione tra le popolazioni ($\leq 25\%$ e $\leq 50\%$). E' possibile osservare che per ogni popolazione il numero di bande specifiche (barre verdi)   prossimo allo zero, al contrario, quelle condivise tra le popolazioni con una percentuale $\geq 5\%$ (barra rossa)   uguale al numero di bande totali in tutte le popolazioni. Questo risultato indica che le piante di pioppo bianco campionate presentano profili genetici alquanto simili. Il numero di bande comuni (al 25% e al 50%) tra le popolazioni campionate (rispettivamente barre viola e azzurra) rappresenta una percentuale variabile tra il 5 e il 10% delle bande totali, con un picco massimo nella popolazione 17. Questo dato suggerisce che le diverse popolazioni condividono una parte rilevante del genoma. L'eterozigosit  attesa, rappresentata dalla linea verde,   nulla per la maggior parte delle popolazioni, tuttavia fanno eccezione le popolazioni 13, 14, 15 e 17 che in ogni caso mostrano valori estremamente bassi nell'ordine di 10^{-2} . Questo dato non sorprende a causa della natura clonale delle popolazioni analizzate. Per stimare la varianza molecolare   stato utilizzato il test statistico AMOVA, i cui risultati sono illustrati di seguito. In particolare il grafico in figura 41 mostra la percentuale di varianza molecolare inter-popolazione rispetto a quella intra-popolazione, calcolate per i dati genetici.

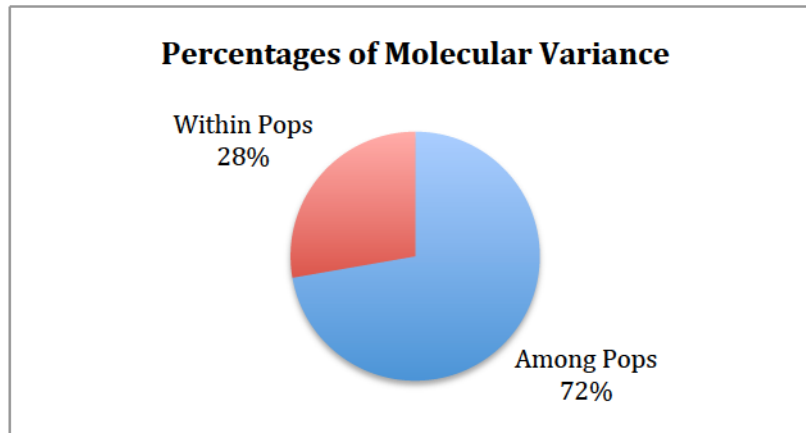


Fig. 41. Il grafico visualizza la percentuale di varianza molecolare tra le popolazioni (blu) e nella popolazione (rosso).

Il grafico a torta evidenzia come la varianza molecolare tra le popolazioni (72%) sia maggiore rispetto a quella intra-popolazione (28%), suggerendo la stabilità dei profili genetici di ogni pianta appartenente a ciascuna popolazione. Contestualmente al test AMOVA è stato eseguito il test di significatività statistica, che ha permesso di rigettare l'ipotesi nulla, secondo cui nessuna differenza esiste tra le popolazioni con un livello di significatività pari a 0,001. Analizzando il grafico della distribuzione degli PhiPT simulati (un analogo di Fst) rispetto ai valori reali (barra rossa), si evince che essi si posizionano fuori dalla distribuzione simulata (Fig. 42).

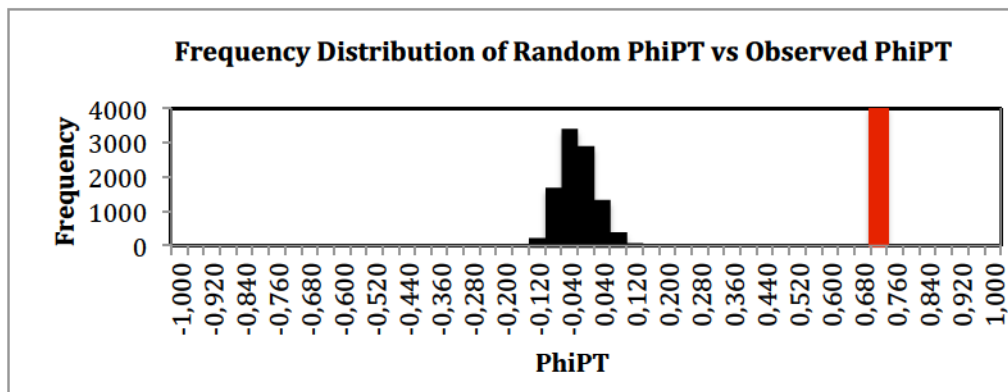


Fig. 42. Grafico della distribuzione dei PhiPT simulati (barre nere) e dei PhiPT reali.

Per analizzare la struttura genetica dell'intera popolazione di pioppi e stimare il numero di popolazioni che costituiscono la popolazione nel suo insieme, è stato utilizzato il software *Structure* (Par. 2.4.2). L'analisi statistica condotta sui dati molecolari, basata su simulazioni e rimescolamenti casuali, ha permesso di calcolare un numero di popolazioni ottimali (K) pari a 6. La simulazione finale, eseguita impostando per l'analisi di strutturazione delle popolazioni il valore K ottimale, ha fornito come *bar plot* la figura qui di seguito riportata (Fig. 43).

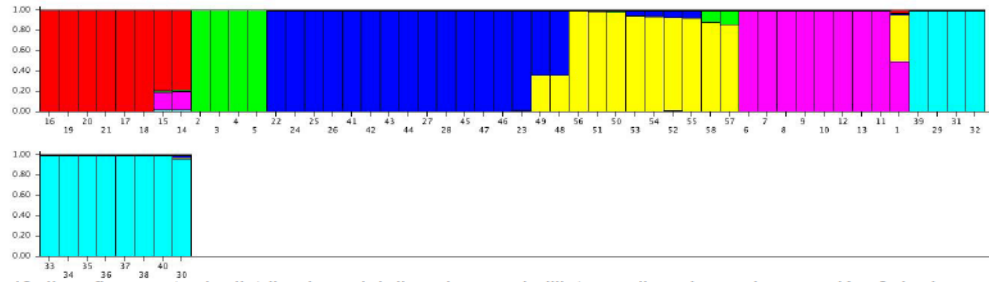


Fig. 43. Il grafico mostra la distribuzione dei diversi genomi all'interno di ogni campione per K = 6. La legenda dei campioni è riportata in Materiali e Metodi.

L'elaborazione statistica condotta mi ha permesso di stimare il numero ottimale di popolazioni dell'insieme di campioni e di analizzare la sottostruttura mediante la stima della proporzione di appartenenza a ciascun sottogruppo di ogni genotipo. In generale, dal grafico a barre è possibile notare la separazione dei diversi pool genici, costituiti da piante con genoma omogeneo. Fanno eccezione alcuni campioni (1, 14, 15, 48 e 49), che mostrano un genoma più eterogeneo. I dati ottenuti mediante un'analisi *whole genome* dimostrano quindi che le popolazioni di pino bianco della Sardegna sono di tipo clonale e confermano la bontà del precedente lavoro (Brundu et al. 2008). Grazie al software *Structure* e alla sua implementazione interna del software *PHYLIP* (Felsenstein 2005) è stato possibile costruire anche un albero di similarità raffigurante la distanza genetica tra le sei popolazioni stimate attraverso l'algoritmo *Neighbor joining* (Fig. 44) (Saitou and Nei 1987).

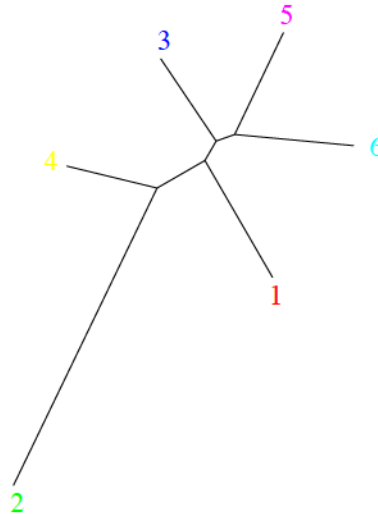


Fig. 44. Albero raffigurante la distanza genetica tra le sei popolazioni stimata mediante l'algoritmo *Neighbor joining*.

La stima di K è stata effettuata anche utilizzando un software diverso, implementato in R. In particolare è stato usato il programma Geneland che, oltre a stimare il numero ottimale di popolazioni in un insieme, permette di analizzare i dati genetici in relazione alle

coordinate geografiche (*Landscape Genetics*). I dati molecolari sono stati quindi elaborati per stimare il valore K, che dopo un elevato numero di permutazioni, è risultato essere pari a sette. Questo dato, anche se leggermente diverso da quello ottenuto in precedenza con l'analisi di *Structure*, non ci sorprende se si tiene conto dei diversi algoritmi matematici usati dai due software. Dopo aver stimato il numero di gruppi (popolazioni) è stata disegnata una mappa raffigurante la probabilità di ritrovare gli stessi in Sardegna (Fig. 45).

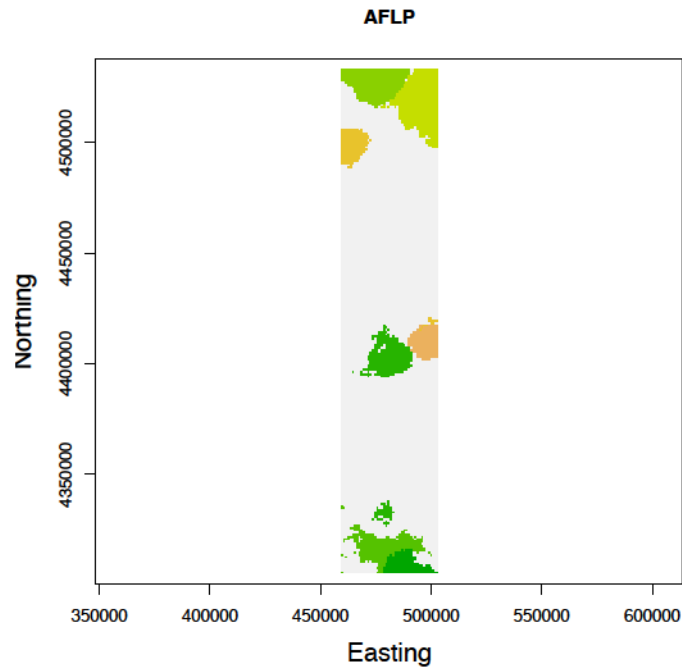


Fig. 45. Mappa delle probabilità di distribuzione delle popolazioni. E' stato assegnato un colore diverso ad ogni popolazione.

Per stimare in che modo le sette popolazioni siano connesse tra loro è stato disegnato un albero delle connessioni ottimali (*Minimum Spanning Tree*) in base alla distanza genetica. Il risultato è stato implementato con le coordinate geografiche e sovrainposto alla mappa della Sardegna importata da *Google Earth*. Il risultato è mostrato in figura 46.

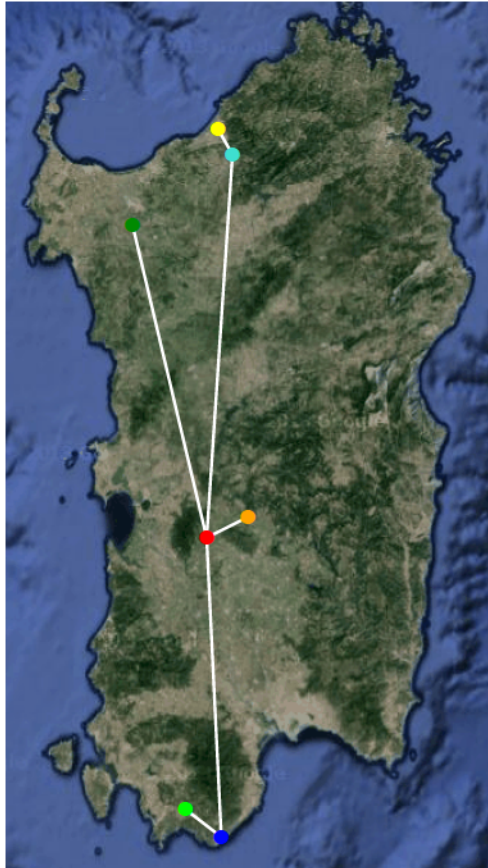


Fig. 46. L'immagine riporta la mappa della Sardegna su cui sono disegnati i centroidi delle sette popolazioni stimate e su cui è stato sovrainposto l'albero delle connessioni ottimali ottenuto in base alla distanza genetica.

La figura 46 mostra la mappa della Sardegna su cui è stato sovrainposto l'albero delle connessioni ottimali. Da questa immagine si evidenzia come tutte le sette popolazioni abbiano origine dalla popolazione centrale indicata con il colore rosso; in particolare nella zona centrale della Sardegna sono state campionate più popolazioni (Par. 2.1.2) che, grazie alle elaborazioni statistiche, sono state raggruppate in soli due gruppi separati dal punto di vista genetico, ma non molto distanti geograficamente. Quest'area della Sardegna è anche la meno antropizzata dell'isola; inoltre, proprio in quest'area, durante il nostro campionamento sono stati ritrovati gli unici due cloni di pioppo bianco di sesso femminile. Queste informazioni supportano l'ipotesi che dal centro della Sardegna si siano originate ancestralmente poi tutte le altre popolazioni dell'isola.

Dal momento che sono stati utilizzati due differenti software per determinare il numero ottimale di popolazioni e la loro struttura, sono state confrontate le rispettive *membership* allo scopo di verificare la riproducibilità e la bontà dei risultati ottenuti sia con *Structure* che con *Geneland* (Tab. 15) .

CONFRONTO DELLE MEMBERSHIP STRUCTURE E GENELAND						
1	2	3	4	5	6	7
2J22	3J1	23N/A	51,1J9	52,1J15?	55H22	66H22
7J20	4J1	24N/A	51,2J9	52,2J15?	59H22	67H22
8J20	5J1	PA29J1(2 004)	51,3J9	52,3J15?	60H22	
9N/A	6J1	PA30J1(2 004)	51,4J9	52,4J15?	61H22	
10N/A		PA31J1(2 004)	51,5J8	52,5J15?	62H22	
11J20		PA32J1(2 004)	51,6J9	52,6J15?	63H22	
12J20		PA33J1(2 004)	51,7J9	53,1J14	64H22	
13N/A		PA34J1(2 004)	54,1J9	53,2J14		
14N/A			54,2J9	53,4J14		
			54,3J9	53,4BISJ1 4		
			54,4J9	53,5J14		
			54,5J9	53,6J14		
			54,6J9			
			54,11J9			
			54,13J9			
			54,14J9			

Tab 15. La tabella mostra la *membership* ottenuta grazie ai software *Structure* e *Geneland*. Il colore verde indica corrispondenza tra i 2 software, mentre il colore giallo indica differenza.

Dalla tabella 15 si può notare come le *membership* stimate dai due software siano molto simili; l'unica differenza riguarda i campioni 66 e 67, che in *Structure* sono collocati nella popolazione sei, mentre con *Geneland* sono attribuiti alla popolazione sette. Questa differenza, che corrisponde a circa il 3,0%, è del tutto plausibile dato che i risultati sono stati ottenuti attraverso due programmi differenti con utilizzo di algoritmi diversi e dopo ricampionamento casuale dei dati (Par. 2.4.6).

I profili molecolari ottenuti sono stati ulteriormente elaborati con il software NTSys-PC che ha permesso di calcolare la matrice di similarità basata sul coefficiente di *Jaccard*. Alla matrice di similarità ottenuta con tale indice è stato applicato un algoritmo di *clusterizzazione* allo scopo di raffigurare tramite un dendrogramma i risultati ottenuti (Fig. 47).

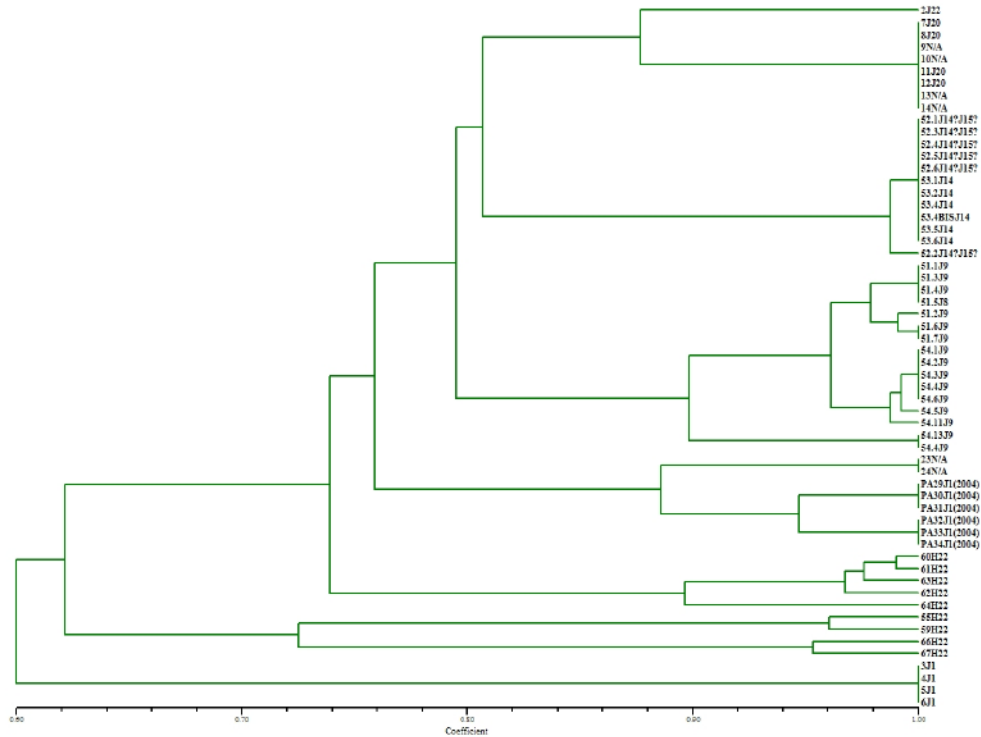


Fig. 47. Albero di similarità genetica ottenuto dalla matrice di similarità calcolata con l'indice di Jaccard. Sulle ascisse il coefficiente di similarità di Jaccard.

Dal dendrogramma riportato in figura 47 emerge ancora una volta l'elevata similarità genetica esistente tra i campioni di pioppo bianco sardi. Per molti campioni delle popolazioni indagate la similarità è addirittura massima (pari a 1), mentre il valore minimo è pari a 0,60, attribuito a un limitato numero di campioni raccolti nella parte meridionale (3, 4, 5 e 6) dell'isola.

La fase successiva dell'indagine da me condotta ha previsto lo studio dello stato di metilazione degli stessi campioni. I profili elettroforetici MSAP, ottenuti con la digestione *EcoRI-HpaII* (digestione sensibile allo stato di metilazione del DNA), sono stati codificati in una matrice binaria ed elaborati come descritto in precedenza per i dati genetici. La figura 48 illustra graficamente i vari indici calcolati mediante il software GenAEx (eterozigosità attesa, frequenza dei frammenti per singola popolazione, ecc.).

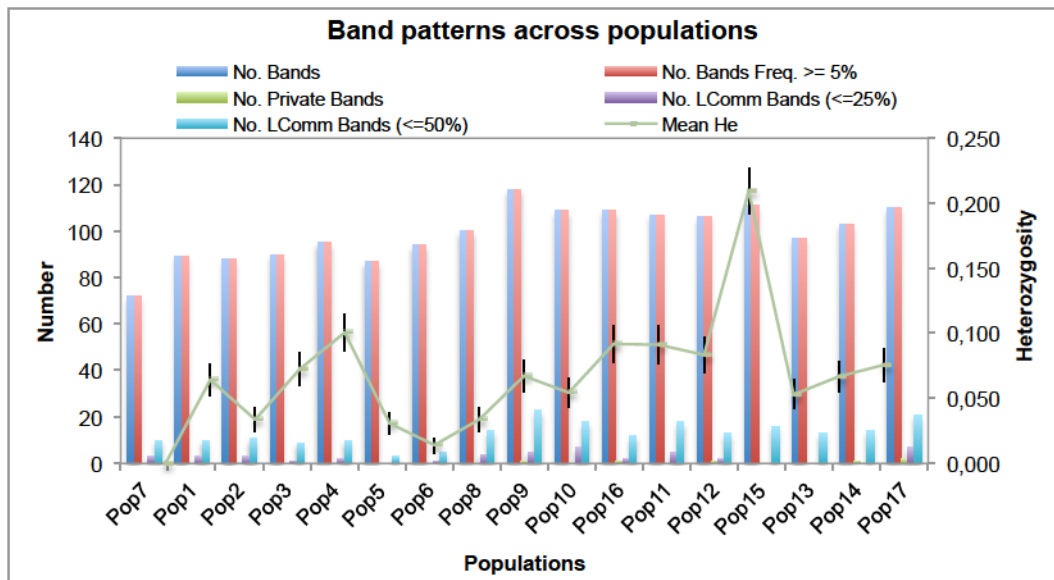


Fig. 48. I valori riportati in figura corrispondono a: No. Bands = Numero di bande, No. Bands Freq. >= 5% = Numero di bande con una frequenza >= 5%; No. Private Bands = Numero di bande specifiche di una singola popolazione; No. LComm Bands (<=25%) = Numero di bande comuni (Freq. >= 5%) trovate nel 25% o meno delle popolazioni; No. LComm Bands (<=50%) = Numero di bande comuni (Freq. >= 5%) trovate nel 50% o meno delle popolazioni; Mean He = Eterozigosità attesa media.

Il numero di bande private è risultato essere pari a 0 per tutte le popolazioni analizzate, al contrario di quelle condivise con una frequenza >=5%, che corrisponde al totale delle bande polimorfiche. Questo dato conferma che anche i profili epigenetici dei pioppi, appartenenti alle diverse popolazioni, condividono un rilevante numero di bande. Il numero di bande condivise tra le popolazioni (con una frequenza <=50%) supera in alcuni profili i 20 frammenti (Pop 9, 11 e 17) .

I profili epigenetici, ottenuti mediante la digestione *EcoRI/HpaII*, sono stati sottoposti al test statistico della varianza molecolare (AMOVA) allo scopo di valutarne la percentuale intra- e inter-popolazione. Il risultato è riportato nella seguente figura (Fig. 49).

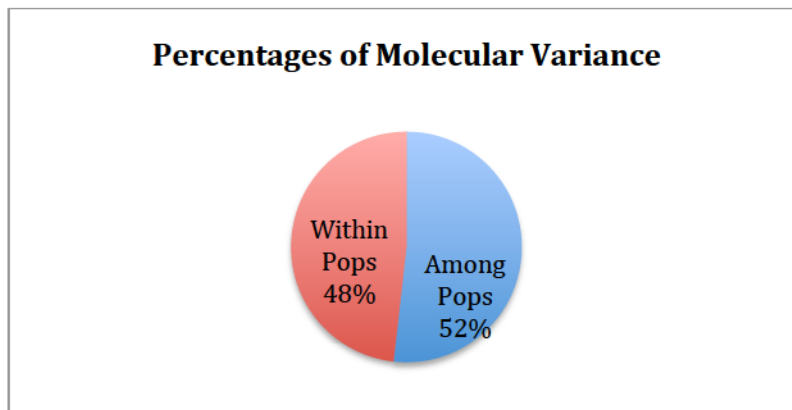


Fig. 49. Il grafico mostra la percentuale di varianza molecolare tra le popolazioni (blu) e nella popolazione (rosso).

Il grafico della figura 49 mostra come la percentuale di varianza intra-popolazione sia alquanto simile a quella inter-popolazione. Contestualmente al test AMOVA è stato

eseguito il test di significatività statistica che ha permesso di rigettare, anche in questo caso, l'ipotesi nulla (ovvero nessuna differenza tra le popolazioni) con un livello di significatività pari a 0,001; infatti, analizzando il grafico della distribuzione degli PhiPT simulati, si può notare che i valori reali (barra rossa) si posizionano al di fuori dalla distribuzione simulata, suggerendo l'esistenza di popolazione epigeneticamente diversificate nell'areale della Sardegna da me indagato (Fig. 50).

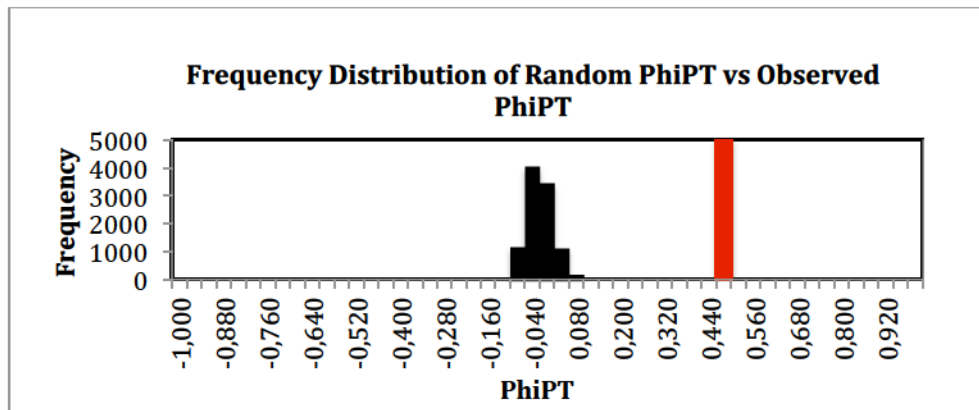


Fig. 50. Grafico della distribuzione dei PhiPT simulati (barre nere) e dei PhiPT reali (barra rossa).

Allo scopo di stimare il numero di popolazioni ottimali e come gli epigenomi identificati dall'analisi di strutturazione di queste popolazioni siano distribuiti nei singoli individui è stato utilizzato come in precedenza il software *Structure*. Come suddetto questa analisi statistica permette di stimare la struttura ottimale delle popolazioni e di assegnare ogni albero di pioppo bianco campionato a ciascuna popolazione individuata. Una prima elaborazione è stata quindi necessaria per stimare il numero ottimale di popolazioni (K) a partire dai profili epigenetici ottenuti. Il risultato di questa prima simulazione ha permesso di stimare il valore di K che è risultato essere pari a 4. La simulazione finale è stata quindi eseguita impostando il valore di K=4 (Fig. 51).

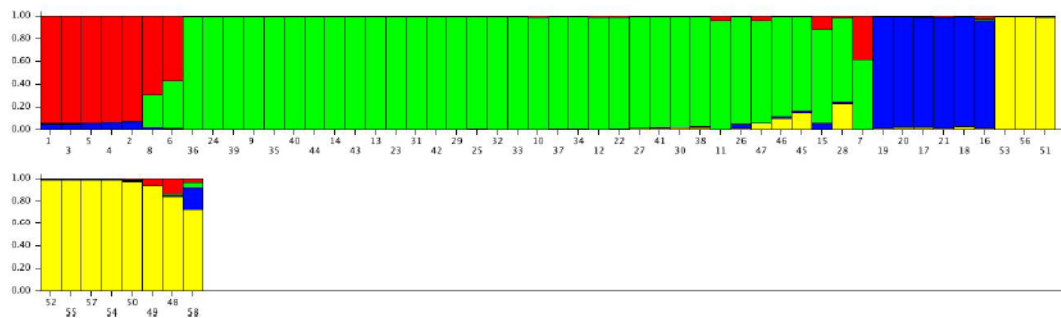


Fig. 51. Il grafico mostra la distribuzione dei diversi genomi all'interno di ogni campione per K = 4 a partire dai profili genetici influenzati dallo stato della metilazione. La legenda dei campioni è riportata in Materiali e Metodi (Par. 2.22).

L'analisi statistica con il software *Structure* ha permesso di stimare in quattro il numero di popolazioni ottimali. L'analisi suggerisce che lo stato di metilazione è stabile tra le popolazioni. Inoltre va evidenziato che il valore di K ottenuto dall'analisi dei dati genetici corrispondeva a sei, mentre quello stimato sui dati epigenetici è pari a quattro.

Il software *Geneland*, utilizzato per l'analisi dei dati genetici, è stato impiegato anche per elaborare i dati epigenetici e ha permesso, anche in questo caso, di stimare il numero di popolazioni e di calcolare, grazie alle coordinate geografiche, la probabilità di distribuzione sul territorio delle popolazioni. Anche l'elaborazione *Geneland* ha stimato un valore di K pari a quattro. In figura 52 è riportata una mappa della distribuzione delle quattro popolazioni.

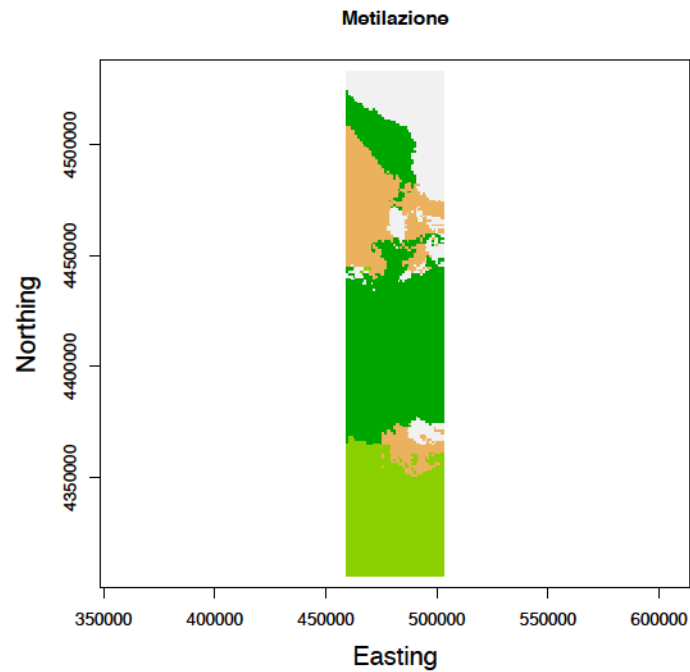


Fig. 52. Mappa delle probabilità di distribuzione delle popolazioni di pioppo bianco della Sardegna.

La figura 52 mostra la probabilità di ritrovare la medesima popolazione nelle aree rappresentate con ugual colore. Il risultato è stato implementato con le coordinate geografiche e sovrapposto alla mappa della Sardegna, importata da *Google Earth*, e illustrata in figura 53.

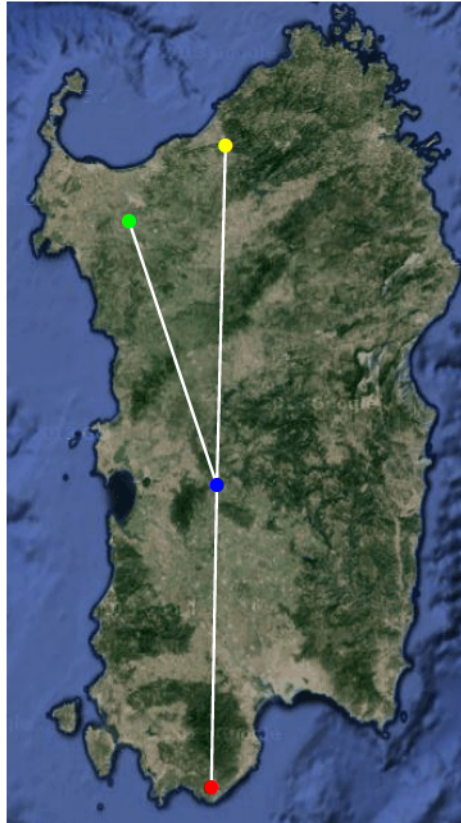


Fig. 53. L'immagine riporta la mappa della Sardegna su cui sono disegnati i centroidi delle quattro popolazioni stimate e su cui è stato sovrainposto l'albero delle connessioni ottimali ottenuto in base alla distanza genetica.

La figura 53 mostra i centroidi delle popolazioni raggruppate, rappresentativi della moda dei profili genetici influenzati dallo stato di metilazione, e le loro connessioni in funzione della distanza genetica. Grazie a questa analisi siamo stati in grado anche di identificare quali sono le popolazioni che si raggruppano e che riducono il valore di K da sette a quattro.

Anche in questo caso come per i dati genetici sono state confrontate le *membership* di appartenenza ai sottogruppi, attribuite dai due software utilizzati per le analisi statistiche ai diversi campioni indagati. Il risultato è mostrato in tabella 16.

MSAP CONFRONTO DELLE MEMBERSHIP STRUCTURE E GENELAND			
1	2	3	4
2J22	23N/A	PA29J1	55H22
3J1	24N/A	PA30J1	59H22
4J1	51,1J9	PA31J1	60H22
5J1	51,2J9	PA32J1	61H22
6J1	51,3J9	PA33J1	62H22
7J20	51,4J9	PA34J1	63H22
8J20	51,5J8		64H22
9N/A	51,6J9		66H22
10N/A	51,7J9		67H22
11J20	52,1J15?		
12J20	52,2J15?		
13N/A	52,3J15?		
14N/A	52,4J15?		
	52,5J15?		
	52,6J15?		
	53,1J14		
	53,2J14		
	53,4J14		
	53,4BJ14		
	53,5J14		
	53,6J14		
	54,1J9		
	54,2J9		
	54,3J9		
	54,4J9		
	54,5J9		
	54,6J9		
	54,11J9		
	54,13J9		
	54,14J9		

Tab 16. La tabella riporta la *membership* ottenuta con i software *Structure* e *Geneland*. Il colore verde indica corrispondenza dei valori ottenuti dalle elaborazioni indipendenti dei due software, mentre il colore giallo indica differenza.

Come evidenziato dalla tabella 16 non esistono sostanziali differenze tra le due *membership*, infatti le uniche discrepanze riguardano i campioni 8, 11, 12, 13 e 14, che rappresentano circa il 10% dei campioni.

Per stimare il grado di similarità del profilo epigenetico tra i vari pioppi campionati è stata calcolata la matrice di similarità utilizzando l'indice di Jaccard come fatto in precedenza per i dati genetici. Il dendrogramma ottenuto è illustrato in figura 54.

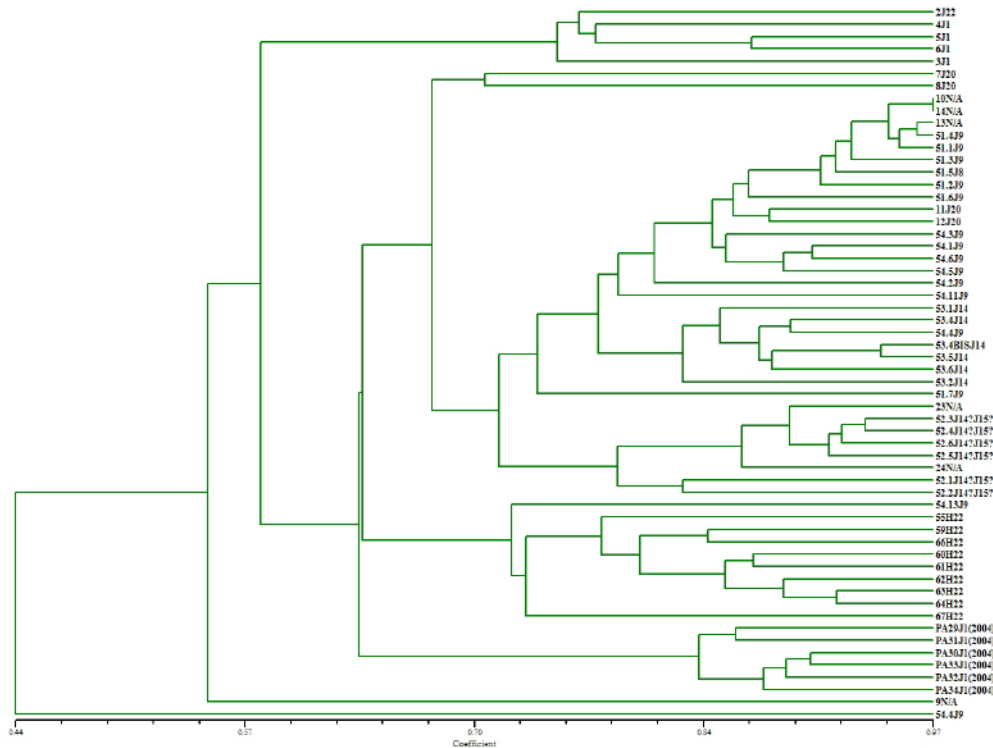


Fig. 54. Albero di similarità epigenetica ottenuto dalla matrice di similarità calcolata con l'indice di Jaccard. Sulle ascisse il coefficiente di similarità di Jaccard.

I valori di similarità, pur essendo elevati, raggiungono il massimo solamente in due campioni (9 e 14).

Infine è stato utilizzato l'algoritmo, elaborato dal Dr. Pérez-Figueroa (Pérez-Figueroa 2013)(Par. 2.4.6) e applicato ai dati epigenetici delle popolazioni di pioppo bianco della Sardegna allo scopo di stimare la frazione di frammenti MSAP suscettibili o meno a fenomeni di metilazione (Tab. 17). Questo pacchetto di R (*MSAP*) permette anche l'analisi delle componenti principali (PCoA).

Analisi dello stato di metilazione				
Report del livello di metilazione				
Popolazioni	A	B	C	D
HPA+/MSP+ (non metilato)	0.37	0.32	0.44	0.33
HPA+/MSP- (emi-metilato)	0.11	0.22	0.20	0.23
HPA-/MSP+ (metilazione della citosina interna)	0.40	0.36	0.25	0.31
HPA-/MSP- (metilazione completa o assenza del frammento)	0.12	0.10	0.11	0.14

Tab. 17. Analisi del grado di metilazione delle quattro popolazioni stimate dall'analisi dei profili di metilazione.

La tabella 17 riassume i dati relativi al grado di metilazione del genoma attraverso il confronto dei profili ottenuti dalle digestioni *EcoRI/MspI* e *EcoRI/HpaII*. La tecnica MSAP permette di distinguere i principali stati di metilazione del DNA, in funzione dell'assenza/presenza delle bande (frammenti di DNA) ottenute con la digestione *EcoRI-MspI* (digestione 1) rispetto alla presenza/assenza delle bande ottenute con la digestione *EcoRI-HpaII* (digestione 2). In particolare quando il frammento è presente in entrambe le digestioni (1 e 2) esso si intende non metilato. Si assume che il frammento sia emi-metilato quando è presente nel caso della digestione 2 e assente nel caso della digestione 1; quando accade invece il contrario si stima che sia la C interna a essere metilata, mentre nel caso in cui il frammento sia assente in entrambe le digestioni (1 e 2) si intende non informativo o comunque iper-metilato. Dalla tabella 17 si evince che il maggior grado di metilazione riguarda le C interne. Inoltre, anche dal confronto statistico eseguito con il software "MSAP", emerge che le quattro popolazioni indagate per il loro stato di metilazione risultano essere tra loro significativamente differenti ($p < 0.001$) in base ai polimorfismi dello stato di metilazione, mentre il confronto risulta statisticamente non significativo considerando i profili non sensibili alla metilazione. Nella figura 55 è mostrato il risultato dell'analisi delle componenti principali (PCoA) ottenuta elaborando i dati epigenetici relativamente ai frammenti indipendenti dallo stato di metilazione. Il dato conferma che dal punto di vista genetico le quattro popolazioni stimate sono molto simili.

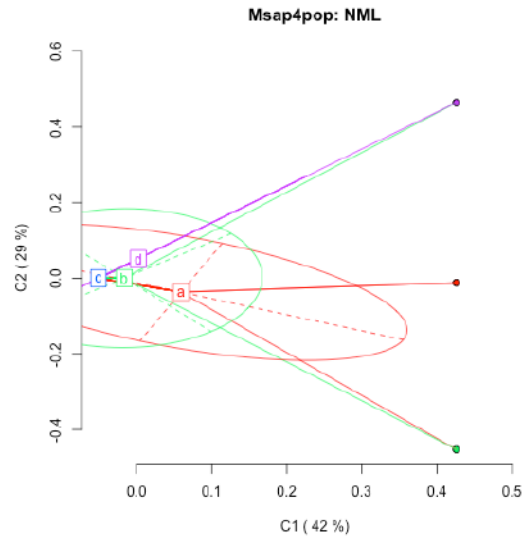


Fig. 55. Analisi delle componenti principali ottenuta dai profili genetici insensibili alla metilazione. C1 e C2 sono indicatori della varianza. Le lettere rappresentano le quattro diverse popolazioni; l'asse maggiore dell'ellisse indica la massima dispersione e l'asse minore la minima dispersione.

Le quattro popolazioni sono sovrapposte, sebbene gli assi delle ellissi e i valori di C1 e C2 indichino un'elevata varianza molecolare. Questo dato conferma, ancora una volta, la considerevole somiglianza genetica tra queste popolazioni.

In figura 56 è riportata la PCoA ottenuta invece dall'analisi dei polimorfismi sensibili allo stato di metilazione.

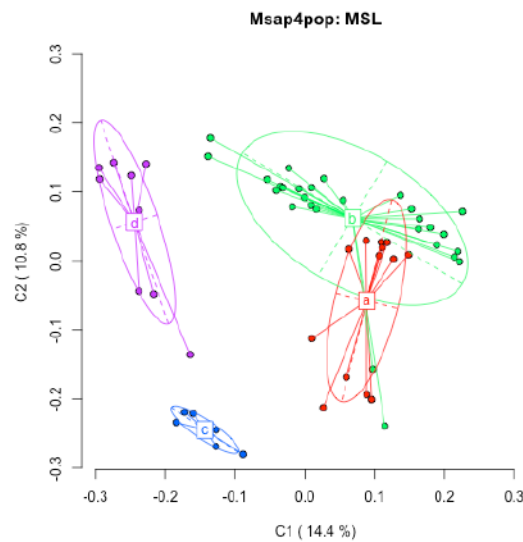


Fig. 56. Analisi delle componenti principali ottenuta dai profili genetici insensibili alla metilazione. C1 e C2 sono indicatori della varianza. Le lettere rappresentano le quattro diverse popolazioni; l'asse maggiore dell'ellisse indica la massima dispersione e l'asse minore la minima dispersione.

In questo caso, si può rilevare come le quattro popolazioni siano ben distinte e anche ben separate tra di loro, ad eccezione di a e b che presentano alcune zone di sovrapposizione. Inoltre la varianza intra-popolazione è assai ridotta considerando le dimensioni degli assi delle ellissi e i valori di C1 e C2.

Capitolo IV

Discussione

Il progetto di dottorato precedentemente illustrato è rivolto specificatamente allo studio della biodiversità genetica vegetale. In particolare sono stati analizzati ecotipi orticoli autoctoni della regione Campania ma anche popolazioni naturali di pioppo bianco della Sardegna. L'analisi della biodiversità degli ecotipi orticoli campani ha previsto uno studio molecolare su diversi ecotipi appartenenti specie vegetali: aglio, cipolla, mais, scarola e zucchino. Lo studio degli ecotipi orticoli campani rientra in un progetto quinquennale co-finanziato dall'Unione Europea e dalla regione Campania dal titolo AGRIGENET (Network per la Salvaguardia e la Gestione delle Risorse Genetiche Agro-Alimentari Campane). AGRIGENET si propone di indagare i profili genetici degli ecotipi orticoli campani in modo da caratterizzare il germoplasma, che sarà conservato presso Università ed enti di ricerca Campani, allo scopo di allestire una banca di DNA. Lo scopo della ricerca ha previsto la stima del livello di biodiversità genetica degli ecotipi mediante marcatori co-dominanti SSR, e la conservazione del patrimonio genetico (DNA) presso il Dipartimento di Chimica e Biologia dell'Università di Salerno in super-congelatori – 80°C. Sulla base della letteratura scientifica, circa 10 loci SSR sono stati selezionati tra i più informativi per ciascuna specie indagata. Le elaborazioni statistiche eseguite sui profili molecolari hanno permesso di stimare che per tutti gli ecotipi esiste un'elevata biodiversità genetica, sia inter-popolazione che intra-popolazione, ad eccezione dei due ecotipi di aglio, "Schiacciato" e "Tondo di torella". In questo caso tutti gli indici calcolati e le elaborazioni biostatistiche volte a definire la struttura della popolazione hanno suggerito che i due ecotipi in realtà sono geneticamente identici e evidenziando pertanto un caso di sinonimia. Questa ipotesi è avvalorata dal fatto che i fiori di aglio sono sterili così come i semi. L'unica via di propagazione e produzione è affidata alla divisione e alla piantumazione dei singoli bulbi. In tal modo si ha quindi una riproduzione clonale e vegetativa (Keller and Senula 2013). A ciò va aggiunto che i cloni di aglio possiedono un adattamento specifico alle differenti regioni agro-pedo-climatiche (McCollum 1976; Figliuolo et al. 2001). Sebbene l'aglio impiegato per scopi agronomici sia coltivato e propagato vegetativamente in tutto il mondo, ne è stata studiata l'origine che è risultata probabilmente localizzata in Asia centrale (Hong and Etoh 1996; Vavilov 2009). Lo studio condotto mi porta a concludere che le accessioni di aglio analizzate non mostrano nessuna varianza molecolare tra gli ecotipi, evidenziando che non vi sono sostanziali differenze genetiche tra i due ecotipi. Dai profili sono emersi dei valori di eterozigosità osservata unitari per la maggior parte dei loci SSR analizzati. Questo dato supporta la tesi di Ipek et al. (2008) secondo cui l'aglio ha un genoma eterogeneo. I nostri dati, a differenza di quelli ottenuti da Zhao et al. (2011) che ha analizzato i medesimi loci SSR in diversi campioni di aglio, non hanno evidenziato nessuna iper-variabilità; il valore del PIC è circa la metà di quello ottenuto da Zhao et al. (2011) confermando l'esigua biodiversità genetica esistente tra questi due ecotipi di aglio. Anche la matrice di similarità di Jaccard, raffigurata dal dendrogramma illustrato nella sezione dei **Risultati** mostra l'analogia totale dei due ecotipi indagati, ad eccezione del campione a4 che presenta una similarità genetica pari a 0,83 rispetto a tutti gli altri campioni indagati. Questa unica variazione di similarità, imputabile a un singolo locus SSR, è dovuta a mutazione somatica, fenomeno riscontrabile soprattutto nelle specie con forte o esclusiva propensione alla riproduzione vegetativa (e.g. pioppo e salicacee in genere). La situazione cambia se consideriamo le altre quattro specie orticole da me indagate. Nel caso della cipolla, infatti, i valori dell'eterozigosità attesa, dell'indice di

fissazione e del PIC evidenziano come gli ecotipi non siano geneticamente stabilizzati. Infatti, i valori di F suggeriscono fenomeni di incroci, alleli nulli, o un eccesso di eterozigotità dovuta probabilmente a selezione basata su tecniche tese a esaltare il fenomeno dell'eterosi. Fenomeno comunemente sfruttato dai programmi di ibridazione genetica che consistono in un evento dove due genomi distinti vanno incontro a una fase di "turbolenza genetica" prima di entrare in una fase di omeostasi. Questa fase di "Shock Genomico", proposto da McClintock (1984), potrebbe essere in relazione con l'espressione genetica e con l'attivazione di trasposoni (Baranwal et al. 2012). La cipolla come illustrato nella sezione **Materiali e Metodi** si riproduce per via sessuata e la fecondazione è operata dagli insetti pronubi, in particolar modo dalle api. Questa caratteristica potrebbe spiegare i valori di eterozigotità osservata maggiori di quella attesa. I risultati ottenuti sono comunque in linea con quelli illustrati dal lavoro di Baldwin et al. (2012). Analizzando i quattro ecotipi di cipolla considerati dal progetto AGRIGENET appare evidente come, per la maggior parte dei loci esaminati, lo scambio genetico sia compatibile con fenomeni di flusso genico naturale e casuale (per esempio i loci 1, 3 e 5 della cipolla "Febbraiese"). Anche i valori stimati di D_{est} supportano questa tesi variando da un minimo di 0 del locus 2, ad un massimo di 0.55 del locus 4. I risultati ottenuti per la cipolla "Marzatica" e per alcuni loci degli altri ecotipi evidenziano fenomeni di ibridazione compatibili con tecniche di selezione basate ancora una volta sull'eterosi (locus 7 pop 2 e locus 3 pop 4). Le varietà testimone a loro volta non risultano varietà stabilizzate, ma sembrano sottoposte a fenomeni di incroci casuali, come si evince dai valori di F e di PIC. La cipolla spesso va in contro a fenomeni di autoimpollinazione, infatti ha evidenziato una rilevante depressione da *inbreeding* addirittura maggiore di quella riportata nel lavoro pubblicato da Vigouroux et al. (2008) per il mais, che è considerata una delle specie più tollerante a tale fenomeno. Il livello di impollinazione dovuto a fenomeni di out-crossing è stato saggiato da Meer and Bennekom (1972) utilizzando due varietà di cipolla, una a bulbo giallo (carattere recessivo), contro una a bulbo rosso (carattere dominante), ed è stato evidenziato che, nell'arco dei quattro anni di studio, essa può variare dal 29-92% in relazione a variazioni climatico-ambientali che avvengono stagionalmente. Le elaborazioni biostatistiche eseguite sulle accessioni degli ecotipi di cipolla da me analizzati mediante il software Structure e NT-Sys PC, e illustrate nella sezione **Risultati**, confermano la notevole variabilità genetica ma anche l'elevato rimescolamento genetico, che comunque caratterizza anche le accessioni delle varietà testimone.

Gli ecotipi di mais, sebbene in maniera differente rispetto alle altre specie ed ecotipi da me indagati, mostrano una variabilità molecolare considerevole. Il mais riveste un'importanza agroalimentare strategica per tutti i paesi industrializzati ma non solo, in quanto, oltre al consumo umano, la sua granella è usata anche per l'allevamento del bestiame sia da macello che da latte. In questo contesto, lo studio della biodiversità e la caratterizzazione del germoplasma locale, appaiono azioni necessarie nell'ambito di una strategia che miri alla conservazione e tutela oltre che allo sfruttamento delle risorse genetiche del posto, indispensabili al miglioramento della qualità, resistenza e sostenibilità della coltura. In questo studio sono stati indagati sei ecotipi di mais mediante sette loci SSR. Tutti gli ecotipi e i loci analizzati mostrano valori di eterozigotità attesa e osservata confrontabili, tuttavia i valori dell'indice di fissazione evidenziano fenomeni di accoppiamento casuale. Per ogni locus genico sono stati riscontrati in media sei alleli in accordo con quanto riportato da Ranatunga et al. (2009) e da (Pejic et al. 1998); al contrario Molin et al. (2013) hanno osservato in media circa 3 alleli per locus. I valori del PIC riportati nella sezione **Risultati** variano in relazione ai loci analizzati ma, in media, sono paragonabili a quelli riportati da Adeyemo et al. (2011). Anche in questo caso gli autori hanno osservato una biodiversità genetica elevata tra le piante di mais analizzate, come evidenziato dai valori della distanza genetica. Valori simili a quelli illustrati nel lavoro sopracitato sono stati ottenuti anche nel presente studio, ove la varianza molecolare intra-ecotipo si è rivelata

circa il doppio di quella inter-ecotipo. Nonostante l'elevata varianza molecolare esistente tra i sei ecotipi indagati, l'analisi di strutturazione delle popolazioni ha permesso di raggrupparli in due macrogruppi. Questo risultato ottenuto con il software Structure non deve trarre in inganno; infatti, è probabile che la *clusterizzazione* in due gruppi separati sia dovuta proprio all'elevata biodiversità genetica, che non permette una netta separazione delle popolazioni per valori di K maggiori. Questo risultato è stato confermato dalla rappresentazione grafica della similarità genetica rappresentata graficamente dal dendrogramma riportato nella sezione dei **Risultati**. Nel dendrogramma, oltre a dimostrare la presenza di due macrogruppi, è anche possibile osservare quale sia il grado di similarità tra le piante delle popolazioni dei sei ecotipi. Il dendrogramma, infatti, evidenzia come i sei ecotipi di mais analizzati non siano linee stabilizzate e diversificate. Esse presentano molti e diversi alleli all'interno degli stessi ecotipi, tanto che la biodiversità intra-ecotipo è addirittura maggiore di quella inter-ecotipo. Nonostante l'elevata biodiversità genetica riscontrabile dal dendrogramma, è possibile notare come alcuni ecotipi, che potevano sembrare dai loro nomi delle sinonimie, sono chiaramente separati. In particolare i campioni dell'ecotipo "mais bianco di Acerra" sono raggruppati in un cluster disgiunto da quello formato dagli ecotipi "Spiga napoletana bianca" e "Spiga bianca", a loro volta distinti in due sottogruppi. Tuttavia, a testimonianza della variabilità genetica che caratterizza le piante analizzate molecolarmente mediante gli SSR, si può osservare dal dendrogramma un cluster basale dove si raggruppano differenti ecotipi. I dati ottenuti sono in accordo tra loro, e confermano l'oggettiva difficoltà di separare i sei ecotipi di mais. Tuttavia grazie a questi marcatori è stato possibile di stimare un elevato grado di biodiversità genetica presente tra gli ecotipi campani di mais.

Le altre due specie studiate sono la scarola e lo zucchini. Per la scarola è stato analizzato un unico ecotipo chiamato scarola "Riccia schiana". La scarola è una pianta diploide ($2n = 18$) e allogama (Kingsley R. Stern et al. 2009), la cui morfologia florale favorisce lo scambio genetico. Questa specie è stata per lungo tempo usata come una pianta medicinale (Kingsley R. Stern et al. 2009) e ora è coltivata per numerose e differenti impieghi, tra cui quello alimentare. A seconda dei luoghi e delle tradizioni colturali, sono stati selezionati e prodotti diversi ecotipi; oltre a quello studiato e analizzato in questo progetto è sicuramente famoso e conosciuto il radicchio, coltivato principalmente in nord Italia. Come per gli altri ecotipi anche in questo caso sono stati selezionati e amplificati un numero sufficientemente informativo di loci SSR. L'analisi molecolare è stata condotta sul DNA estratto dalle foglie campionate da 10 piante di scarola "Riccia schiana". Analizzando i risultati ottenuti possiamo concludere che è presente una discreta variabilità genetica tra le piante indagate, probabilmente favorita dal meccanismo riproduttivo. Gli indici di eterozigosità attesa e osservata calcolati sui dati molecolari risultano paragonabili; inoltre l'eterozigosità osservata risulta essere molto bassa a testimonianza dell'elevata omozigosi dei loci indagati. Anche l'indice di Shannon suggerisce che la biodiversità intra-ecotipo sia ridotta. Il dendrogramma presente nella sezione **Risultati** mostra come, ad eccezione dei campioni 1, 9 e 10, le piante di scarola costituiscano un gruppo abbastanza omogeneo con valori di similarità elevati. Per sette campioni su 10, infatti, la percentuale di similarità è pari a circa l'85%. Confrontato il numero di alleli individuati e l'intervallo delle dimensioni possiamo concludere che i nostri dati sono in linea con quelli ottenuti per gli stessi loci da Cadalen et al. (2010). In particolare l'intervallo di peso molecolare dei vari alleli corrisponde nella maggior parte dei loci analizzati, mentre il numero di alleli, per alcuni loci, si dimezza o, come nel caso del locus 2, che è omozigote per tutti i campioni, si riduce a 1 rispetto ai 5 amplificati da Cadalen et al. (2010).

Gli ultimi due ecotipi indagati appartengono alla specie *C. pepo*. Lo zucchini è coltivato in tutto il mondo utilizzando differenti varietà, ognuna adattata al clima e all'ambiente della regione in cui è coltivata (Robinson and Decker-Walters 1997). Gli ecotipi di zucchini

studiati sono esempi di questo adattamento, in quanto essi mantengono le caratteristiche organolettiche legate all'ambiente in cui sono coltivati. I due ecotipi di zucchini analizzati chiamati "Cilentano" e "San Pasquale" hanno evidenziato valori di eterozigosità osservata mediamente bassi per tutti i loci SSR analizzati e in linea con quelli della eterozigosità attesa, anche se lo zucchini "San pasquale" per i loci 1 e 7 ha mostrato valori di eterozigosità massimi. L'indice di Shannon e i valori di PIC confermano che i due ecotipi non sono caratterizzati da elevata biodiversità, probabilmente per via di una selezione ancestrale indirizzata alla stabilizzazione genetica e fenotipica delle linee. Infatti il valore medio del PIC è circa la metà di quello osservato da Barzegar et al. (2013), mentre l'indice di Shannon è un terzo circa. Anche dal dendrogramma riportato nella sezione **Risultati** emerge come i campioni mostrino un'elevata similarità, soprattutto le piante dell'ecotipo zucchini "Cilentano". Barzegar et al. (2013), al contrario, ha ottenuto, esaminando 26 accessioni di *C. pepo*, con 14 primer SSR, valori di similarità ancora minori e con un massimo pari a circa 0,5. Il numero di alleli per locus ottenuti nello studio da me condotto varia da un minimo di uno a un massimo di cinque in linea con i dati ottenuti da Paris et al. (2003). Yu et al. (1999) riportano che il numero di alleli, evidenziabili attraverso marcatori SSR, è direttamente correlato al numero di ripetizioni: maggiore è la lunghezza del *repeat* maggiore risulterà il numero di polimorfismi. Nello studio qui condotto la lunghezza delle ripetizioni è compresa tra sei e otto basi nucleotidiche, tuttavia il locus 2 si è rilevato addirittura monomorfo. Le considerazioni fatte da Yu e collaboratori (1999) sono contraddette in una qual misura da Barzegar et al. (2013) che ha evidenziato un maggior numero di alleli polimorfici in *C. pepo* in una ripetizione dinucleotidica (AG)₈. In conclusione, le analisi molecolari condotte sugli ecotipi zucchini hanno permesso di stimare il livello di biodiversità esistente tra i due ecotipi e da esse emerge una buona stabilizzazione delle linee volta a ottenere colture omogenee.

Il progetto di dottorato ha previsto lo studio di popolazioni naturali di *P. alba* della Sardegna precedentemente studiate dal gruppo di ricerca del Prof. Castiglione e del Dr. Brundu mediante marcatori molecolari SSR. I risultati di quello studio hanno indicato una prevalenza quasi assoluta di riproduzione vegetativa su quella sessuata. Inoltre in Brundu et al. (2008) si è sottolineato come le popolazioni sarde di pioppo bianco siano geneticamente isolate da quelle dell'Italia continentale, e in particolare da quelle del parco del Ticino. I campioni di pioppo analizzato provenivano, oltre che dalla Sardegna, dal Parco Regionale Lombardo del Ticino, dalla Sicilia, dalla Macedonia e da altre regione dell'Europa meridionale. Questi profili erano stati quindi confrontati con campioni di riferimento di *P. tremula* provenienti dalla Scozia. Alla luce dei risultati pregressi, i dati molecolari sono stati rielaborati nell'ambito del progetto di dottorato, attraverso il software *Structure* allo scopo di analizzare la struttura dell'intero gruppo di pioppi campionati senza utilizzare alcuna conoscenza a priori riguardo il numero di cluster presenti nel data set. Una volta stimato il numero di cluster (K) l'analisi è stata ripetuta per ottenere informazioni sul grado di rimescolamento genetico delle piante indagate. Il risultato ha confermato l'isolamento dei pioppi bianchi sardi rispetto a quelli di riferimento scozzesi e a quelli del parco del Ticino. Tuttavia i campioni di *P. alba* sardi sembrano condividere gran parte del genoma con alcuni pioppi bianchi del bacino Mediterraneo (Sicilia, Rodi e Corsica). Come suggerito da Petit et al. (2002), sulla base di profili SSR cloroplastici, la Sardegna potrebbe essere stata un rifugio di alcune specie vegetali durante l'ultima era glaciale, come nel caso studiato di una specie di quercia. Il pioppo bianco potrebbe aver seguito lo stesso destino di quella specie di quercia rifugiandosi nelle regione dell'Europa meridionale come fatto anche dal pioppo tremulo e dal pioppo nero (Petit et al. 2003; Cottrell et al. 2005). L'analisi presentata nel lavoro di Brundu et al. (2008) ha, infatti, evidenziato la presenza di soli tre aplotipi distinti da quelli di tutti gli altri campioni analizzati. Gli aplotipi di pioppo bianco del bacino del Mediterraneo, ad eccezione di quelli sardi, potrebbero

essersi diffusi in tutto il bacino e in altri siti dell'Europa dalle zone di rifugio come dimostrano i reperti paleo-botanici ritrovati nel sud della Francia da Roiron et al. (2004). La particolare struttura genetica delle popolazioni di *P. alba* della Sardegna ha ispirato il progetto di dottorato illustrato in questa tesi. La natura clonale di queste è idonea a valutare se, e come, differenti condizioni ambientali possano alterare lo stato di metilazione del DNA, che a sua volta, attraverso la regolazione dell'espressione genica, permette l'adattamento all'ambiente attraverso risposte a breve-medio termine (Zhao and Zhou 2012), rispetto alla classica, e ben più documentata, mutazione genetica implicata nella selezione naturale operata nel corso dei millenni (Darwin 1872). A questo scopo è stato eseguito un nuovo campionamento dei pioppi sardi che ha visto non soltanto il recupero di materiale fogliare dalle popolazioni indagate nel precedente lavoro, ma anche da piante raccolte in nuovi siti dell'isola. Le popolazioni di pioppo bianco sono state analizzate attraverso la metodica molecolare *whole genome* chiamata MSAP che permette di stimare la biodiversità genetica indagando tutto il genoma, e allo stesso tempo, di verificare la stabilità o la variabilità dello stato di metilazione del DNA. I risultati della variabilità genetica confermano i dati ottenuti in precedenza da Brundu et al. (2008), infatti il numero di frammenti comuni tra le popolazioni con una percentuale $\geq 5\%$ è uguale al numero di bande totali in tutte le popolazioni, mentre il numero di bande specifiche per ogni popolazione è prossimo allo zero. Il numero di bande comuni (al 25% e al 50%) tra le popolazioni campionate rappresenta una percentuale variabile tra il 5 e il 10% delle bande totali, con un picco massimo nella popolazione 17. Questo dato suggerisce che le diverse popolazioni condividono una parte rilevante del genoma. L'eterozigosità attesa è nulla per la maggior parte delle popolazioni, tuttavia fanno eccezione le popolazioni 13, 14, 15 e 17 che in ogni caso mostrano valori estremamente bassi nell'ordine di 10^{-2} . Questo dato suggerisce che i pioppi, diffusi in tutta la Sardegna, condividono una parte rilevante del genoma, e, mostra una ridotta biodiversità genetica ancora maggiore di quella ritrovata da Vandepitte et al. (2013) per *Leopoldia gussonei* Parl. di cui sono rimasti solo 3000 esemplari in Sicilia. I risultati ottenuti elaborando i dati derivati dalla digestione *EcoRI-MspI* sono confermati da quelli ottenuti dall'amplificazione dei frammenti ottenuti mediante la digestione con gli enzimi *EcoRI-HpaII*. In questo caso l'eterozigosità attesa aumenta di un ordine di grandezza ma resta comunque molto bassa, con un massimo per la popolazione 15 che raggiunge il valore di 0,2; inoltre anche l'indice di Shannon (dati non mostrati) risulta essere in sostanza pari a zero suggerendo di fatto la che la riproduzione avvenga quasi esclusivamente per via vegetativa, come suggerito anche da Fussi et al. (2012). In particolare, questo studio gli autori hanno evidenziato che 28 pioppi bianchi campionati nell'arcipelago di Malta appartengono ad un unico clone e, inoltre, gli SSR cloroplastici hanno mostrato che questi pioppi sono imparentati con quelli cresciuti in alcune zone italiane. Sia i profili ottenuti dalla digestione con *EcoRI-MspI* che quelli ottenuti con *EcoRI-HpaII* sono stati sottoposti al test statico della varianza molecolare (AMOVA) affiancato dal test di ipotesi per verificare se, le popolazioni indagate, possano essere raggruppate in un'unica popolazione. I risultati sono stati diversi per i profili *EcoRI-MspI* e *EcoRI-HpaII*. La varianza molecolare intra-popolazione dei profili *EcoRI-MspI*, infatti, è risultata circa un terzo di quella ottenuta elaborando i profili *EcoRI-HpaII*. Tuttavia la situazione è invertita se consideriamo la varianza molecolare inter-popolazione. In questo caso, infatti, sono i profili *EcoRI-HpaII* a essere più omogenei rispetto a quelli *EcoRI-MspI*. Questo dato indica che la stabilità epigenetica è maggiore tra le popolazioni campionate rispetto a quella genetica. In entrambi i casi il test di ipotesi associato a quello AMOVA ha permesso di rigettare l'ipotesi nulla di nessuna differenza statistica tra le popolazioni. Per analizzare la struttura genetica dell'intera popolazione di pioppi bianchi e stimare il numero di sottopopolazioni che costituiscono la popolazione nel suo insieme, è stato utilizzato il software *Structure*. L'analisi statistica condotta sui dati molecolari derivanti dalla digestione con *EcoRI-MspI* e *EcoRI-HpaII*, basata su simulazioni e rimescolamenti casuali, ha permesso di calcolare un

numero ottimale di popolazioni (K). I risultati sono differenti: per i profili ottenuti mediante la digestione *EcoRI-MspI* (enzima insensibile alla metilazione) il numero di popolazioni è pari a sei; per i profili ottenuti dalle amplificazione del DNA digerito con *EcoRI-HpaII* il software ha stimato in quattro il numero di popolazioni ottimali. Questo dato risulta molto interessante e conferma la maggiore stabilità molecolare dei profili epigenetici evidenziata dal test AMOVA.

Dai *bar plot* dei profili sia genetici che epigenetici, illustrati nella sezione **Risultati**, si può osservare la stima del grado di rimescolamento genetico dei campioni. In entrambi i casi, i profili risultano essere molto omogenei.

La stima del numero ottimale di popolazioni (K) è stata eseguita anche utilizzando un software diverso implementato in ambiente R. In particolare è stato utilizzato il programma *Geneland* che, oltre a stimare il numero ottimale di popolazioni in un insieme, permette di analizzare i dati genetici in relazione alle coordinate geografiche (*landscape genetics*). I dati molecolari sono stati quindi elaborati per stimare il valore K, che è stato pari a sette per i profili ottenuti dalle amplificazione del DNA digerito con *EcoRI-MspI*, e, invece, pari a quattro per quelli ottenuti con la digestione enzimatica *EcoRI-HpaII*. Trattandosi di due software differenti una piccola differenza nella stima del numero di popolazioni era ipotizzabile. Per valutare queste discrepanze, nel numero K di popolazioni stimate, sono state confrontate le rispettive *membership*. Dal confronto è emerso che la differenza del numero di popolazione stimate per i profili *EcoRI-MspI* è dovuta a due campioni che escono da una popolazione e ne formano una *ex novo*. Anche se i profili genetici generatisi dalla digestione *EcoRI-HpaII* non abbiano evidenziato sostanziali differenze nel numero di popolazioni tra i due software, il confronto delle *membership* ha evidenziato sei campioni che migrano da una popolazione a un'altra, secondo il programma usato per l'analisi.

Questo dato potrebbe essere dovuto a un artefatto statistico, dovuto alle migliaia di simulazioni nel calcolo degli *Fst* utilizzati per l'analisi, e i campionamenti casuali degli *Fst* medesimi, oltre che, ovviamente, a una questione puramente biologica. *Geneland*, quindi, ha confermato che il numero ottimale di popolazioni decresce sensibilmente se si analizzano i profili epigenetici rispetto a quelli genetici. Il risultato è perciò al quanto interessante proprio perché la metilazione sembrerebbe stabilizzare i profili molecolari di piante altrimenti differenti dal punto di vista strettamente genetico. Il risultato è stato molto sorprendente anche perché l'ipotesi iniziale era quella di verificare se, dei pioppi geneticamente identici (cloni), che vivevano in differenti condizioni ambientali, mostrassero un profilo epigenetico alterato; al contrario, dai dati di diversità genetica ottenuti in questo studio è stato verificato che piante simili da un punto di vista genetico, ma non cloni, che vivono nelle stesse condizioni ambientali presentano poi dei profili epigenetici pressoché identici tanto da permettere il loro raggruppamento in popolazioni separate.

Questo dato è avvalorato e confermato dai dendrogrammi di similarità riportati nella sezione **Risultati**. L'analisi condotta ed elaborata mediante la statistica UPGMA, sostanzialmente diversa da quella impiegata sia dal software *Structure* che *Geneland*, conferma la ridotta biodiversità genetica dei pioppi sardi e, soprattutto, conferma che il numero di cluster diminuisce se si prendono in considerazione i profili epigenetici. Allo scopo di stimare come le popolazioni individuate siano connesse, si è proceduto alla costruzione di un albero delle connessioni ottimali. In particolare per ogni popolazione è stato calcolato il profilo modale, mentre le coordinate geografiche delle popolazioni sono state usate per calcolare il centroide del macro gruppo che le comprende. Le distanze genetiche calcolate sui profili modali rispettivamente delle sette e delle quattro popolazioni hanno permesso di disegnare l'albero delle connessioni ottimali che è stato implementato sulla mappa geografica della Sardegna. Grazie a questa elaborazione è stato possibile verificare come le popolazioni, che vivono in aree più vicine, mostrino profili epigenetici tali da permettere il loro raggruppamento in un minor numero di singoli gruppi. I dendrogrammi

inoltre confermano anche il dato dell'AMOVA. Dal grado di similarità rappresentato, infatti, si evince che, per i profili genetici, la similarità è massima per vari campioni, rendendo, quindi, la varianza molecolare intra-popolazione maggiormente ridotta. Il dendrogramma che illustra la similarità degli stessi campioni in base ai profili epigenetici evidenzia un elevato grado di similarità, ma nessun campione raggiunge l'identità, evidenziando in questo modo che, la varianza inter-popolazione è si ridotta, ma a scapito della varianza molecolare intra-popolazione che aumenta di circa i due terzi. Questo risultato può indicare che, sebbene i profili di metilazione tendano a essere simili tra i campioni delle diverse popolazioni, soprattutto in quelle che vivono in condizioni ambientali relativamente simili, essi conservano una porzione di genoma peculiare e distintiva dagli altri cloni.

Per la stima della frazione di frammenti MSAP suscettibili o meno a fenomeni di metilazione è stato utilizzato l'algoritmo elaborato dal Dr. Pérez-Figueroa (2013) e applicato ai dati epigenetici delle popolazioni di pioppo bianco della Sardegna da me campionate.

Questo programma permette di confrontare simultaneamente i profili ottenuti dalla digestione *EcoRI-HpaII* rispetto a quelli ottenuti mediante la digestione *EcoRI-MspI*. I risultati ottenuti dimostrano che il maggior grado di metilazione riguarda le C interne. Inoltre, anche dal confronto statistico eseguito dal software "MSAP" emerge che le popolazioni discriminate in base al loro profilo epigenetico (4) risultano essere tra loro significativamente differenti ($p < 0.001$) in base ai polimorfismi MSAP, mentre il confronto risulta statisticamente non significativo qualora si considerino i profili genetici. Sempre il programma MSAP mi ha permesso di disegnare dei grafici delle componenti principali (PCoA) in base al numero di popolazioni inserito. Nel nostro caso si è calcolato i valori della PCoA sia considerando le sette popolazioni (emerse dalle elaborazioni dei profili genetici), che le quattro (numero di popolazioni stimate in base invece ai profili epigenetici). L'analisi dei risultati ha portato alle medesime conclusioni ottenute mediante le precedenti elaborazioni biostatistiche. Tuttavia, analizzando il problema dal punto di vista biologico mi è sembrato più opportuno riportare le PCoA ottenute inserendo quattro come numero di popolazioni. Questa decisione è stata presa allo scopo di evidenziare la stabilità dei profili epigenetici rispetto a quelli genetici. Osservando le due PCoA riportate nella sezione **Risultati** si può infatti notare come le ellissi, che rappresentano le popolazioni, siano sovrapposte nel caso di profili genetici, e invece chiaramente separate per quelli epigenetici, ad eccezione delle popolazioni identificate nel grafico PCoA come a e b, che presentano, invece, zone di sovrapposizione. Inoltre, analizzando la lunghezza dell'asse maggiore e dell'asse minore delle ellissi, che rappresentano il grado di dispersione dei campioni, è palese come nei grafici ottenuti analizzando i profili epigenetici la dispersione sia minore; questo dato è confermato anche dai valori di C1 e C2, che sono minori nel caso in cui il grafico della PCoA rappresenti i profili epigenetici. Sebbene il livello di metilazione è elevato in molti organismi vegetali, è pur vero che questo dipende dalla specie considerata come dimostrato da Kovařík et al. (1997). In questo studio Kovařík e collaboratori hanno indagato diverse specie di angiosperme usando differenti enzimi di restrizione. Le specie indagate sono caratterizzate da differenti livelli di metilazione del genoma. Inoltre, Ma et al. (2012), studiando la correlazione tra fenotipo e metilazione del pioppo bianco cinese, ha evidenziato come il livello di metilazione, a carico della C interna fosse di poco maggiore dell'emimetilazione; al contrario sempre gli stessi autori in un'altra pubblicazione (Ma et al. 2013) hanno evidenziato come il fenomeno dell'emimetilazione fosse maggiore della metilazione della C interna anche se non in maniera statisticamente significativa. Queste differenze, evidenziate dagli stessi autori in due diverse pubblicazioni, sono probabilmente riconducibili al fatto che la metilazione del genoma può essere tessuto-specifica. Nel lavoro illustrato in questa tesi di dottorato è stato calcolato che il livello di metilazione maggiore nei pioppi bianchi da me campionati è a carico della C interna. Hauben et al. (2009) hanno rilevato che lo stato di metilazione del DNA estratto

dai cotiledoni e dalle foglie della medesima pianta non era lo stesso; anche Teyssier et al. (2008) hanno dimostrato che il livello di metilazione varia durante le fasi di sviluppo del frutto, nelle foglie mature e nel pericarpo. La percentuale di metilazione, inoltre, mostra cambiamenti dinamici nel castagno durante le fasi di crescita e dormienza (Valledor et al. 2004). Per escludere le variabili attribuibili ai tessuti indagati, Ma et al. (2013) hanno analizzato solo lo xilema dei pioppi bianchi cinesi. Questo approccio è stato diverso dal nostro, in quanto obiettivo di questa tesi era di verificare l'effetto ambientale sul profilo epigenetico di popolazioni di pioppo bianco molto simili dal punto di vista genetico. Le foglie, essendo il centro di reazione principale della pianta che sicuramente mostra più sensibilità alla luce, all'altitudine, alla temperatura, agli inquinanti atmosferici, può fornire informazioni di carattere generale riguardante l'adattamento epigenetico all'ambiente. E' stata nostra cura effettuare un campionamento quanto più rappresentativo possibile della pianta, prelevando sei foglie dalla chioma, alla medesima altezza e tutte allo stadio giovanile nell'arco di pochissimi giorni e in condizioni climatiche costanti di bel tempo, in modo da ridurre al minimo l'effetto delle condizioni climatiche, della fenologia fogliare e dello stato di crescita delle piante. E' noto dalla letteratura che alterazioni dello stato di metilazione in risposta a condizioni ambientali differenti sono state riscontrate, per esempio, in piante di mangrovia cresciute nei presso dei fiumi, rispetto a quelle cresciute nelle paludi salate (Lira-Medeiros et al. 2010). I fattori ambientali, infatti, sono ritenuti in grado di influire sulla crescita e la vita della pianta. Per esempio elevate temperature e livelli di salinità, possono causare stress all'organismo vegetale, a cui, reagisce attraverso la modifica dello stato di metilazione delle C, che a sua volta regola l'espressione genica indotta dall'alterazione ambientale (Ahuja et al. 2010). Come dimostrato da Downen et al. (2012), la regolazione dell'espressione genica è chiaramente mediata dalla metilazione del DNA; in quello studio Downen ha osservato che, in *Arabidopsis*, numerose regioni del DNA erano metilate in relazione allo stress dovuto all'infezione di organismi patogeni, con conseguente espressione genica differenziale. E' stato ipotizzato che queste risposte dell'organismo vegetale allo stress biotico o abiotico, mediate da metilazione o demetilazione di regioni del DNA regolatrici l'espressione genica, possano essere addirittura conservate in una sorta di memoria della pianta addirittura trasmissibile alla progenie. Sebbene i precisi meccanismi epigenetici in relazione alla memoria dello stress ambientale non siano ancora del tutto chiariti dalla comunità scientifica (Saze 2008), Thellier and Lüttge (2013) hanno sviluppato una loro ipotesi di risposta allo stress da parte della pianta attraverso episodi di metilazione ipotizzando un meccanismo di memorizzazione del fenomeno stressante a cui l'organismo è andato incontro. Secondo il loro schema, le variazioni epigenetiche possono essere attivate da stimoli esterni come già ampiamente dimostrato in letteratura (Chinnusamy and Zhu 2009; Alvarez et al. 2010; Chen et al. 2010; Yaish et al. 2011). Queste variazioni potrebbero coinvolgere la produzione di segnali chimici come fito-ormoni, segnali elettrici e attivazione dei canali del calcio (Trewavas 2003). Questi segnali possono essere trascritti in particolari RNA regolativi, detti microRNA indotti dallo stress, per esempio in condizioni particolari di salinità, siccità e freddo (Sunkar and Zhu 2004; Shen et al. 2010). Gli studi di epigenetica hanno evidenziato che gli *small interfering RNAs* (siRNAs), aventi una lunghezza di 24 – 26 nucleotidi, sono coinvolti direttamente nella metilazione del DNA e nelle modificazioni istoniche (Zhang et al. 2006; Saze 2008; Chinnusamy and Zhu 2009). I siRNAs, inoltre, sono mobili nelle vie simplastiche e nel floema e possono essere trasferiti da tessuto a tessuto e da organo a organo nella pianta funzionando come segnali sistemici prodotti in seguito a stress (Saze 2008). Affinché si mantenga memoria dello stress attraverso la conservazione dello stato di metilazione, è necessario che, quando lo stress cessa, lo stato di metilazione non sia riportato a livello antecedente lo stress (Chinnusamy and Zhu 2009). Il trasferimento potrebbe avvenire attraverso divisioni cellulari, sia mitotiche che meiotiche; entrambe le ipotesi sembrano essere plausibili (Molinier et al. 2006). Il

trasferimento mediante la mitosi costituisce una memoria dello stress epigenetico di breve durata. Tuttavia, in piante perenni, compresi alberi longevi, la memoria può essere mantenuta per la durata di una o più stagioni. Sasek et al. (1991) riportano che, gli effetti delle piogge acide e l'esposizione all'ozono sulla fotosintesi sono stati osservati per due stagioni su alberi di *Pinus taeda* L. cresciuti in North Caroline. Un altro esempio della trasmissione epigenetica ereditaria ci è fornito da uno studio di Suter and Widmer (2013), i quali hanno verificato che in semenzali di *Arabidopsis* lo stress, provocato da elevata salinità, induce potenziali adattamenti fenotipici ereditabili indipendentemente dalla variazione genetica. Gli stessi autori, infatti, hanno fatto crescere piantine di *A. thaliana* per tre generazioni in due condizioni di stress, causato da eccessiva salinità e calore. Nella quarta generazione, e negli ibridi ottenuti dall'incrocio delle tre generazioni precedenti, è stata verificata la trasmissione dei caratteri fenotipici. In questo modo è stato dimostrato che i cambiamenti fenotipici erano dovuti a trasmissioni materne o paterne. Infine, la generazione cinque e le due seguenti sono state fatte crescere in assenza di stress. I risultati dello studio hanno permesso di evidenziare come lo stress da calore indotto nelle prime tre generazioni abbia accelerato la fioritura nella generazione successiva. Tuttavia i caratteri epigenetici ereditati sono scomparsi nelle successive generazioni cresciute in assenza di stress. Le esposizioni delle prime generazioni a stress salinità elevata ha incrementato la tolleranza allo stesso in uno dei due ibridi. In assenza di variabilità genetica, l'ereditarietà materna e paterna, e la reversibilità degli effetti transgenerazionali indicano che lo stress può indurre cambiamenti adattativi di tipo fenotipico probabilmente attraverso meccanismi epigenetici (Suter and Widmer 2013)

E' stato recentemente pubblicato sulla rivista Nature (Latzel et al. 2013) che l'epigenetica, oltre a essere implicata nei meccanismi di memoria dello stress, è fondamentale anche nella valutazione della biodiversità. Gli autori sostengono che la diversità epigenetica incrementa la produttività e la stabilità delle popolazioni vegetali. In particolare, nel loro studio, Latzel et al. (2013) hanno stimato che, popolazioni di *A. thaliana* epigeneticamente diverse, producono fino al 40 % di biomassa in più rispetto a popolazioni uniformi dal punto di vista epigenetico. Questo dato risulta molto più evidente qualora le popolazioni siano cresciute insieme a organismi competitori o se infettate da patogeni. Gli autori dello studio suggeriscono, in conclusione, che sarebbe opportuno includere le ricerche di tipo epigenetico nello studio dei sistemi ecologici quantificando la naturale diversità epigenetica e testando le sue conseguenze tra numerose e differenti specie. Nel nostro caso la diversità epigenetica tra le popolazioni di pioppo bianco della Sardegna diminuisce permettendo di raggruppare le popolazioni campionate in un numero minore di cluster rispetto a quello stimato analizzando la diversità genetica. L'ambiente, evidentemente, altera lo stato di metilazione delle foglie di pioppo sardo indagate, rendendo i profili epigenetici delle piante che vivono in condizioni ambientali simili, più omogenei sia per piante geneticamente uguali, ma, fatto assai rilevante, anche dei pioppi maggiormente biodiversi dal punto di vista genetico. Naturalmente sono molte le variabili ambientali che potrebbero influire sull'omogeneità dei profili epigenetici; per esempio Labra et al. (2002) hanno dimostrato l'influenza dello stress idrico sul profilo di metilazione delle radici di pisello, mentre Aina et al. (2004), Cicatelli et al. (2014) e (2012) hanno evidenziato che l'inquinamento da metalli pesanti altera lo stato di metilazione ma anche l'espressione genica in *Trifolium repens* L., *Cannabis sativa* L. e *P. alba*.

Nel lavoro presentato e discusso in questa tesi di dottorato è stato evidenziato, attraverso differenti strumenti bio-statistici, che la posizione geografica influenza i profili di metilazione dei pioppi bianchi della Sardegna; questo risultato, tenendo conto dell'organo analizzato, potrebbe essere dovuto a molti fattori, principalmente pedo-climatici come l'intensità luminosa, le ore di luce, la temperatura, l'umidità, l'altitudine, etc.. Gli sviluppi futuri di questo studio condotto durante i miei tre anni di dottorato prevedono la pubblicazione su una rivista internazionale dei dati e successivamente il sequenziamento

dei frammenti MSAP amplificati mediante un approccio NGS (*Next Generation Sequencing*). In tal modo sarà possibile identificare porzioni metilate del genoma ed eventualmente collegarle all'espressione di geni sensibili alle variazioni ambientali. In generale il lavoro fin qui eseguito permette di concludere che la biodiversità genetica dei pioppi bianchi sardi è al quanto limitata, e lo è ancora di più quella epigenetica se si considera la varianza molecolare inter popolazione. Tuttavia, è opportuno notare che la biodiversità epigenetica intra-popolazione aumenta rispetto a quella genetica, confermando che gli studi di biodiversità non possono più limitarsi alla sola varianza genetica (Latzel et al. 2013), in quanto l'adattamento più rapido e frequente consiste nel cambiamento dello stato di metilazione e non nelle mutazioni genetiche che, in special modo nelle popolazioni di tipi clonale, avvengono esclusivamente in maniera casuale e sono trasmesse secondo la genetica mendeliana alle generazioni successive e poi selezionate nel lungo periodo (Ossowski et al. 2010).

Ringraziamenti

Gli ultimi tre anni di lavoro e di vita sono il frutto di una scelta non facile che tuttavia ho fatto con serenità e decisione. Per scegliere è necessario che ci siano delle alternative o che queste siano create con il lavoro, la passione e la dedizione. Tutto questo a volte può non bastare: siamo persone e in quanto tali instauriamo quotidianamente relazioni e rapporti con altre persone. Niente si può raggiungere da soli. E' per questo motivo che i miei ringraziamenti sono rivolti al Prof. Castiglione (il Prof.) e alla Dott.ssa Angela Cicutelli con cui già da tanti anni si è instaurata una vera e sincera amicizia. Il Prof e Angela mi hanno trasmesso e continuano a trasmettermi lo spirito, la mentalità e l'entusiasmo di chi crede che tutti i traguardi siano raggiungibili con l'impegno e con lo studio.

Il lavoro che è stato illustrato in questa tesi è stato reso possibile grazie al contributo fondamentale del dott. Giuseppe Brundu; grazie alla sua passione per la scienza e alla sua enorme disponibilità abbiamo potuto realizzare i campionamenti delle popolazioni del pioppo bianco della Sardegna.

Al di là delle conoscenze e del metodo di ricerca, al Prof Castiglione devo parte della mia crescita personale di questi tre anni. I traguardi che ho raggiunto sono il risultato di una crescita e di uno stile di vita sereno e armonioso che i miei genitori si sono impegnati a creare e mantenere: ciò che sono lo devo soprattutto a loro. Un grazie altrettanto speciale lo devo alla mia ragazza Barbara che con pazienza, entusiasmo e ottimismo mi ha sempre appoggiato, incoraggiato e spronato nel mio e nel nostro interesse.

Bibliografia

- Adeyemo, O., A. Menkir, et al. (2011). "Genetic diversity assessment and relationship among tropical-yellow endosperm maize inbred lines using SSR markers." *Maydica* **56**(1): 43-49.
- Ahuja, I., R. C. H. de Vos, et al. (2010). "Plant molecular stress responses face climate change." *Trends in Plant Science* **15**(12): 664-674.
- Aina, R., S. Sgorbati, et al. (2004). "Specific hypomethylation of DNA is induced by heavy metals in white clover and industrial hemp." *Physiologia Plantarum* **121**(3): 472-480.
- Alvarez, M. E., F. Nota, et al. (2010). "Epigenetic control of plant immunity." *Molecular Plant Pathology* **11**(4): 563-576.
- Armenise L. , S. M. C., Landi S. (2010). "La banca centrale del DNA forestale." *Silvae*(11).
- Asrat, S., M. Yesuf, et al. (2010). "Farmers' preferences for crop variety traits: Lessons for on-farm conservation and technology adoption." *Ecological Economics* **69**(12): 2394-2401.
- Baldwin, S., M. Pither-Joyce, et al. (2012). "Development of robust genomic simple sequence repeat markers for estimation of genetic diversity within and among bulb onion (*Allium cepa* L.) populations." *Molecular Breeding* **30**(3): 1401-1411.
- Baranwal, V. K., V. Mikkilineni, et al. (2012). "Heterosis: emerging ideas about hybrid vigour." *Journal of Experimental Botany* **63**(18): 6309-6314.
- Barzegar, R., G. Peyvast, et al. (2013). "Biochemical systematic, population structure and genetic variability studies among Iranian *Cucurbita* (*Cucurbita pepo* L.) accessions, using genomic SSRs and implications for their breeding potential." *Biochemical Systematics and Ecology* **50**: 187-198.
- Barzegar, R., G. Peyvast, et al. (2013). "Biochemical systematic, population structure and genetic variability studies among Iranian *Cucurbita* (*Cucurbita pepo* L.) accessions, using genomic SSRs and implications for their breeding potential." *Biochemical Systematics and Ecology* **50**: 187-198.
- Block, E. (2009). *Garlic and Other Alliums. The Lore and the Science*. Cambridge Royal Society of Chemistry.
- Botstein, D., R. L. White, et al. (1980). "Construction of a genetic linkage map in man using restriction fragment length polymorphisms." *American journal of human genetics* **32**: 314-331.
- Boyko, A. and I. Kovalchuk (2008). "Epigenetic control of plant stress response." *Environmental and Molecular Mutagenesis* **49**(1): 61-72.

- Bremer, B., K. Bremer, et al. (2009). "An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III." *Botanical Journal of the Linnean Society*.
- Brundu, G., R. Lupi, et al. (2008). "The origin of clonal diversity and structure of *populus alba* in sardinia: evidence from nuclear and plastid microsatellite markers." *Annals of Botany* **102**(6): 997-1006.
- Brush, S. B. (1995). "In situ conservation of landraces in centers of crop diversity." *Crop Science* **35**(2): 346-354.
- Buchanan, B. B., W. Gruissem, et al. (2003). *Biochimica e biologia molecolare delle piante*, Zanichelli.
- Bush, G. L. (1975). "Modes of animal speciation." *Annual Review of Ecology and Systematics* **6**: 339-364.
- Cadalen, T., M. Mörchen, et al. (2010). "Development of SSR markers and construction of a consensus genetic map for chicory (*Cichorium intybus* L.)." *Molecular Breeding* **25**: 699-722.
- Camarda, V. (1985). *Alberi e arbusti spontanei della Sardegna*. Gallizzi.
- Chen, M., S. Lv, et al. (2010). "Epigenetic performers in plants." *Development Growth & Differentiation* **52**(6): 555-566.
- Chinnusamy, V. and J.-K. Zhu (2009). "Epigenetic regulation of stress responses in plants." *Current Opinion in Plant Biology* **12**(2): 133-139.
- Chitwood, D. H. and M. C. P. Timmermans (2010). "Small RNAs are on the move." *Nature* **467**(7314): 415-419.
- Chrispeels M. and Sadava D. (1996). *Biologia vegetale applicata*. Piccin.
- Cicatelli, A., G. Lingua, et al. (2012). "Arbuscular mycorrhizal fungi modulate the leaf transcriptome of a *Populus alba* L. clone grown on a zinc and copper-contaminated soil." *Environmental and Experimental Botany* **75**: 25-35.
- Cicatelli, A., V. Todeschini, et al. (2014). "Epigenetic control of heavy metal stress response in mycorrhizal versus non-mycorrhizal poplar plants." *Environ Sci Pollut Res Int* **21**(3): 1723-1737.
- Cohen, J. E. and D. Tilman (1996). "Ecology - Biosphere 2 and biodiversity: The lessons so far." *Science* **274**(5290): 1150-1151.
- Cottrell, J. E., V. Krystufek, et al. (2005). "Postglacial migration of *Populus nigra* L.: lessons learnt from chloroplast DNA." *Forest Ecology and Management* **219**(2): 293-312.
- Darwin, C. (1872). *The Origin of Species by means of Natural Selection, or the preservation of favoured Races in the struggle for Life*. Sixth edition, with additions and corrections.

- Di Falco, S., I. Penov, et al. (2010). "Agrobiodiversity, farm profits and land fragmentation: Evidence from Bulgaria." *Land Use Policy* **27**(3): 763-771.
- Di Falco, S. and C. Perrings (2005). "Crop biodiversity, risk management and the implications of agricultural assistance." *Ecological Economics* **55**(4): 459-466.
- Diaconis, P. (2009). "The markov chain monte carlo revolution." *Bulletin of the American Mathematical Society* **46**(2): 179-205.
- Downen, R. H., M. Pelizzola, et al. (2012). "From the Cover: PNAS Plus: Widespread dynamic DNA methylation in response to biotic stress." *Proceedings of the National Academy of Sciences* **109**(32): E2183-E2191.
- Doyle, J. J. and J. L. Doyle (1990). "Isolation of plant DNA from fresh tissue. ." *Focus* **12**: 13-15.
- Engelmann, F. (1979). "Insect vitellogenin: identification, biosynthesis and role in vitellogenesis." *Adv. Insect Physiol* **14**: 49-108.
- Evanno, G., S. Regnaut, et al. (2005). "Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study." *Molecular ecology* **14**(8): 2611-2620.
- Excoffier, L., P. E. Smouse, et al. (1992). "Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data." *Genetics* **131**(2): 479-491.
- Felsenstein, J. (2005). "PHYLIP (Phylogeny Inference Package) version 3.6." Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.
- Figliuolo, G., V. Candido, et al. (2001). "Genetic evaluation of cultivated garlic germplasm (*Allium sativum* L. and *A-ampeloprasum* L.)." *Euphytica* **121**(3): 325-334.
- Fineschi, S., M. Anzidei, et al. (2002). "Molecular markers reveal a strong genetic differentiation between two European relic tree species: *Zelkova abelicea* (Lam.) Boissier and *Z-sicula* Di Pasquale, Garfi & Quezel (Ulmaceae)." *Conservation Genetics* **3**(2): 145-153.
- Fujimoto, R., T. Sasaki, et al. (2012). "Molecular Mechanisms of Epigenetic Variation in Plants." *International Journal of Molecular Sciences* **13**(12): 9900-9922.
- Fussi, B., J. Bonello, et al. (2012). "Combining the use of molecular techniques and archival documentary evidence to trace the origin of *Populus alba* in a Central Mediterranean archipelago." *European Journal of Forest Research* **131**(2): 347-354.
- Gower, J. C. and G. J. S. Ross (1969). "Minimum spanning trees and single linkage cluster analysis." *Applied statistics*: 54-64.
- Grassi, F., M. Labra, et al. (2006). *Introduzione alla biodiversità del mondo vegetale*. Piccin.

- Grativol, C., A. S. Hemerly, et al. (2012). "Genetic and epigenetic regulation of stress responses in natural plant populations." *Biochimica Et Biophysica Acta-Genetic Regulatory Mechanisms* **1819**(2): 176-185.
- Guillot, G., F. Santos, et al. (2011). Population genetics analysis using R and the Geneland program.
- Hauben, M., B. Haesendonckx, et al. (2009). "Energy use efficiency is characterized by an epigenetic component that can be directed through artificial selection to increase yield." *Proceedings of the National Academy of Sciences of the United States of America* **106**(47): 20109-20114.
- Hoag, J. C. and T. A. Landis (2001). "Plant materials for riparian revegetation." *Native Plant Journal* **2**(1).
- Hong, C. J. and T. Etoh (1996). "Fertile clones of garlic (*Allium sativum* L.) abundant around the Tien Shan mountains." *Breeding Science* **46**(4): 349-353.
- Howden, S. M., J.-F. Soussana, et al. (2007). "Adapting agriculture to climate change." *Proceedings of the National Academy of Sciences* **104**(50): 19691-19696.
- Ipek, M., A. Ipek, et al. (2008). "Molecular characterization of Kastamonu garlic: An economically important garlic clone in Turkey." *Scientia Horticulturae* **115**: 203-208.
- Jaccard, P. (1901). "Étude comparative de la distribution florale dans une portion des Alpes et des Jura." *Bulletin del la Société Vaudoise des Sciences Naturelles* **37**: 547-579.
- Jost, L. O. U. (2008). "GST and its relatives do not measure differentiation." *Molecular Ecology* **17**(18): 4015-4026.
- Kalia, R. K., M. K. Rai, et al. (2010). "Microsatellite markers: an overview of the recent progress in plants." *Euphytica* **177**(3): 309-334.
- Karan, R., T. DeLeon, et al. (2012). "Salt Stress Induced Variation in DNA Methylation Pattern and Its Influence on Gene Expression in Contrasting Rice Genotypes." *Plos One* **7**(6).
- Keller, E. R. J. and A. Senula (2013). "Micropropagation and cryopreservation of garlic (*Allium sativum* L.)." *Methods in molecular biology* (Clifton, N.J.) **11013**: 353-368.
- Kimura, M. (1983). "Rare variant alleles in the light of the neutral theory." *Molecular biology and evolution* **1**(1): 84-93.
- Kingsley R. Stern, James E. Bidlack, et al. (2009). *Introduzione alla biologia vegetale*. The McGraw-Hill Companies.
- Kovařík, A., R. Matyášek, et al. (1997). "Variability in CpNpG methylation in higher plant genomes." *Gene* **204**(1-2): 25-33.

- Labra, M., A. Ghiani, et al. (2002). "Analysis of cytosine methylation pattern in response to water deficit in pea root tips." *Plant Biology* 4(6): 694-699.
- Landis, T., D. Dreesen, et al. (2003). "Sex and the Single Salix: Considerations for Riparian Restoration." *Native Plants Journal* 4(2): 110-117.
- Latzel, V., E. Allan, et al. (2013). "Epigenetic diversity increases the productivity and stability of plant populations." *Nature Communications* 4: 7.
- Lira-Medeiros, C. F., C. Parisod, et al. (2010). "Epigenetic Variation in Mangrove Plants Occurring in Contrasting Natural Environment." *Plos One* 5(4).
- Liu, K. and S. V. Muse (2005). "PowerMarker: an integrated analysis environment for genetic marker analysis." *Bioinformatics* 21(9): 2128-2129.
- Lukens, L. N. and S. Zhan (2007). "The plant genome's methylation status and response to stress: implications for plant improvement." *Current Opinion in Plant Biology* 10(3): 317-322.
- Ma, K., Y. Song, et al. (2012). "Photosynthetic response to genome methylation affects the growth of Chinese white poplar." *Tree Genetics & Genomes* 8(6): 1407-1421.
- Ma, K., Y. Song, et al. (2013). "Variation in Genomic Methylation in Natural Populations of Chinese White Poplar." *Plos One* 8(5).
- Mayr, E. (1963). "Names given to hybrids." *Bulletin of Zoological Nomenclature* 20: 50-51.
- Mazzio, E. A. and K. F. A. Soliman (2012). "Basic concepts of epigenetics Impact of environmental signals on gene expression." *Epigenetics* 7(2): 119-130.
- McClintock, B. (1984). "The significance of responses of the genome to challenge." *Science (New York, N.Y.)* 226(4676): 792-801.
- McCollum, G. D. (1976). "Onion and allies." *Evolution of Crop Plants*: 186-190.
- Medina, R. F., P. Barbosa, et al. (2006). "Number of individuals and molecular markers to use in genetic differentiation studies." *Molecular Ecology Notes* 6(4): 1010-1013.
- Meer, Q. P. and J. L. v. Bennekom (1972). "Influence of the environment on the percentage of self-fertilisation in onions and some consequences for breeding." *Euphytica* 21(3): 450-453.
- Meirmans, P. G. and P. W. Hedrick (2011). "Assessing population structure: FST and related measures." *Molecular Ecology Resources* 11(1): 5-18.
- Michalakis, Y. and L. Excoffier (1996). "A Generic Estimation of Population Subdivision Using Distances Between Alleles With Special Reference for Microsatellite Loci." *Genetics* 142(3): 1061-1064.

- Molin, D., C. J. Coelho, et al. (2013). "Genetic diversity in the germplasm of tropical maize landraces determined using molecular markers." *Genetics and molecular research: GMR* **12**(1): 99.
- Molinier, J., G. Ries, et al. (2006). "Transgeneration memory of stress in plants." *Nature* **442**(7106): 1046-1049.
- Mullis, K. B., F. A. Faloona, et al. (1992). "Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction." *Biotechnology series*: 17-17.
- Nei, M. and R. K. Chesser (1983). "Estimation of fixation indices and gene diversities." *Annals of human genetics* **47**(3): 253-259.
- Ossowski, S., K. Schneeberger, et al. (2010). "The Rate and Molecular Spectrum of Spontaneous Mutations in *Arabidopsis thaliana*." *Science* **327**(5961): 92-94.
- Panickar, K. S. (2013). "Bioactive Components of Plant Products Including Polyphenols Exert Neuroprotective Effects and Benefit Neural Function." *Central Nervous System Agents in Medicinal Chemistry* **13**(1): 2-2.
- Paris, H. S., N. Yonash, et al. (2003). "Assessment of genetic relationships in *Cucurbita pepo* (Cucurbitaceae) using DNA markers." *Theoretical and Applied Genetics* **106**(6): 971-978.
- Peakall, R., P. E. Smouse, et al. (1995). "Evolutionary implications of allozyme and RAPD variation in diploid populations of dioecious buffalograss *Buchloe dactyloides*." *Molecular Ecology* **4**(2): 135-148.
- Peakall, R. O. D. and P. E. Smouse (2006). "genalex 6: genetic analysis in Excel. Population genetic software for teaching and research." *Molecular Ecology Notes* **6**(1): 288-295.
- Pejic, I., P. Ajmone-Marsan, et al. (1998). "Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs." *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik* **97**(8): 1248-1255.
- Pérez-Figueroa, A. (2013). "msap: a tool for the statistical analysis of methylation-sensitive amplified polymorphism data." *Molecular ecology resources*.
- Petit, R. J., I. Aguinagalde, et al. (2003). "Glacial refugia: hotspots but not melting pots of genetic diversity." *Science* **300**(5625): 1563-1565.
- Petit, R. J., S. Brewer, et al. (2002). "Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence." *Forest ecology and management* **156**(1): 49-74.
- Pritchard, J. K., M. Stephens, et al. (2000). "Inference of population structure using multilocus genotype data." *Genetics* **155**(2): 945-959.

- Putiri, E. and K. Robertson (2011). "Epigenetic mechanisms and genome stability." *Clinical Epigenetics* 2(2): 299-314.
- Ranatunga, M. A. B., P. Meenakshisundaram, et al. (2009). "Genetic diversity analysis of maize (*Zea mays* L.) inbreds determined with morphometric traits and simple sequence repeat markers." 54: 113-123.
- Reyna-Lopez, G. E., J. Simpson, et al. (1997). "Differences in DNA methylation patterns are detectable during the dimorphic transition of fungi by amplification of restriction polymorphisms." *Molecular and General Genetics MGG* 253(6): 703-710.
- Ricciardi L. and Filippetti A. (2003). L'erosione genetica di specie agrarie in ambito mediterraneo: rilevanza del problema e strategie d'intervento. Cahiers Options Méditerranéennes. Ciheam, Bari. 53.
- Robinson, R. W. and D. S. Decker-Walters (1997). Cucurbits. Cab international.
- Roiron, P., A. A. Ali, et al. (2004). "Preuve de l'indigénat de *Populus alba* L. dans le Bassin méditerranéen occidental." *Comptes Rendus Biologies* 327(2): 125-132.
- Saitou, N. and M. Nei (1987). "The neighbor-joining method: a new method for reconstructing phylogenetic trees." *Molecular biology and evolution* 4(4): 406-425.
- Sasek, T. W., C. J. Richardson, et al. (1991). "Carryover effects of acid-rain and ozone on the physiology of multiple flushes of loblolly-pine seedlings." *Forest Science* 37(4): 1078-1098.
- Saze, H. (2008). "Epigenetic memory transmission through mitosis and meiosis in plants." *Seminars in Cell & Developmental Biology* 19(6): 527-536.
- Shannon, C. E. (1948). "A mathematical theory of communication." *Bell system technical journal* 27.
- Shen, J., K. Xie, et al. (2010). "Global expression profiling of rice microRNAs by one-tube stem-loop reverse transcription quantitative PCR revealed important roles of microRNAs in abiotic stress responses." *Molecular Genetics and Genomics* 284(6): 477-488.
- Simpson, E. H. (1949). "Measurement of diversity." *Nature*.
- Sunkar, R. and J. K. Zhu (2004). "Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*." *Plant Cell* 16(8): 2001-2019.
- Suter, L. and A. Widmer (2013). "Environmental Heat and Salt Stress Induce Transgenerational Phenotypic Changes in *Arabidopsis thaliana*." *Plos One* 8(4).
- Taramino, G. and S. Tingey (1996). "Simple sequence repeats for germplasm analysis and mapping in maize." *Genome* 39(2): 277-287.
- Tautz, D. and M. Renz (1984). "Simple sequences are ubiquitous repetitive components of eukaryotic genomes." *Nucleic acids research* 12(10): 4127-4138.

- Teyssier, E., G. Bernacchia, et al. (2008). "Tissue dependent variations of DNA methylation and endoreduplication levels during tomato fruit development and ripening." *Planta* **228**(3): 391-399.
- Thellier, M. and U. Lüttge (2013). "Plant memory: a tentative model." *Plant Biology* **15**(1): 1-12.
- Trewavas, A. (2003). "Aspects of plant intelligence." *Annals of Botany* **92**(1): 1-20.
- Vaillant, I. and J. Paszkowski (2007). "Role of histone and DNA methylation in gene regulation." *Current Opinion in Plant Biology* **10**(5): 528-533.
- Valdes, A. M., M. Slatkin, et al. (1993). "Allele frequencies at microsatellite loci - the stepwise mutation model revisited." *Genetics* **133**(3): 737-749.
- Valledor, L., M. Berdasco, et al. (2004). In Vitro Proliferation and Genome DNA Methylation in Adult Chestnuts. III International Chestnut Congress 693.
- Vandepitte, K., A. S. Gristina, et al. (2013). "Conservation genetics of an endemic from the Mediterranean Basin: high genetic differentiation but no genetic diversity loss from the last populations of the Sicilian Grape Hyacinth *Leopoldia gussonei*." *Conservation Genetics* **14**(5): 963-972.
- Vavilov, N. (2009). "Geographical regularities in the distribution of the genes of cultivated plants." *Comparative Cytogenetics* **3**(1): 71-78.
- Vigouroux, Y., J. C. Glaubitz, et al. (2008). "Population structure and genetic diversity of new world maize races assessed by dna microsatellites." *American Journal of Botany* **95**(10): 1240-1253.
- Vos, P., R. Hogers, et al. (1995). "AFLP - A new technique for dna-fingerprinting." *Nucleic Acids Research* **23**(21): 4407-4414.
- Wang, B., Y. Li, et al. (2012). "Cadmium and Its Epigenetic Effects." *Current Medicinal Chemistry* **19**(16): 2611-2620.
- Whitlock, M. C. (2011). "G'ST and D do not replace FST." *Mol Ecol* **20**(6): 1083-1091.
- Williams, J. G. K., A. R. Kubelik, et al. (1990). "DNA polymorphisms amplified by arbitrary primers are useful as genetic markers." *Nucleic acids research* **18**(22): 6531-6535.
- Wright, S. (1946). "Isolation by distance under diverse systems of mating." *Genetics* **31**(1): 39.
- Yaish, M. W., J. Colasanti, et al. (2011). "The role of epigenetic processes in controlling flowering time in plants exposed to stress." *Journal of Experimental Botany* **62**(11): 3727-3735.
- Yu, K., S. J. Park, et al. (1999). "Abundance and variation of microsatellite DNA sequences in beans (*Phaseolus* and *Vigna*)." *Genome* **42**(1): 27-34.

Zhang, M. S., J. N. Kimatu, et al. (2010). "DNA cytosine methylation in plant development." *Journal of Genetics and Genomics* **37**(1): 1-12.

Zhang, X., J. Yazaki, et al. (2006). "Genome-wide high-resolution mapping and functional analysis of DNA methylation in Arabidopsis." *Cell* **126**(6): 1189-1201.

Zhao, W.-G., J.-W. Chung, et al. (2011). "Molecular genetic diversity and population structure of a selected core set in garlic and its relatives using novel SSR markers." *Plant Breeding* **130**: 46-54.

Zhao, Y. and D. X. Zhou (2012). "Epigenomic Modification and Epigenetic Regulation in Rice." *Journal of Genetics and Genomics* **39**(7): 307-315.