

UNIVERSITÀ DEGLI STUDI DI CATANIA

Dottorato di ricerca in:
PRODUTTIVITÀ DELLE PIANTE COLTIVATE IN AMBIENTE MEDITERRANEO
(XXVII ciclo: 2011-2014)

GIUSEPPE DILIBERTO

Impiego di tecniche innovative per la valorizzazione
e caratterizzazione molecolare di specie vegetali
ornamentali

Tesi di dottorato

Tutor: *Dott. Antonio Giovino*
Dott. Giovanni Gugliuzza
Coordinatore: *Prof. Daniela Romano*

INDICE

PRESENTAZIONE	pag.	3
PARTE INTRODUTTIVA	“	4
1. Biodiversità e Classificazione	“	4
2. Certificazione genetica: tipizzazione e tracciabilità	“	7
3. Identificazione delle specie vegetali	“	13
3.1. Identificazione fenotipica	“	13
3.2. Identificazione genetica	“	14
3.2.1. Il DNA	“	16
3.2.2. Marcatori molecolari	“	19
3.2.3. DNA Barcoding	“	24
3.2.3.1. Utilità ed applicazioni del DNA barcoding	“	27
3.2.3.2. Caratteristiche di un marcatore barcoding e approccio multilocus	“	29
3.2.3.3. CBOL	“	33
3.2.3.4. BOLD data system	“	36
3.2.4. PCR	“	38
3.2.5. Sequenziamento	“	40
3.2.6. Filogenesi e cluster analysis	“	42
PARTE SPERIMENTALE	“	47
4. Materiali e metodi	“	47
4.1. Selezione del materiale vegetale	“	48
4.2. Conservazione del materiale vegetale	“	54
4.3. Estrazione del DNA da materiale vegetale	“	55

4.4. Valutazione quantitativa e qualitative del DNA estratto	“	58
4.5. Amplificazione PCR	“	60
4.6. Elettroforesi	“	63
4.7. Sequenziamento ed analisi dei dati	“	65
4.8. Valutazione delle performance dei marcatori barcoding	“	66
5. RISULTATI E DISCUSSIONI	“	68
5.1. Conservazione del materiale vegetale	“	68
5.2. Estrazione del DNA da materiale vegetale	“	69
5.3. Valutazione delle performance tecniche dei marcatori barcoding	“	70
5.4. Identificazione tramite DNA barcoding	“	72
5.5. Problematiche riscontrate e soluzioni	“	94
5.6. DNA barcoding nella tracciabilità delle specie vegetali ornamentali	“	95
6. CONCLUSIONI	“	97
7. BIBLIOGRAFIA	“	98
8. APPENDICE	“	106

PRESENTAZIONE

Il presente elaborato è relativo alle attività di studio e di ricerca svolte nel corso del triennio 2011-2014, nell'ambito della frequenza al corso di dottorato di ricerca internazionale in "produttività delle piante coltivate in ambiente mediterraneo", ricadente nell'ambito dell'area formazione del progetto SO.PRO.ME-PON01_01611.

L'argomento affrontato, a suo tempo sottoposto all'approvazione del consiglio dei docenti, riguarda *"l'impiego di tecniche innovative per la valorizzazione e caratterizzazione molecolare di specie vegetali ornamentali"*.

In particolare la tesi affronta i problemi connessi all'identificazione a livello genotipico del materiale florovivaistico con metodi di analisi molecolare standardizzati, a basso costo ed eseguibili in tempi alquanto rapidi, al fine di approfondire problematiche attinenti:

- 1) Controversie tassonomiche (sinonimie/omonimie) di specie affini.
- 2) Identificazione precoce di specie con cicli fenologici molto lunghi.
- 3) Controlli di corrispondenza di specie con rilevante impatto commerciale.

Questo lavoro può rappresentare la base per:

- sviluppare un sistema di certificazione genetica delle piante .
 - sviluppare un sistema di tracciabilità molecolare delle piante da proporre alle aziende.
- Con l'obiettivo di rispondere alle nuove e crescenti esigenze di mercato, improntate all'innovazione di prodotto ed alla sua tipizzazione, nonché a supportare gli aspetti commerciali legati all'identità e tracciabilità delle piante.

Tale sistema rappresenterebbe, sia un utile strumento per la valorizzazione della produzione vivaistica, accrescendo in tal modo la competitività delle aziende vivaistiche sui mercati, sia uno strumento di tutela per i vari protagonisti del settore, dai produttori e propagatori di piante ornamentali fino ai consumatori finali.

PARTE INTRODUTTIVA

1. Biodiversità e Classificazione

La diversità biologica, meglio conosciuta come biodiversità, è sinonimo di ricchezza, di specie, della coesistenza di svariate forme di vita, non etichettate come migliori o peggiori, normali o anormali, belle o brutte, ma tutte utili e selezionate nel corso dei millenni.

La prima definizione di biodiversità fu coniata durante la Conferenza delle Nazioni Unite sull'ambiente e sullo sviluppo, tenutasi a Rio de Janeiro nel 1992: "Per diversità biologica si intende la variabilità degli organismi viventi, degli ecosistemi terrestri, acquatici e i complessi ecologici che essi costituiscono; la diversità biologica comprende la diversità intraspecifica, interspecifica e degli ecosistemi" .

Questa ricchezza è il frutto dei lenti processi evolutivi che, sotto la spinta della selezione naturale, agiscono sulle caratteristiche genetiche e morfologiche delle specie, permettendo così alle forme di vita di adattarsi al cambiamento delle condizioni ambientali.

La biodiversità è fondamentale non solo per noi, ma anche per i nostri discendenti e per tutti gli esseri viventi della Terra, è il pilastro della salute del nostro pianeta. Dalla varietà di forme di vita animali e vegetali, infatti, dipendono sia la qualità dell'esistenza umana sia la nostra stessa possibilità di sopravvivenza. Se la varietà della vita è più ampia, infatti, ogni ecosistema reagisce meglio agli stimoli negativi, siano essi rappresentati dai cambiamenti climatici, dai dissesti idrogeologici, dall'introduzione di Organismi Geneticamente Modificati o dall'invasione di specie aliene. Ma quali sono le cause che determinano la perdita di biodiversità? Inquinamento, specie aliene, la perdita e la frammentazione degli habitat ed i cambiamenti climatici sono tra le cause principali, in quanto non solo possono alterare in modo irreversibile i delicati equilibri del nostro ecosistema, ma possono anche amplificare gli effetti di questo processo. Tutelare la biodiversità vuol dire anche contrastare il cambiamento climatico, una delle più grandi sfide del 21° secolo.

(<http://www.legambiente.it/temi/natura/biodiversità>)

Pertanto la conservazione e l'utilizzo sostenibile della diversità biologica sono quindi due elementi indispensabili per conseguire uno sviluppo sostenibile e per realizzare gli obiettivi di sviluppo in materia di povertà, salute e ambiente (i cosiddetti Millennium Goals).

Le informazioni e le conoscenze relative alla diversità biologica sono in genere insufficienti. È quindi necessario sviluppare capacità scientifiche, tecniche ed istituzionali per ottenere le conoscenze basilari grazie alle quali programmare ed attuare opportuni provvedimenti per conservare la diversità biologica.

La basilare importanza della tassonomia, la scienza che si occupa di descrivere le affinità evolutive, classificare e dare un nome agli esseri viventi, nel progresso delle scienze biologiche è sotto gli occhi di tutti, "Nomina si nescis, perit et cognitio rerum" (se si ignora il nome delle cose, scompare anche quello che di esse si conosce) affermava il padre della nomenclatura zoologica Carlo Linneo.

La tassonomia ci consente di distinguere gli animali e le piante tra loro, di intenderci in qualsiasi lingua (grazie alla nomenclatura binomia in latino), di monitorare e conservare la biodiversità (altrimenti impossibile senza un nome per ciascuna specie), di studiare l'evoluzione delle specie e la loro ecologia ed etologia, di riconoscere le specie utili per lo sfruttamento antropico e molto altro.

L'identificazione della specie è essenziale per il monitoraggio e la conservazione della biodiversità su ampia scala, e la misura della ricchezza delle specie è un utile indicatore di tale biodiversità.

Inizialmente, la maggior parte delle specie sono state differenziate grazie ai caratteri morfologici dell'adulto, ma nel tempo si sono sviluppati approcci molto più sofisticati per questo scopo. Il microscopio elettronico, l'analisi di caratteristiche comportamentali e marker biochimici sono divenuti tutti strumenti che i tassonomi hanno acquisito ed utilizzato per migliorare la scienza della tassonomia.

Il primo sistema di catalogazione di specie è stato fondato più di 250 anni fa dal naturalista svedese Carl von Linné (1707-1778) che ha iniziato la tassonomia formale attraverso l'introduzione della nomenclatura binomiale della specie, principalmente basata sui caratteri morfologici per descrivere la biodiversità (Linneo, 1756).

Il suo lavoro pionieristico ha rappresentato una pietra miliare come sistema di classificazione della specie, anche se egli ha in effetti sottostimato la reale diversità biologica sulla Terra.

Ancora oggi siamo molto lontani da una classificazione completa della biodiversità esistente sul nostro pianeta.

Nel corso di oltre due secoli i tassonomi e naturalisti sono stati in grado, utilizzando osservazioni morfologiche e più recentemente marcatori biochimici, di descrivere e classificare solo una piccola parte delle specie potenzialmente esistenti sul pianeta terra.

Lo straordinario catalogo della vita sul nostro pianeta viene mantenuto da un team di scienziati organizzati nel Catalogue of Life (www.catalogueoflife.org) : La checklist del Catalogue of Life attualmente indica il dato di circa 1 milione e 500.000 specie verificate, anche se si tratta di una cifra inevitabilmente incompleta per i continui aggiornamenti che hanno luogo nella tassonomia.

I maggiori esperti in materia, nelle numerose pubblicazioni scientifiche prodotte sul tema, ritengono che una stima credibile delle specie presenti rispetto al numero che si suppone possa esistere, ma che al momento è ancora ignoto, possa farci affermare che sulla Terra siano presenti intorno a 5 milioni di specie, con un range che può essere di più o meno 3 milioni.

Quindi il deficit tassonomico (inteso come rapporto tra specie attese e quelle identificate) è sicuramente rilevante.

2. Certificazione genetica: Tipizzazione e Tracciabilità

La certificazione genetica del materiale vegetale è, oggi più che mai, un requisito fondamentale per incrementare la competitività delle aziende, anche nel settore ornamentale, essendo uno strumento unico ed effettivo per l'identificazione certa delle specie vegetali.

Essa contribuisce alla valorizzazione del florovivaismo attraverso:

- Tipizzazione (piante autoctone): individua il legame fra risorse genetiche di interesse ornamentale ed il relativo territorio di origine, promuovendo quindi i prodotti in sintonia con il territorio e con i criteri di sostenibilità.
- Tracciabilità (piante autoctone e piante d'importazione): caratterizzazione dei prodotti autoctoni ai fini della tracciabilità e rintracciabilità della filiera, controlli di identità del materiale non autoctono in entrata.

Per tanto la certificazione genetica del materiale vegetale è un aspetto di estrema importanza per la risoluzione di problematiche attinenti:

1. controversie tassonomiche (sinonimie/omonimie) specie affini
2. programmi di nuova introduzione, miglioramento genetico
3. identificazione precoce di specie con cicli fenologici molto lunghi
4. controlli di corrispondenza di specie vegetali in entrata nei mercati
5. tutela della biodiversità, di specie autoctone o a rischio di estinzione

Tutto ciò ovviamente a tutela dei diversi protagonisti del settore, quali:

- A tutela di produttori e propagatori di piante ornamentali che spesso lavorano con materiale di incerta identità generando controversie per l'assegnazione, registrazione ed il miglioramento genetico delle varietà.
- A tutela del consumatore, risolvendo il problema dell'origine e delle molteplici identità attribuite ad una stessa pianta, che generano confusione e deprezzano il prodotto.

- A tutela di tutti gli utenti del settore, visto l'elevato volume di import/export sia di materiale di riproduzione (talee, sementi), sia di prodotto finito e pronto alla commercializzazione.

Bisogna inoltre considerare oltre all'importanza della certificazione genetica legata agli aspetti di valorizzazione del prodotto, competitività dell'azienda e tutela dei vari protagonisti della filiera floro-vivaistica, anche l'importanza della certificazione genetica sotto gli aspetti legati alle normative vigenti in ambito vivaistico.

In ambito nazionale l'attività di commercializzazione delle piante e fiori non ha prodotto una grande azione legislativa; la più importante è collegata alla applicazione delle direttive europee in merito al controllo della diffusione delle patologie vegetali ad alla conservazione della qualità del prodotto, in particolare in ambito nazionale si fa riferimento al Decreto Legislativo 19 Maggio 2000 n.151 recante disposizioni normative di attuazione della direttiva 98/56/CE relativa alla commercializzazione dei materiali di moltiplicazione delle piante ornamentali, le cui norme si applicano alla commercializzazione dei materiali di moltiplicazione delle piante ornamentali ed affida al Ministero delle politiche agricole e forestali il coordinamento nazionale nella materia ai fini della tutela della qualità dei materiali di moltiplicazione.

Il Ministero delle politiche agricole e forestali con proprio decreto n.33043 del 9 Agosto 2000 ha regolamentato gli aspetti di propria competenza in base a quanto previsto dal citato DLgs n.151/2000, dettando le norme di applicazione per la commercializzazione dei materiali di moltiplicazione delle piante ornamentali.

Secondo tale decreto il materiale di moltiplicazione delle piante ornamentali deve essere, almeno ad una ispezione visiva, sostanzialmente privo di organismi nocivi o malattie – nonché dei relativi indizi o sintomi – tali da compromettere la sua qualità e da ridurre la possibilità di utilizzarlo come materiale di moltiplicazione; quindi si fa riferimento agli aspetti sanitari del prodotto, ma non solo, in quanto si parla anche di identità genetica del materiale di moltiplicazione commercializzato, come si evince infatti dall'art.4 del decreto n.33043 del 9 Agosto 2000 "I materiali di moltiplicazione commercializzati devono avere l'identità e la purezza del genere o della specie o, se del caso, del gruppo di piante cui appartengono e, se commercializzati o destinati ad essere commercializzati con riferimento alla varietà, devono avere altresì l'identità e la purezza della stessa varietà".

(Decreto Legislativo 19 Maggio 2000 n.151; Decreto n.33043 del 9 Agosto 2000).

Altro importantissimo aspetto da considerare ai fini dell'utilità di una caratterizzazione molecolare delle specie vegetali (ma anche animali) sotto il punto di vista giuridico è senza dubbio legato alla commercializzazione delle specie vegetali inserite nelle liste CITES.

La Convenzione di Washington sul commercio internazionale delle specie di fauna e flora minacciate di estinzione, denominata in sigla C.I.T.E.S., è un trattato internazionale nato dall'esigenza di controllare e disciplinare il commercio degli animali e delle piante (vivi, morti o parti e prodotti derivati), in quanto lo sfruttamento commerciale è, assieme alla distruzione degli ambienti naturali nei quali vivono, una delle principali cause dell'estinzione e rarefazione in natura di numerose specie.

Tale convenzione è stata siglata a Washington il 3 marzo 1973 ed è entrata in vigore il 1° luglio 1975. La CITES è stata ratificata dallo Stato italiano con la legge del 19 dicembre 1975, n. 874.

La Convenzione, entrata in vigore per l'Italia il 31 dicembre 1979, è stata inizialmente siglata da 21 Stati.

Attualmente più di 150 Stati sono membri della CITES. Inoltre l'Unione Europea, pur non essendo di per sé membro della CITES ha dettato norme uniformi per la sua applicazione fra gli Stati Membri dell'Unione attraverso alcuni Regolamenti Comunitari, dapprima con il Regolamento (CEE) n. 3626/82 del 3 dicembre 1982 ed ora dal 1997 con il Regolamento (CE) 338/97 del Consiglio del 9 dicembre 1996 modificato da ultimo dal Regolamento (CE) 2704/2000 della Commissione del 30 Novembre 2000. L'ultimo regolamento attuativo è il CE 1808/2001 della Commissione del 30 agosto 2001. Per un quadro completo dell'applicazione della CITES nell'Unione Europea è possibile consultare il sito della Commissione Europea. (Randi E., C. Tabarroni e S. Rimondi, 2002).

In Italia l'attuazione della Convenzione di Washington è affidata a diversi Ministeri: Ambiente, Finanze Commercio con l'Estero, ma la parte più importante è svolta dal Ministero delle Politiche Agricole, come prevede la legge, tramite il Servizio CITES, che cura la gestione amministrativa ai fini della certificazione e del controllo tecnico-specialistico per il rispetto della Convenzione.

La CITES si fonda sul principio che il controllo del commercio sostenibile degli animali e delle piante e dei loro prodotti, costituisca uno strumento di conservazione delle

popolazioni selvatiche, soprattutto se il concetto dell'uso sostenibile delle specie viventi è posto alla base delle legislazioni internazionali e nazionali.

La CITES non è nata solo come strumento per tutelare la conservazione di specie-bandiera, vale a dire di specie particolarmente minacciate e di grande impatto sull'opinione pubblica.

La CITES intende anche tutelare una miriade di altre specie, apparentemente meno carismatiche, ma di straordinaria importanza per la conservazione dell'integrità e della funzionalità degli ecosistemi.

L'estinzione di queste specie (ad es., chiroteri, coralli) avrebbe effetti negativi a cascata sull'intero ecosistema e potrebbe provocare gravi crisi nella diversità biologica regionale. La CITES quindi assume un ruolo importante anche come strumento per la conservazione della biodiversità.

Come opera la CITES?: tutte le importazioni, esportazioni, riesportazioni ed introduzioni dal mare delle specie coperte dalla convenzione deve essere autorizzata mediante un sistema di licenze. Ogni parte (per parte si intende uno Stato per il quale la CITES è entrata in vigore) della Convenzione deve designare una o più autorità di gestione incaricata di gestire tale sistema di licenze e di una o più autorità scientifiche per investigare e comunicare gli effetti del commercio sullo status della specie.

L'attività della CITES si esplica tramite il rilascio di permessi d'importazione ed esportazione di individui vivi o dei loro prodotti, appartenenti a specie coperte dalla CITES, che sono elencate in tre appendici, secondo il grado di protezione di cui hanno bisogno.

L'appendice I comprende specie minacciate di estinzione. Il commercio di esemplari di tali specie è consentita solo in circostanze eccezionali, infatti il commercio delle specie d'Appendice I è di solito vietato, mentre gli scambi di campioni per scopi scientifici o gli scambi di individui fra giardini zoologici possono essere autorizzati.

L'Appendice II include le specie non necessariamente minacciate di estinzione, ma il cui commercio deve essere controllato attraverso permessi di importazione ed esportazione al fine di evitare uno sfruttamento incompatibile con la loro sopravvivenza.

Inoltre elenca specie che pur non essendo direttamente minacciate appartengono a generi, famiglie o ordini che potrebbero essere confuse con specie di Appendice I.

L'Appendice III elenca specie protette localmente, e consente alla comunità internazionale di collaborare al controllo del commercio di specie che godono di protezione a livello nazionale.

Circa 5.600 specie di animali e 30.000 specie di piante sono protette dalla CITES contro l'eccessivo sfruttamento attraverso il commercio internazionale. Essi sono elencati nelle tre appendici CITES. Le specie sono raggruppate nelle appendici a seconda di come sono minacciate dal commercio internazionale.

La **Tab.1** mostra il numero approssimativo delle specie incluse nelle Appendici CITES aggiornata al 2 ottobre 2013

Tab.1 Appendici CITES

	Appendix I	Appendix II	Appendix III
FAUNA			
Mammals	300 spp. (incl. 11 popns) + 23 sspp. (incl. 3 popns)	501 spp. (incl. 16 popns) + 7 sspp. (incl. 2 popns)	45 spp. + 10 sspp.
Birds	154 spp. (incl. 2 popns) + 10 sspp.	1278 spp. (incl. 1 popn) + 3 sspp.	25 spp.
Reptiles	80 spp. (incl. 8 popns) + 5 sspp.	673 spp. (incl. 6 popns)	40 spp.
Amphibians	17 spp.	126 spp.	3 spp.
Fish	16 spp.	87 spp.	-
Invertebrates	63 spp. + 5 sspp.	2162 spp. + 1 sspp.	22 spp. + 3 sspp.
FAUNA TOTAL	630 spp. + 43 sspp.	4827 spp. + 11 sspp.	135 spp. + 13 sspp.
FLORA	301 spp. + 4 sspp.	29592 spp. (incl. 162 popns)	12 spp. (incl. 2 popns) + 1 var.
GRAND TOTAL	931 spp. + 47 sspp.	34419 spp. + 11 sspp.	147 spp. + 13 sspp. + 1 var.

spp. = species sspp.= subspecies var. = variety popn= population

(<http://www.cites.org/eng/disc/species.php>)

Gli aggiornamenti alle Appendici CITES sono costantemente riportati nel sito CITES (<http://cites.org/>). Gli aggiornamenti ai Regolamenti Comunitari sono riportati nel sito della Commissione Europea, citato in precedenza.

Le linee guida ed i criteri per l'iscrizione (o la cancellazione) delle specie delle appendici, vengono di tanto in tanto stabiliti ed aggiornati dalla Conferenza delle Parti, ovvero il complesso degli stati membri (CoP), che redige ed approva delle risoluzioni, che indicano linee-guida e raccomandano specifiche azioni agli Stati membri della CITES.

L'inclusione delle specie nelle tre Appendici vincola le parti (stati membri) ad applicare specifici controlli sull'importazione e l'esportazione, e di conseguenza i singoli stati emettono regolamenti in applicazione della CITES, e tali legislazioni nazionali devono anche essere basate sulle risoluzioni approvata dalla CoP.

Gli Stati partner devono istituire una o più Autorità di gestione ed un'Autorità scientifica, le quali operano autonomamente l'una dall'altra.

L'Autorità di gestione rilascia i necessari permessi CITES e compila le informative annuali per la CITES, partecipa ai lavori delle CoP, ecc. L'Autorità scientifica, tramite le delibere della Commissione Scientifica CITES (CSC) determina se il commercio di una particolare specie sia dannoso alla sua sopravvivenza, controlla i volumi di commercio nel rispetto delle quote e ne valuta gli impatti sulle popolazioni naturali.

In Italia si sono istituite tre Autorità di gestione: la principale è il Ministero dell'Ambiente e della Tutela del Territorio; altri organi di gestione sono il Ministero per le Politiche Agricole, Divisione II, il Servizio CITES del Corpo Forestale dello Stato ed il Ministero delle Attività Produttive.

L'Autorità scientifica ha sede presso la Divisione II della Direzione Conservazione della natura del Ministero dell'Ambiente e della Tutela del Territorio.

3. Identificazione delle specie vegetali

3.1 Identificazione fenotipica

È l'identificazione basata sul fenotipo della pianta, ovvero l'insieme delle sue caratteristiche morfologiche e funzionali. I controlli sul fenotipo consistono nel rilevare i caratteri morfologici, fenologici ed agronomici della pianta, il che viene fatto solitamente compilando le schede pomologiche.

Anche se questo tipo di analisi è fondamentale ai fini dell'identificazione, la stessa presenta comunque alcuni limiti:

- Innanzitutto perché l'identificazione delle specie con i tradizionali metodi di riconoscimento morfologico, risulta essere molto lungo, laborioso e richiede un certo livello di competenza non comune.
- Inoltre i sistemi di riconoscimento morfologico risultano essere spesso inadeguati per il riconoscimento delle specie che caratterizzano la biodiversità sulla terra a causa di alcuni fattori come l'omoplasia (Vences et al., 2005), e la plasticità fenotipica (Saunders, 2005) che sono frutto dell'adattamento di una data specie all'ambiente e che quindi possono generare la comparsa di strutture morfologiche simili anche in individui non collegati filogeneticamente o viceversa a strutture morfologiche diverse anche nell'ambito dello stesso genotipo e quindi indurre ad errori di identificazione morfologica di alcuni taxa.
- Altro fattore da tenere sicuramente in considerazione è che la caratterizzazione su base morfologica spesso è risolutiva solo in una particolare fase della vita e pertanto molti individui negli stadi giovanili sono difficili da identificare con esattezza (Pegg et al., 2006).
- A questo si va ad aggiungere anche, che in particolare per quanto riguarda il mondo vegetale, la difficoltà di caratterizzazione su base morfologica può risultare complessa anche in conseguenza del fatto che i rilievi condotti sul fenotipo per alcuni caratteri possono essere eseguite solo in determinate fasi fenologiche basta ad esempio pensare all'epoca di fioritura di una pianta che può essere stabilita solo nella stagione primaverile, oppure i caratteri del frutto che possono essere rilevati solo una volta superata la fase iniziale improduttiva, etc.

- Altro limite importante alla identificazione su base fenotipica, risiede nel fatto che il fenotipo della pianta può essere influenzato oltre che dall'ambiente, anche da altri fattori, come lo stato sanitario (es. presenza/assenza di virus o di altri organismi patogeni), o come i fattori epigenetici (es. fenomeni di ringiovanimento, di inerzia ormonale, etc.).

Per ovviare a tutti questi limiti, a supporto dell'identificazione fenotipica, è importante affiancare una corretta identificazione genotipica, attraverso tecniche di biologia molecolare, che sia da supporto alla caratterizzazione a livello fenotipico.

3.2 Identificazione genetica

È l'identificazione basata sull'analisi e lo studio del genotipo della pianta.

Il genotipo non è altro che l'insieme di tutti i geni che compongono il DNA di un organismo (corredo genetico).

In questo ambito assume notevole importanza l'obiettivo di valorizzare la produzione vivaistica grazie ai nuovi strumenti di riconoscimento molecolare delle piante, al fine di rispondere alle nuove e crescenti esigenze di mercato, improntate all'innovazione di prodotto ed alla sua tipizzazione nonché a supportare gli aspetti commerciali legati all'identità ed alla tracciabilità delle piante.

Così la tassonomia basata sul DNA si è rivelata un prezioso supporto alla tassonomia classica permettendo di far fronte al crescente bisogno di accurate e accessibili informazioni tassonomiche (Tautz et al., 2003).

Già da diversi anni le azioni di miglioramento genetico si avvalgono di "strumenti genetici" individuati attraverso lo studio molecolare del genoma delle specie vegetali.

In particolare, l'avvento dei marcatori molecolari ha segnato una notevole svolta nel mondo della genetica vegetale consentendo la costruzione di mappe genetiche di associazione e l'individuazione di geni responsabili di caratteri agronomici (Quarta et al., 1998).

L'identificazione molecolare ha applicazioni in diversi campi. Nella conservazione del germoplasma, per esempio, i marcatori molecolari consentono di individuare alleli rari garantendo la salvaguardia della biodiversità, di effettuare il controllo dell'identità degli esemplari presenti o di individuare eventuali errori di classificazione. Negli studi

tassonomici, i marcatori sono importanti per le classificazioni botaniche e l'analisi della relazioni filogenetiche (Varshney et al., 2005).

L'identificazione a livello molecolare presenta molti vantaggi che altro non sono che i limiti legati alla caratterizzazione a livello fenotipico, di cui si è già discusso in precedenza, ma presenta anch'essa dei limiti legati spesso alla scarsa capacità per svariati motivi di discriminare taxa diversi, che invece potrebbero essere più facilmente discriminati con l'analisi del fenotipo. Pertanto l'identificazione a livello molecolare non deve assolutamente sostituire quella basata sul fenotipo ma essere da supporto ad essa per una corretta caratterizzazione dei diversi individui, quindi per una corretta identificazione sarebbe opportuno ***integrare i due approcci***, per colmare i limiti di entrambi.

3.2.1 Il DNA

Il DNA è un acido nucleico (acido desossiribonucleico), dal punto di vista chimico è un polimero organico composto da monomeri denominati nucleotidi.

I nucleotidi sono formati principalmente da tre componenti: Il deossiribosio (ovvero zucchero pentoso), una base azotata collegata al deossiribosio con legame denominato N-glicosidico e un gruppo fosfato. Nella formazione dei nucleotidi, possono essere impiegate quattro diverse basi azotate: adenina (A), citosina (C), guanina(G) e timina (T).

È una molecola costituita da una doppia elica, cioè due eliche (ognuna composta da una sequenza di nucleotidi) avvolte a spirale l'una sull'altra e tenute unite dai legami che si instaurano tra le basi azotate complementari presenti sulle due eliche A-T e C-G.

L'ordine nella disposizione sequenziale dei nucleotidi costituisce l'informazione genetica, la quale è tradotta nella sintesi di proteine importanti per costruire i tessuti e gli organi e per poter mettere in atto tutti quei processi biologici e chimici che garantiscono la sopravvivenza dell'organismo. Tutte queste informazioni formano insieme il codice genetico.

Nello specifico, sono i geni che contengono le informazioni genetiche per la sintesi delle proteine (codificano per la sintesi di proteine), che altro non sono che porzioni del codice genetico, ovvero porzioni della sequenza nucleotidica del DNA e sono organizzati in strutture dette cromosomi.

Nelle cellule eucariotiche il gene è costituito da sequenze codificanti dette *esoni* che da sequenze non codificanti, dette *introni*.

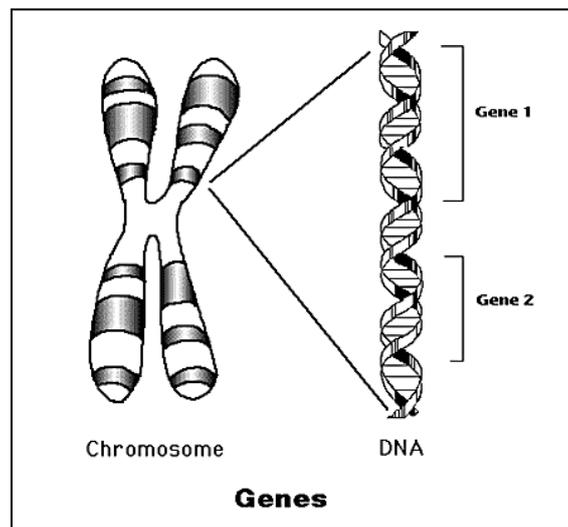


Fig 1 .Rappresentazione della relazione che intercorre tra gene, DNA e cromosoma

Nelle cellule vegetali il DNA è localizzato in diversi organuli, oltre che nel nucleo, infatti possiamo distinguere:

- **DNA nucleare**, il nucleo cellulare ha una funzione essenziale nella trasmissione dei caratteri ereditari e nel controllo del metabolismo cellulare, poiché la maggior parte del DNA e quindi delle informazioni genetiche è localizzato nel nucleo. Nel nucleo il DNA non è sparso, ma ben organizzato in un complesso dato dall'associazione del DNA stesso con delle proteine, dette istoni, tale associazione prende il nome di *cromatina*. Ma oltre che nel nucleo, nelle cellule vegetali il DNA è localizzato anche in altri organuli, si parla in questo caso di "DNA extra nucleare" o "DNA citoplasmatico", nello specifico esso è presente in altri due organuli cellulari:

- **DNA mitocondriale**, è quel DNA contenuto nei mitocondri, che sono organelli cellulari di forma allungata (reniforme), di importanza fondamentale perché impegnati nei processi di respirazione cellulare (ciclo di Krebs e fosforilazione ossidativa). Il DNA mitocondriale reca il codice genetico necessario per la sintesi di alcune proteine che si trovano esclusivamente all'interno del mitocondrio stesso.

- **DNA del cloroplasto**, è quel DNA contenuto all'interno dei cloroplasti, che sono organelli cellulari, presenti nelle cellule vegetali al cui interno si svolge il processo di fotosintesi clorofilliana, i cloroplasti così proprio come i mitocondri presentano un loro proprio DNA, solitamente di tipo circolare (cpDNA, Chloroplast DNA). Il numero di molecole di cpDNA in una cellula vegetale è molto variabile: dipende sia dal numero di

molecole di cpDNA presenti in ciascun cloroplasto, sia dal numero di cloroplasti per cellula.

Il sequenziamento completo del cpDNA della pianta del tabacco, *Nicotiana tabacum*, ha mostrato che i cloroplasti di questa specie contengono 150 geni.

Le conoscenze sul genoma mitocondriale sono molto più dettagliate rispetto a quelle del genoma del cloroplasto. E' disponibile la sequenza completa del DNA mitocondriale dell'uomo e di numerosi altri animali e la sequenza del DNA cloroplastico di riso e tabacco.

L'esistenza del DNA extra nucleare, supporta la teoria dell'endosimbiosi, che sostiene che le cellule eucariote apparvero per la prima volta quando una cellula procariote venne assorbita da un'altra cellula senza essere digerita. Si pensa che queste cellule siano poi entrate in un rapporto simbiotico.

3.2.2 Marcatori molecolari

I marcatori molecolari (M.M.) possono essere definiti, come frammenti di DNA cromosomico (variabile tra 50-3000 bp) compreso tra due regioni oligonucleotidiche note (4-30 bp), che in virtù della loro presenza caratterizzano in modo univoco quella regione del DNA in questione e quindi l'individuo a cui appartiene.

Queste regioni del DNA (M.M.) possono essere rilevate con l'ausilio di sonde (probe) o con l'ausilio di inneschi (primers) specifici, che vanno a riconoscere le regioni oligonucleotidiche a sequenza nota che fiancheggiano il marcatore stesso. Maggiore sarà il numero di polimorfismi associati ad un determinato marcatore, migliore sarà il potere discriminante del marcatore stesso e quindi le informazioni che fornirà.

I marcatori molecolari non sono necessariamente riferibili all'attività di specifici geni, in quanto possono essere ubicati sia in regioni espresse, sia in porzioni spaziatrici dei diversi geni (regioni non espresse).

In tempi recenti si è assistito ad una notevole proliferazione dei sistemi molecolari per lo studio genomico, che si differenziano per il tipo di sequenze analizzate e/o per il tipo di tecnica impiegata, alcuni tipi di marcatori, come RFLP, sono basati sul procedimento di ibridazione tipo "Southern" (*Southern Blot Hybridization, SBH*), mentre altri, RAPD, SSR (microsatelliti), AFLP, sono basati sulla "Reazione a Catena della Polimerasi" (*PCR, Polimerase Chain Reaction*).

Tra tutte queste tecniche è inoltre possibile un'ulteriore distinzione tra: i) marcatori di tipo dominante (ad esempio, RAPD, AFLP, ecc.), ad ogni locus si può evidenziare la presenza o l'assenza della banda, ma non è possibile distinguere la situazione eterozigote da quella omozigote per lo stesso allele marcatore; ii) marcatori di tipo co-dominante (ad esempio, RFLP e SSR) che permettono di distinguere i loci omozigoti da quelli eterozigoti.

Tra i marcatori molecolari, maggiormente utilizzati, abbiamo:

- **RAPDs** (Random Amplified Polymorphic DNA), marcatori di tipo dominante, basati sull'analisi simultanea di molti loci genomici, che applicano l'amplificazione di tratti cromosomici casuali (sequenze di DNA ignote) mediante l'impiego di inneschi oligonucleotidici a 10 bp (frammenti decamerici) a composizione arbitraria, e che pertanto non richiedono alcuna conoscenza della sequenza stessa. Nel caso in cui i

primer in seguito ad amplificazione riescono ad associarsi al DNA, l'amplificazione produrrà una serie di frammenti, che interpretati, permetteranno di ricostruire un profilo dell'organismo in analisi. E' una metodica facile e veloce (nessuna conoscenza della sequenza bersaglio), con basse esigenze in termini di costo e di tempo sia per lo sviluppo che per l'impiego, è di contro una metodica che presenta una scarsa riproducibilità, quindi non standardizzata, presenta un potere discriminatorio (polimorfismo) non particolarmente alto, a differenza dei primer specie-specifici, oltre ad essere dei marcatori di tipo "dominante", ad ogni locus permettono di evidenziare la presenza o assenza della banda ma non permettono di distinguere la situazione omozigote da quella eterozigote per lo stesso allele marcatore, a differenza invece dei marcatori di tipo codominanti.

- **AFLPs** (Amplified Fragment Length Polymorphism), come i RAPD sono marcatori di tipo dominante, basati sull'analisi simultanea di molti loci genomici, che applicano l'amplificazione di tratti cromosomici casuali. La tecnica AFLP è basata sulla combinazione di due diverse fasi: (i) digestione del DNA, con una coppia di enzimi di restrizione (che catalizzano la rottura dei legami fosfodiesterici del DNA), (ii) amplificazione selettiva dei frammenti ristretti mediante l'ausilio di primers specifici complementari alle sequenze di taglio. È una tecnica che presenta dei vantaggi, come il fatto che non richiede informazioni sulle sequenze, è universalmente applicabile a tutti gli organismi (non presenta specificità) e presenta un livello di polimorfismo abbastanza elevato, di contro è una metodica che non ha un'alta riproducibilità, è tecnicamente complessa e quindi richiede tempi molto lunghi per lo sviluppo e l'impiego e come i RAPD sono marcatori di tipo "dominante".

- **RFLPs** (Restriction Fragment Length Polymorphism), a differenza dei primi due, sono marcatori di tipo co-dominante, che prevedono l'ibridazione di tratti cromosomici a sequenza nota, mediante l'impiego di sonde specifiche per determinati loci genomici, si basano sullo studio del polimorfismo associato alla variazione della lunghezza dei frammenti di restrizione del DNA, mediante l'ausilio di enzimi di restrizione e di sonde appropriate con sequenza nota, specifiche per il sito di restrizione, e marcate con fluorocromi, con il procedimento di ibridazione tipo "Southern" (Southern Blot Hybridization, SBH).

Presenta tanti limiti dovuti al fatto che richiede DNA di alta qualità e quantità, ma soprattutto richiede tempi lunghi e costi elevati, inoltre la comparazione dei pattern

ottenuti non è sempre di facile interpretazione, e inoltre bisogna cercare delle sonde abbastanza polimorfiche, il che ha un costo non indifferente, però ha anche punti di forza come, un'alta riproducibilità ed il fatto che sono inoltre marcatori di tipo codominante.

- **SSR** (*Simple Sequence Repeats*), i micro satelliti sono brevi unità ripetute di 1-6 bp. E' stato stimato facendo un confronto tra piante ed animali (Beckmann e Weber, 1992; Lagercrantz et al., 1993; Morgante e Olivieri, 1993) che i microsattelliti sono molto più abbondanti nel genoma degli animali. Nei mammiferi, infatti, è possibile riscontrare un microsattellite ogni 6Kb, mentre nelle piante se ne ritrova uno soltanto ogni 50 Kb. Inoltre secondo (Wang et al, 1994), nelle piante i microsattelliti sono molto abbondanti nel DNA nucleare e rari nel DNA extracromosomico.

Attraverso l'utilizzo di primer specifici, complementari alle regioni che fiancheggiano le sequenze ripetute (SSR), è possibile amplificare queste zone del genoma e renderle quindi rilevabili mediante tecniche elettroforetiche. I primer consentono di amplificare singoli microsattelliti che possono differire tra individui non per il motivo di base, ma per il numero di volte che questo motivo è ripetuto.

Uno dei campi di maggior interesse in cui vengono utilizzati i microsattelliti è la ricerca di polimorfismi. Nel caso dei microsattelliti, tale polimorfismo è dato dalla variazione del numero di sequenze SSR ripetute.

L'elevato polimorfismo (gli SSR sono una classe di marcatori molecolari capaci di mettere in evidenza una diversità genetica difficilmente rilevabile con altri tipi di marcatori, poiché i polimorfismi sono localizzati in regioni ipervariabili e non espresse del genoma.), la codominanza (presenza in caso di eterozigosi di segnale per entrambi gli alleli), l'uniformità di distribuzione nel genoma, la riproducibilità e la facile realizzazione delle indagini condotte per rilevarli, rappresentano dei vantaggi per questa tipologia di marcatori.

Di contro però nonostante i tempi e i costi per l'impiego sono relativamente bassi, risultano però alti i tempi ed i costi necessari per il loro sviluppo, poiché per la loro alta specificità richiedono indagini abbastanza accurate sulla specie oggetto di studio, e quindi non è una tecnica da poter applicare su larga scala, ovvero per condurre indagini in tempi brevi su un elevato numero di taxa (non universalità).

- **SNPs** (Single Nucleotide Polymorphism) costituiscono una nuova ed alternativa tecnica di identificazione molecolare.

Uno SNP, è una differenza in una singola coppia di basi in un sito del DNA, quindi sono dei potenti marcatori in grado di individuare variazioni di un singolo nucleotide in un determinato sito del genoma.

Gli SNPs possono presentarsi all'interno di una sequenza codificante di un gene, all'interno di una regione intronica o in una regione intergenica.

Questo tipo di marcatori si è rilevato fortemente discriminante sia in campo umano (Collins et al., 1998), sia in campo vegetale su mais (Rafalski et al., 2001) e soia (Rafalski, 2002).

Gli SNPs possono essere identificati attraverso diversi approcci che includono il risequenziamento degli ampliconi ottenuti in PCR, con o senza pre-screening e lo sviluppo di SNP da librerie EST (Expressed Sequence Tag) (Rafalski, 2002).

Per le loro caratteristiche, gli SNPs, si presentano come dei potenti marcatori molecolari, grazie al loro elevato polimorfismo, alla loro alta densità lungo tutto il genoma ed alla loro alta riproducibilità, nonché il fatto che sono dei marcatori di tipo codominante, come i microsatelliti.

Di contro hanno anche loro dei limiti, dovuti alla loro alta specificità e quindi alla conoscenza delle sequenze degli alleli da studiare che comporta un lungo lavoro di messa a punto.

- **Marcatori Barcode**, Nel 2003 un nuovo sistema di identificazione, DNA barcoding, è stato sviluppato dai ricercatori dell'Università di Guelph (Canada). Questo approccio si basa sull'analisi della variabilità all'interno di una regione genomica nota e standardizzata. Questa tecnica, fornisce la possibilità di avere un unico identificatore diagnostico per le differenti specie viventi, sia animali che vegetali (alto tasso di universalità).

Questo nuovo approccio ha la potenzialità di contribuire in modo radicale al lavoro che i biologi stanno svolgendo al fine di catalogare la vita sulla terra, rendendo possibile l'identificazione di qualsiasi organismo a partire da un piccolo pezzo di tessuto.

Come vedremo dettagliatamente più avanti, nel capitolo dedicato alla metodica del *DNA barcoding*, la tecnica in questione presenta tanti punti di forza ma anche punti di debolezza.

In definitiva i requisiti che un marcatore molecolare ideale deve avere sono:

- Elevato polimorfismo
- Espressione stabile (non deve essere influenzato da fattori come l'ambiente, altri geni o caratteristiche transitorie degli organismi tra cui l'età, le dimensioni, ecc..)
- Facile osservazione e determinazione (basso input in termini di risorse economiche e di tempo)
- Codominanza
- Riproducibilità entro e fra diversi laboratori
- Determinazione attraverso metodiche applicabili a specie diverse (universalità)

Il marcatore perfetto non esiste!!!

Le metodiche di identificazione genetica, basate sull'impiego di marcatori molecolari sono diverse, ognuna con dei vantaggi e dei limiti, pertanto l'impiego di una tecnica piuttosto che di un'altra, deve ovviamente tenere in considerazione le finalità stesse dello studio genetico.

3.2.3 DNA barcoding

Tra le tecniche molecolari, si è diffuso, ormai a livello mondiale, un nuovo approccio allo studio della biodiversità con tutte le problematiche connesse a tale studio: il *DNA barcoding*, letteralmente “codice a barre del DNA”. Il nome di questo approccio si riferisce al metodo di identificazione con cui uno scanner distingue i vari prodotti commerciali utilizzando i codici a barre lineari o “UPC” (Universal Product Code).

Tale approccio di indagine molecolare è stato proposto per la prima volta all’attenzione della comunità scientifica nel 2003 dal genetista di popolazione Paul Hebert dell’ University of Guelph (Canada) (Hebert et al 2003). In tale lavoro veniva utilizzato per l’identificazione delle specie, una sequenza genica situata nella regione del gene mitocondriale COI, codificante per la sub unità I della citocromo-c ossidasi (noto anche come frammento respiratorio di Warburg, è un enzima che interviene nella fase terminale della catena respiratoria), quindi viene sfruttata la variabilità di un marcatore molecolare per l’identificazione di identità biologiche.

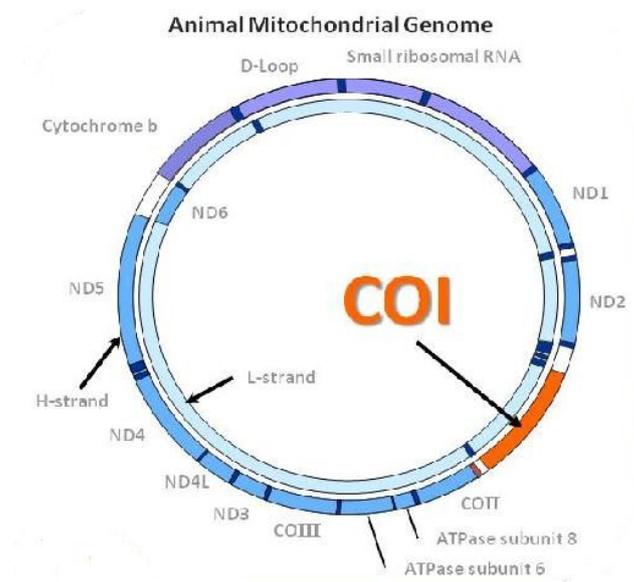


Fig 2 Genoma mitocondriale animale. COI rappresenta il marcatore *barcode* utile per l’identificazione delle specie

Nel corso degli anni COI è stato usato con successo in diversi taxa animali, inclusi uccelli (Hebert et al., 2004b), artropodi (Barrett and Hebert, 2005; Hogg and Hebert, 2005), pesci (Ward et al., 2005) e lepidotteri (Hebert et al., 2003; Hebert et al., 2004a; Janzen et al., 2005).

Nei vegetali COI non si è rivelato un ottimo marker identificativo delle specie, a causa del basso tasso evolutivo del genoma mitocondriale, rispetto a quello degli animali.

Nasceva quindi l'esigenza di scegliere un nuovo genoma e una nuova regione adatta al mondo vegetale che avrebbe dovuto possedere le stesse caratteristiche del frammento COI.

Al fine di ovviare a questo problema, negli ultimi anni sono stati identificati altri marcatori per il DNA barcoding delle piante. Si tratta di sequenze di DNA presenti in alcuni tratti del genoma del cloroplasto, come la regione intergenica trnH-psbA, il gene matK o il gene rbcL, che presentano caratteristiche simili alla coxI utili per l'identificazione delle specie.

Occorre però sottolineare che questo approccio all'identificazione degli organismi viventi non è sostitutivo dei tradizionali metodi di classificazione e descrizione, basati su minuziose analisi dei tratti morfologici e di altre caratteristiche cariologiche, ecologiche, etologiche e geografiche, ma ne rappresenta un'integrazione, con alcuni importanti ed esclusivi vantaggi:

- E' innanzitutto un *sistema di discriminazione generalista* (universalità) con uno o pochi marcatori molecolari (DNA barcode) è possibile differenziare un vasto pannello di organismi viventi. In linea generale si può affermare che il DNA barcoding non è un sistema di discriminazione pensato per un esperto di un certo gruppo tassonomico. E' invece lo strumento ideale per chi non ha particolari conoscenze nel gruppo in esame, grazie anche alla presenza di una banca dati di riferimento.

- Il DNA barcoding è *molecolarizzazione e informatizzazione della tassonomia*. In particolare, la molecolarizzazione è un processo iniziato decenni prima del lancio del DNA barcoding, e non costituisce quindi la vera innovazione. Dall'altra parte invece, l'informatizzazione è il vero cuore del DNA barcoding. Tutte le campagne dell'iBOL (International Barcode of Life) hanno come fine ultimo la generazione di database estesi, in cui tassonomia tradizionale, molecolare, biogeografia e altre informazioni si integrino.

L'efficacia dell'identificazione genetica è legata quindi alla costituzione di un ampio e robusto dataset molecolare, come base di confronto di segmenti di DNA ottenuti.

- Il DNA barcoding è *standardizzazione*. Per la prima volta da secoli uno stesso laboratorio di ricerca può occuparsi di differenti tematiche, unificate dallo stesso modo di: raccogliere e collezionare gli organismi; estrarre, amplificare e sequenziare il DNA; analizzare i dati; diffondere i risultati.

- Permette di risalire alla determinazione di una specie anche analizzandone una piccola porzione. È, infatti, sufficiente un frammento fogliare per l'analisi, fatto che rende il metodo meno distruttivo;

- Grazie alle attuali tecnologie è possibile eseguire analisi di DNA barcoding molto velocemente e a costi accessibili;

- Le moderne tecniche biomolecolari di estrazione del DNA permettono di lavorare anche su materiale molto vecchio, quale quello conservato negli erbari.

Anche tale tecnica, come tutti gli altri approcci di indagine molecolare, presenta oltre a dei punti di forza, anche dei punti di debolezza, quali:

- Assume che la variazione intraspecifica è trascurabile, o in ogni caso minore di quella interspecifica.

- Nessun singolo gene funzionerà per tutti i taxa.

- Alcuni degli aspetti più interessanti sono basati su tecnologie future (ad es. sequenziatori portatili).

In definitiva la grande innovazione non è la capacità discriminatoria e quindi il potere di identificazione delle specie, il DNA barcoding non è perfetto, perché un sistema perfetto non esiste, abbiamo tecniche molecolari che sono più sensibili del DNA barcoding, la vera forza del DNA barcoding è che si tratta del più grande e generale sistema discriminatorio che ha introdotto dei rilevanti automatismi e schematizzazioni nella tassonomia. Il tentativo di fornire delle metodiche universali di catalogazione dei reperti biologici, di analisi del dato, di fruizione dei risultati è la vera forza del sistema.

Quindi seppure non rivoluzionario dal punto di vista metodologico, la grande novità del DNA barcoding è la scala di analisi e la standardizzazione del metodo.

3.2.3.1 Utilità ed applicazioni del DNA barcoding

La possibilità di discriminare velocemente un campione vegetale o animale porta benefici non solo a studi di tassonomia di base, ma le applicazioni potenziali includono la capacità di un'immediata identificazione di specie criptiche e aliene, il monitoraggio di specie a rischio, le dimensioni e la crescita di popolazioni, il monitoraggio della qualità ambientale.

Pertanto notevoli e svariati risultano essere le applicazioni pratiche del DNA barcoding: Nel campo delle scienze naturali questo approccio rivoluziona i metodi di stima e analisi della diversità biologica, in quanto permette di identificare gli organismi di una determinata area semplicemente esaminando parte del loro DNA, capire se si spostano, come si spostano, da dove provengono e se le popolazioni presenti sono eterogenee o costituite da uno solo o da pochi genotipi; Le autorità sanitarie potrebbero identificare velocemente gli organismi vettori (es.zanzare) di malattie infettive, permettendo la tempestiva adozione di metodi di controlli mirati; Gli organi preposti al controllo delle frodi alimentari potrebbero avvalersi di questa metodica per accertarsi che non vi siano delle irregolarità nel commercio dei prodotti alimentari; Gli esperti di tassonomia potrebbero individuare esemplari geneticamente distinti, accelerando la catalogazione delle nuove specie prima che si estinguano; I musei potrebbero usare la tecnica per analizzare vaste collezioni di esemplari che hanno accumulato nel tempo, come ausilio per la scoperta di nuove specie conservate nelle loro teche.

Questi sono alcuni dei motivi che spiegano la rapida diffusione della tecnica anche al di fuori della comunità scientifica, tanto che il numero dei nuovi progetti è difficilmente calcolabile. Il DNA barcoding si è, infatti, dimostrato una tecnica estremamente flessibile, che ha rapidamente lasciato l'ambito della ricerca pura per estendersi a svariati campi, dal controllo della qualità merceologica all'ambito forense e delle frodi alimentari, dalle indagini ambientali alla diagnostica medica e veterinaria.

La determinazione dei codici a barre genetici sta contribuendo ad una più rapida catalogazione della biodiversità. Uno dei principali risultati è l'esistenza di un numero di specie maggiore di quanto si credesse.

Tale scoperta si deve allo studio dei codici a barre offerti dalle *criptospecie*, organismi che pur essendo fenotipicamente uguali presentano differenze genetiche tali da farli ascrivere a specie diverse.

L'uso dei codici a barre genetici è funzionale alla realizzazione di una mappa della diversità a livello del DNA, e rappresenterà la base per la scoperta di nuove specie. Ma quali potranno essere le implicazioni di questa iniziativa? I difensori del DNA barcoding asseriscono che questa tecnologia darà impulso alle collezioni biologiche, abbreviando i tempi richiesti per l'identificazione delle nuove specie. I detrattori, invece, restano scettici obiettando che l'identificazione effettuata su basi genetiche contribuirà a metter definitivamente in soffitta la sistematica tradizionale, già in crisi, e sarà facile preda di speculazioni commerciali. Certamente sussistono aspetti metodologici ancora irrisolti. Ad esempio, il gene COI non rappresenta il marcatore ideale per le piante, dato che il suo tasso di sostituzione è troppo lento e non sempre consente il riconoscimento di specie affini. Inoltre, il DNA barcoding permette di affrontare un ambito molto circoscritto della tassonomia, in quanto si limita ad attribuire determinate sequenze di DNA a specie note, identificate attraverso metodologie tradizionali (basate, ad esempio, su caratteristiche morfologiche e fisiologiche). In tale contesto la sua funzione è quella di assegnare campioni ignoti a specie già tradizionalmente caratterizzate.

Laddove la specie sia ancora indefinita, la tecnica del DNA barcoding potrebbe rivelare i propri limiti, sebbene sarebbe sicuramente in grado di scoprire nuove specie "criptiche". Il DNA barcoding non sostituisce il sistema linneano di classificazione; il suo scopo non è creare un sistema di classificazione basato esclusivamente sul DNA. Lo scopo è di aiutare l'identificazione di nuove specie nell'ambito di taxa ancora non ben definiti. Le nuove specie scoperte dovranno sempre possedere un nome in base al sistema linneano di classificazione, essere studiate accuratamente con immagini ad alta risoluzione di alcuni esemplari, ed essere accompagnate da informazioni precise sui siti di raccolta. Il DNA barcoding non segna pertanto la fine dell'alfa tassonomia ma la assiste e la coadiuva: il DNA barcoding ha pertanto valore solo se accompagnato da una corretta tassonomia tradizionale (Hebert et al. 2004a). Lo studio molecolare è efficiente, infatti, solo quando è possibile il confronto tra l'esemplare sconosciuto oggetto dello studio e un esemplare di riferimento ben descritto dal punto di vista tassonomico, morfologico ed ecologico (Dupont et al. 2007).

Auspicabile e molto prezioso, è quindi il contributo di tassonomi e naturalisti alla ricerca e possibile risoluzione dei misteri della biodiversità.

3.2.3.2 Caratteristiche di un marcatore barcoding e approccio *multilocus*

Esistono, diversi requisiti affinché un marker possa essere considerato appropriato per il *DNA barcoding*, in primo luogo è opportuno che il marcatore possieda un'ampia copertura tassonomica (detta anche universalità), che permetterebbe l'applicabilità, del gene scelto come marcatore barcode, ad un numero di taxa più ampio possibile e avere un alto successo di PCR e di sequenziamento, importante è anche l'alta capacità di risoluzione del gene, ovvero l'abilità di un dato barcode di differenziare le specie e si basa sulla quantità di differenze interspecifiche tra le sequenze di DNA (Polimorfismo). Altro presupposto fondamentale è che il marker molecolare scelto come barcode dovrebbe mostrare un variabilità interspecifica più alta di quella intraspecifica e che le variabilità inter e intraspecifiche siano separate da una certa distanza (discontinuità tra le variabilità intra e interspecifiche) detta "barcoding gap" (Meyer and Paulay, 2005). Il marcatore ideale consiste pertanto in una regione altamente variabile, che prevede la discriminazione delle specie, fiancheggiata da regioni altamente conservate per le quali possono essere progettati i primer adeguati (Saunders and Kucera, 2010).

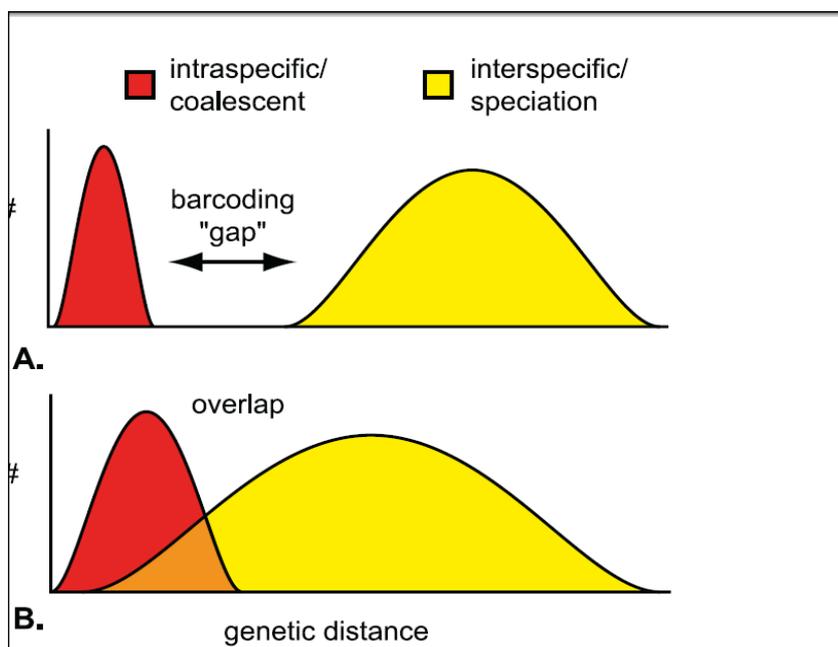


Fig.3 Schema del “Barcoding gap” second Meyer & Paulay (2005). La variazione intraspecifica è rappresentata dalla curva rossa, mentre la variazione interspecifica in giallo. (A) Situazione ideale per il barcoding con le distribuzioni discrete delle divergenze e assenza di overlap. (B) Situazione alternativa con assenza di gap e sovrapposizione tra le due distribuzioni.

Ogni marcatore barcode presenta punti di forza e punti di debolezza.

Pertanto per le piante, i protocolli del Barcoding fanno riferimento alle indicazioni del Plant working Group, che suggerisce l'impiego di un approccio multilocus (Hollingsworth et al., 2011; Fazekas et al., 2008, 2009).

La preferenza dell'ampio gruppo di lavoro CBOL (Consortium for the Barcode of Life), per l'identificazione a livello genotipico, delle specie vegetali, testando numerosi marcatori e seguendo 3 criteri: universalità (amplificazione e sequenziamento applicabile ad un gran numero di taxa), qualità della sequenza ottenuta e potere discriminante, è stata quella costituita dall'associazione dei due marcatori *rbcL*+*matK*, indicati come “core markers”, a cui affiancare eventualmente marcatori aggiuntivi, in grado di discriminare gruppi tassonomici che presentano scarsa variabilità genetica, come la regione intergenica *trnH-psbA*.

- Il marcatore ***rbcl***, è un gene presente nel cloroplasto delle specie vegetali, che codifica per la sub-unità della ribulose-1.5- bisphosphate carboxylase/oxygenase (RUBISCO), il Rubisco è un enzima appartenente alla classe delle liasi, tale enzima interviene nella fase di organizzazione del Ciclo di Calvin (stadio della fotosintesi clorofilliana che porta alla sintesi di glucosio) fissando un atomo di carbonio proveniente da una molecola di anidride carbonica gassosa su una molecola a 5 atomi di carbonio, il ribulosio-1,5-bisfosfato, quindi catalizza la reazione del ribulosio-1,5-bisfosfato con O₂ e CO₂.

Tale marcatore, presenta parecchi punti di forza, quali:

- Facile da amplificare
- Facile da sequenziare
- Facile da allineare
- Alta universalità (applicabile con successo ad un gran numero di taxa)

Presenta però anche punti di debolezza:

- Basso livello di discriminazione

- Il marcatore ***matK***, è un gene presente nel cloroplasto, che codifica per la sintesi della proteina maturase K, è una proteina che interviene nello “*splicing*”, con questo termine in biologia molecolare e genetica, si intende una modifica del nascente pre-mRNA (pre-RNA messaggero), in cui gli introni vengono rimossi e gli esoni uniti, tutto questo è necessario per ottenere mRNA maturo che a sua volta può produrre una corretta sintesi proteica (traduzione), in quanto l’ mRNA maturo è formato da soli esoni.

Il punto di forza di questo marcatore è:

- L’elevato potere discriminatorio.

Però di contro:

- E’ più difficile da amplificare (non c’è regione di amplificazione universale).

Questa combinazione ***rbcl+matK***, rappresenta un complesso compromesso tra:

- Universalità
- Qualità delle sequenze
- Potere discriminante
- Costo

- Al di là della suddetta associazione, il marcatore plastidiale più utilizzato, è la regione intergenica non codificante **trnH-psbA**.

Nello specifico il gene *psbA*, presente nel DNA del cloroplasto, codifica per un polipeptide (proteina D1 32 KDalton) che è coinvolto nel trasporto di elettroni attraverso il foto sistema II (PSII), che è quel complesso situato nei cloroplasti, che permette, attraverso un meccanismo di ossidazione dell' H₂O, di ricavare tutti gli elettroni necessari allo svolgimento dell'attività fotosintetica.

Mentre il *trnH* è il gene dell' RNA transfer dell'istidina (tRNA HIS), coinvolto nel meccanismo di trasferimento di questo amminoacido (HIS) durante la sintesi proteica.

Questo marcatore presenta punti di forza, quali:

- Una certa universalità (amplificazione e sequenziamento con successo su un gran numero di taxa)
- Un'alta variabilità (alta divergenza interspecifica)

Però di contro, ha anche punti di debolezza, quali:

- Presenza di un livello significativo di inserzioni, delezioni, frammenti ripetuti che si traducono in una difficoltà di allineamento
- Casi di presenza di doppia banda su gel

In definitiva il vantaggio dell'approccio *multilocus*, è quello di ottenere, dalla combinazione della variabilità espressa da ciascun marcatore, un'alta probabilità di successo in termini di identificazione univoca del maggior numero di taxa possibili.

3.2.3.3 CBOL

Come detto in precedenza attualmente sono attivi numerosissimi progetti sul DNA barcoding che riguardano sia il regno animale che vegetale, e che hanno messo a disposizione della comunità scientifica enormi quantità di dati.

Il successo dell'utilizzo di COI-5' come marker identificativo delle specie ha portato alla nascita del Consortium for the Barcode of Life (CBOL -www.barcodeoflife.org), che raccoglie 200 istituzioni di 50 paesi che utilizzano i codici a barre genetici come standard per l'identificazione delle specie. I membri del CBOL coordinano l'iniziativa International Barcode of Life (iBOL - www.ibolproject.org), il cui obiettivo primario è la costruzione di una libreria di sequenze DNA barcoding di riferimento, effettuando uno screening su larga scala di uno o pochi geni, al fine di assegnare esemplari sconosciuti alla rispettiva specie e facilitare la scoperta di nuovi taxa. CBOL inoltre convalida a livello internazionale i barcode ufficiali per i diversi taxa.

Attualmente l'iniziativa iBOL comprende diversi gruppi di lavoro, ognuno dei quali conduce progetti di ricerca sui seguenti gruppi tassonomici o raggruppamenti ecologici:

- WG 1.1 Vertebrati
- WG 1.2 Piante terrestri
- WG 1.3 Funghi
- WG 1.4 Animali Parassiti, Patogeni & Vettori
- WG 1.5 Parassiti Agricoli e Forestali e loro Parassitoidi
- WG 1.6 Impollinatori
- WG 1.7 Bio-sorveglianza d'Acqua dolce
- WG 1.8 Bio-sorveglianza Marina
- WG 1.9 Bio-sorveglianza Terrestre
- WG 1.10 Specie Polari
- WG 1.11 Biota Barcoding

Tutti i progetti iBOL si basano su tre principali componenti: i campioni, i laboratori di analisi e il database centrale.

I campioni possono provenire da varie fonti: alcuni sono raccolti in campo, altri provengono da collezioni museali, zoo, giardini botanici o banche del seme.

In laboratorio viene utilizzato un piccolo pezzo di campione per estrarre il DNA. La regione barcode viene isolata, amplificata tramite PCR e successivamente sequenziata. La sequenza genica risultante è rappresentata da una serie di lettere corrispondenti ai singoli nucleotidi: citosina, adenina, timina e guanina (C-A-T-G).

Uno dei più importanti componenti dell'iBOL è la costruzione di una libreria pubblica contenente le sequenze genomiche delle specie identificate. Attualmente esistono due importanti database di barcode che svolgono questo ruolo: il *Barcode of Life Data System* (BOLD), un database creato dall'Università di Guelph (Ontario, Canada), e l'*International Nucleotide Sequence Database Collaborative*, costituito dall'associazione di tre banche dati intercomunicati: *GenBank*, gestito dal National Center for Biotechnology Information (Stati Uniti), *Nucleotide Sequence Database*, del European Molecular Biology Lab (Europa) e *DNA Data Bank of Japan*, situato presso il National Institute of Genetics (Giappone). Tutte queste banche dati archiviano sequenze di barcode seguendo gli standard del Consortium for the Barcode of Life.

Ogni sequenza genica, ottenuta dai vari centri di ricerca, si riferisce ad un campione la cui raccolta è documentata da precise informazioni (luogo e data di campionamento, raccoglitore, ecc.) ed include un campione di riferimento.

L'esemplare, con la foto e le coordinate geografiche del sito di raccolta, accompagnano ogni sequenza depositata, tranne che in particolari circostanze (per esempio quando i dati di località potrebbero compromettere la sopravvivenza di una specie in pericolo di estinzione). L'elemento chiave della libreria barcode è il livello qualificato di identificazione tassonomica. Ogni esemplare viene infatti sottoposto ad un controllo incrociato con un elenco di nomi convalidati.

PROJECT NAME	TARGET ORGANISM	LOGO	PROJECT WEBSITE
Fish-BOL	Fishes		http://www.fishbol.org/
Lepidoptera Barcode of Life	Butterflies and Moths		http://lepbarcode.org/
Trichoptera Barcode of Life	Caddisflies		http://trichopterabol.org/
Mammalia Barcode of Life	Mammals		http://www.mammalibol.org/
Marine Barcode of Life	Marine fauna		http://www.marinebarcode.org/

Tab.2 Alcuni dei principali progetti sul DNA barcoding a livello mondiale

3.2.3.4 BOLD data system

Una volta ottenuta la sequenza genica, questa viene depositata sul Barcode of Life Data Systems (BOLD - www.barcodinglife.org).

Sebbene BOLD aiuti l'audace raccolta dei dati barcode e ne mantenga i record, una copia di tutti i dati di sequenza e dei campioni viene trasferita all'NCBI o ad altri database depositari (es. DNA Data Bank of Japan, DDBJ; European Molecular Biology Laboratory, EMBL) (Ratnasingham and Hebert, 2007). Alcune funzioni di BOLD, come l'accesso ai dati dei progetti pubblici e l'uso del sistema di identificazione, sono a disposizione di ogni visitatore.

Attualmente BOLD comprende più di 1.700.000 record barcode (<http://www.boldsystems.org>).

In BOLD, la procedura di deposito e pubblicazione dei dati di sequenza si svolge in due fasi. La prima fase comporta il rilascio di tutti i dati di sequenza generati ed il più alto livello di informazioni tassonomiche ottenute. La pubblicazione immediata di questi dati ha lo scopo di mettere a disposizione sufficienti informazioni alla comunità scientifica. La pubblicazione dei dati avviene in seguito a controlli di qualità computerizzati e la generazione automatica di un Barcode Index Numbers (BINs). I requisiti minimi per il rilascio iniziale dei dati comprendono:

- informazioni spaziali e temporali del sito di campionamento;
- informazioni tassonomiche del campione: l'identificazione tassonomica è necessaria per l'inizio del rilascio dei dati. Questa può essere fatta a livello di famiglia o può essere un nome descrittivo (ad esempio "campione ambientale"), e non è destinata ad essere una identificazione definitiva.

L'identificazione accurata al livello del genere e della specie è l'obiettivo finale, ma questo potrebbe non essere raggiunto per qualche tempo, pertanto il rilascio dei dati di sequenza non deve dipendere da una identificazione tassonomica definitiva;

- informazioni di sequenza: regione genomica sequenziata, metodo di assemblaggio delle sequenze, primers utilizzati, ecc.;
- identificatori di database: BOLD processa l'ID e ne identifica il campione (numero di voucher, nome del depositario e codice di raccolta).

La seconda fase comporta il rilascio di dati quando i lavori vengono sottomessi per la pubblicazione.



Fig.4 Home page del Barcode of Life Data Systems (<http://www.barcodinglife.org>)

3.2.4 PCR

La reazione a catena della polimerasi (Polymerase Chain Reaction) è una tecnica di biologia molecolare, che consente la moltiplicazione (amplificazione in vitro) di frammenti di acidi nucleici.

La PCR quindi ricostruisce in vitro uno specifico passaggio della duplicazione cellulare, ovvero la sintesi di un filamento di DNA a partire da un filamento stampo (complementare), mettendo in sequenza i nucleotidi sulla base dell'ordine del filamento complementare che funge da stampo.

Questo processo in natura viene svolto da enzimi chiamati "DNA polimerasi", tali enzimi sono in grado di allungare un filamento, avendo a disposizione dei nucleotidi liberi, ma non possono sintetizzarlo ex-novo.

Quindi con la tecnica PCR è possibile riprodurre in vitro la sintesi di un filamento di DNA a partire da un filamento stampo, ponendo in soluzione:

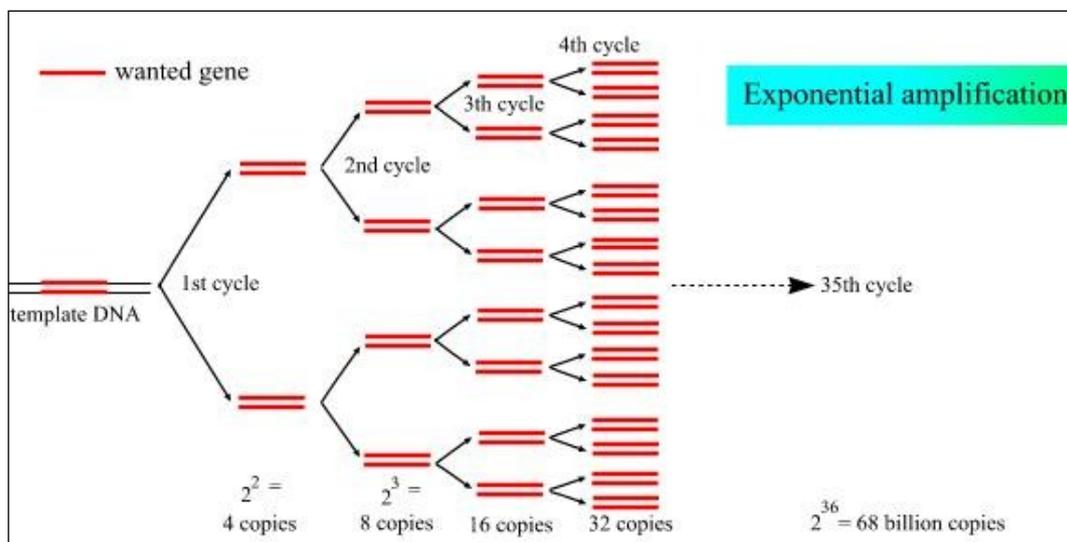
- Opportuni inneschi "primer" (infatti come già detto un filamento di DNA non può essere sintetizzato a partire da zero), che sono brevi sequenze di DNA complementari agli estremi dei filamenti da riprodurre.
- La sequenza di DNA da riprodurre.
- Una quantità opportuna di nucleotidi liberi per costruire i nuovi filamenti.
- Una DNA polimerasi termo-resistente, che è la Taq polimerasi, proveniente dal batterio termofilo "Thermophilus aquaticus".
- Un buffer che serve a mantenere il pH stabile (tampone), necessario a creare l'ambiente adatto alla reazione.
- MgCl₂ che è un indispensabile cofattore per il corretto funzionamento della DNA polimerasi.
- Acqua per portare a volume la soluzione.

L'amplificazione genica tramite PCR, si ottiene grazie ad una successione ciclica di più fasi a diverse temperature, ovvero: di una prima fase, ad elevata temperatura (92-98 °C), di denaturazione termica dei due filamenti del DNA bersaglio; di una seconda fase, a temperatura più bassa (40-55 °C), di appaiamento ("annealing") dei primer alle regioni complementari sulle singole eliche denaturate; di una terza fase a temperatura

ottimale per la taq polimerasi (65-72 °C), di sintesi ed estensione delle sequenze di DNA complementari a quelle bersaglio.

La successione ciclica di tali fasi deve ripetersi almeno 30-35 volte e si svolge all'interno di un termociclatore che opera la variazione di tale temperatura secondo un programma ben preciso, ed il numero di copie della sequenza bersaglio aumenta in maniera esponenziale ad ogni ciclo ($a=2^n$; dove "a" è il numero di copie della sequenza bersaglio ed "n" è il numero di cicli).

Fig.5 Rappresentazione schematica dei cicli di PCR



3.2.5 Sequenziamento

Per sequenziamento del DNA, si intende la determinazione dell'ordine esatto dei quattro nucleotidi che compongono il DNA stesso.

Il metodo più utilizzato per questa analisi è quello di Sanger (Sanger et al., 1977), detto anche metodo della terminazione della catena o dei didesossiribonucleotidi (ddNTP), utilizzando dei sequenziatori automatici.

Il principio del sequenziamento *Sanger*, è il seguente:

Si tratta di una metodica relativamente semplice che consiste nel far sintetizzare frammenti di catena polinucleotidica di lunghezza diversa sullo stampo del DNA che si vuole sequenziare (prodotto di PCR). Ciò si ottiene facendo avvenire la sintesi della nuova catena utilizzando oltre ai 4 nucleotidi dNTP (dATP, dGTP, dCTP, dTTP), anche uno dei 4 ddNTP (dideozinucleoside trifosfato), che altro non è che un nucleotide che non ha il gruppo $-OH$ in posizione 3' dello zucchero. In presenza di DNA polimerasi un ddNTP può essere incorporato all'estremità 3' di una catena nucleotidica in accrescimento su uno stampo di DNA, ma non può poi legare un altro nucleotide perchè non è disponibile l' OH in 3'.

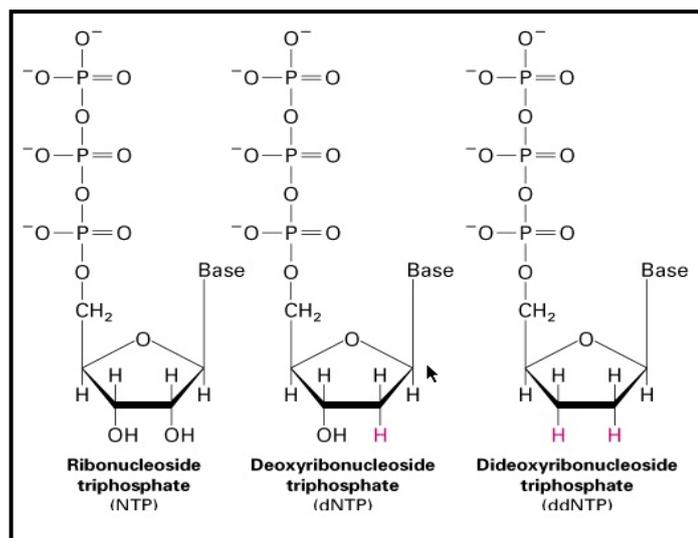


Fig.6 Rappresentazione di un ddNTP (a destra)

Per questo motivo la sintesi si arresta alla posizione in cui un ddNTP è stato incorporato all'estremità in accrescimento di un filamento di DNA (fase di terminazione).

In linea di principio nel sequenziamento manuale, si preparano pertanto nella pratica 4 miscele, contenente: il DNA da sequenziare; i nucleotidi liberi (dNTP); il primer utilizzato per l'amplificazione di PCR; la DNA polimerasi ed un ddNTP diverso (ddATP, ddGTP, ddCTP, ddTTP) per ognuna delle 4 frazioni e marcato con un fluorescente.

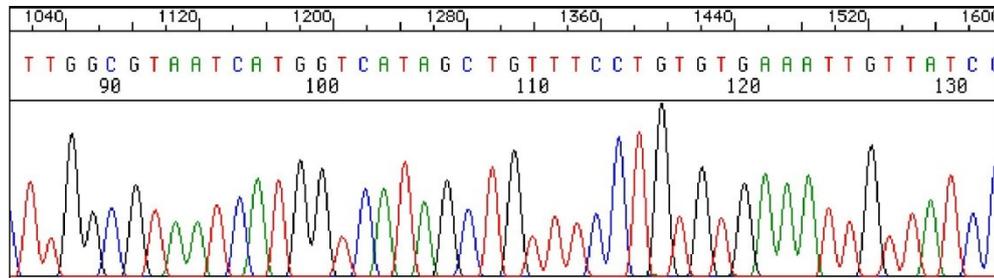
Poichè l'incorporazione del ddNTP nella catena in accrescimento è del tutto casuale, durante l'incubazione si formano in ciascuna frazione frammenti polinucleotidici di lunghezza diversa, aventi tutti come sequenza iniziale quella del primer, sequenza successiva in direzione 5'-3' complementare al segmento duplicato del DNA stampo, e tutti terminanti con il ddNTP (marcato) presente in quella frazione.

In seguito ad elettroforesi, le diverse catene polinucleotidiche neosintetizzate migreranno nel gel verso l'anodo in funzione della loro lunghezza e possono essere facilmente localizzate, poichè radioattive, per autoradiografia. Si possono così evidenziare centinaia di bande e separare catene che differiscono di un solo nucleotide. Nelle 4 corsie del gel le bande si disporranno in ordine di lunghezza dal fondo verso la zona di deposizione. Dalla successione di tutte le bande presenti nelle 4 corsie del gel si può risalire alla sequenza del frammento di DNA usato come stampo.

Nel sequenziamento automatico, invece non è necessario separare le quattro reazioni di terminazione in quattro provette differenti, ma si può allestire una singola reazione inserendo i quattro ddNTPs marcati mediante l'incorporazione di un composto fluorescente, diverso per ogni base, e l'elettroforesi avviene lungo dei capillari, all'interno dei quali è sempre ricostruita la rete di gel o in polimero. Durante la corsa elettroforetica, i frammenti vengono letti in ordine di lunghezza crescente da un fascio laser che eccita i marcatori fluorescenti. Quindi l'intensità della luce emessa viene misurata e tale informazione, cioè quale colore di fluorescenza e dunque il tipo di ddNTP è presente all'estremità di ogni filamento di differente lunghezza, viene inviata ad un computer.

Quindi le emissioni fluorescenti vengono captate da un rilevatore e trasformate in picchi di colore diverso (ogni colore corrisponde ad uno dei 4 ddNTPs presente all'estremità del filamento di lunghezza diversa), con aree proporzionali all'intensità di emissione, ed il tutto si traduce nella costruzione di un *elettroferogramma*.

Fig.7 Rappresentazione di un elettroferogramma



3.2.6 Filogenesi e “cluster analysis”

La **filogenesi** o **filogenetica** o **filogenia**, è il processo di ramificazione delle linee di discendenza nell'evoluzione della vita. La sua ricostruzione è fondamentale per la sistematica che si occupa di ricostruire le relazioni di parentela evolutiva, di gruppi tassonomici di organismi a qualunque livello sistematico.

Quindi ricostruire una filogenesi significa trovare la migliore stima delle relazioni evolutive storiche fra entità tassonomiche usando i dati disponibili.

Il diagramma che mostra i rapporti filogenetici tra specie o gruppi superiori si chiama albero filogenetico.

L'albero filogenetico è un diagramma che mostra le relazioni fondamentali di discendenza comune di gruppi tassonomici di organismi. Essi vengono costruiti sulla base di caratteri morfologici e molecolari (sequenze nucleotidiche o amminoacidiche).

Gli alberi filogenetici possono essere divisi in due grandi tipi: gli *alberi radicati* e gli *alberi non radicati*.

- Un albero *radicato* (in inglese *rooted*), è una struttura ad albero che si sviluppa a partire da un unico nodo, rappresentante il più recente antenato comune delle forme di vita che si trovano alle estremità dell'albero. In questo modo, un albero filogenetico radicato è in grado di fornire informazioni sia sulla correlazione genetica esistente tra gli organismi presenti sulle sue ramificazioni, sia sui rapporti evolutivi che intercorrono tra gli stessi.

- Un albero *non radicato*, al contrario, illustra le relazioni genetiche che intercorrono tra gli organismi che si trovano ai suoi apici, ma non fornisce alcuna informazione in merito alla loro evoluzione.

Quindi un albero filogenetico può fornire anche una rappresentazione, non necessariamente evoluzionistica, delle relazioni esistenti fra le specie (o altre categorie tassonomiche, popolazioni, gruppi sistematici, sequenze di DNA, ecc.) a seconda dei caratteri che si considerano.

Le due metodologie più comunemente adoperate per ricostruire un albero filogenetico sono la fenetica e la cladistica.

- In biologia la *fenetica*, talvolta indicata come tassonomia numerica o tassimetria (Sneath and Sokal 1973) è una tecnica la cui finalità è la classificazione degli organismi sulla base della somiglianza dei caratteri presi in considerazione senza tener conto della filogenia o delle relazioni evolutive.
- La cladistica, ricostruisce la filogenia classificando gli organismi viventi in base a criteri evolutivi. Il padre fondatore di questa disciplina fu l'entomologo tedesco Willi Hennig, che la battezzò sistematica filogenetica.

Nel primo caso le genealogie sono costruite sulla base di indici di distanza, mentre nel secondo caso sulla della condivisione di caratteri

Le rappresentazioni più comuni degli alberi filogenetici, sono:

➤ I Dendrogrammi, sono alberi filogenetici non radicati, sono rappresentazioni grafiche dei coefficienti di similarità.

Con il diffondersi delle indagini molecolari basate sull'analisi di dati di sequenze nucleotidiche, la costruzione dei dendrogrammi è diventata fondamentale per chi si occupa di biologia molecolare, basti pensare al fatto che nel confrontare gruppi di individui molto diversi tra di loro, è molto più facile andare a trovare segmenti di DNA omologhi, piuttosto che caratteri morfologici comparabili.

I Dendrogrammi prodotti a partire da sequenze nucleotidiche hanno preso il nome di alberi filogenetici molecolari.

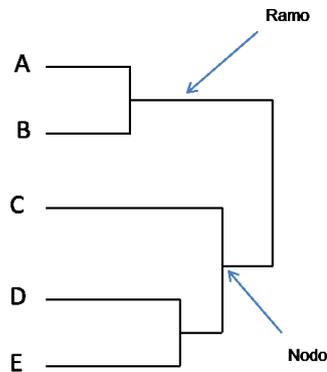


Fig.8 Esempio di rappresentazione di un Dendrogramma

➤ I Cladogrammi, sono alberi filogenetici costruiti sulla base di metodi cladistici, sono delle rappresentazioni non metriche, dove la lunghezza delle branche non ha significato ed illustra solo le relazioni di parentela tra le specie.

I Dendrogrammi nascono e vengono utilizzati all'interno dei "cluster analysis" (analisi dei gruppi), è una tecnica di analisi multivariata dei dati, volta al raggruppamento di unità statistiche in sottoinsiemi (gruppi), le cui unità sono mediamente più simili fra loro di quanto non lo sia ciascuna di esse ad ogni unità degli altri gruppi.

I metodi per la costruzione dei Dendrogrammi si possono suddividere in due grandi gruppi:

➤ **Metodi basati sulle matrici di distanza**, che tengono conto della distanza evolutiva tra specie e specie.

I metodi basati sul calcolo della distanza si compongono generalmente da due fasi: una prima fase dove sulla base dei caratteri presi in esame (set di dati), viene elaborata con l'ausilio di modelli ben determinati una matrice delle distanze tra le specie considerate, se ad esempio il set di dati presi in esame, sono un insieme di sequenze nucleotidiche (ciascuna relativa ad una specie presa in esame), sulla base dell'allineamento delle sequenze fornite, verrà calcolata una matrice della distanza evolutiva tra tutte le coppie di specie considerate; ed una seconda fase dove viene costruito un albero filogenetico che rispetti la matrice delle distanze ottenuta precedentemente.

Tra i metodi che sfruttano la matrice di distanza ricordiamo:

- **UPGMA**, (unweighted pair group method with arithmetic mean), è il metodo più semplice per la costruzione dei Dendrogrammi. Questo metodo consiste nel prendere dalla matrice delle distanze la coppia di taxa che presentano la maggiore similarità (cioè i due taxa che presentano la minore distanza), questi taxa vengono collegati tra loro nell'albero e vengono considerate come un'unica nuovo taxa. Vengono, quindi ricalcolate le distanze genetiche con gli altri taxa, ricostruendo la matrice delle distanze che avrà in questo modo un elemento in meno, quindi verrà nuovamente unita assieme (a formare un'unica entità) la coppia di taxa più affini, e si procede più volte in questo modo, fin quando non saranno collegati tutti i taxa in questione all'albero.
- **Neighbor Joining**, l'approccio è simile a quello dell'UPGMA, ma la ricerca del "neighbor" (il "vicino") di ciascuna specie viene effettuata non scegliendo la coppia con distanza minima, ma attraverso un processo di minimizzazione della somma di tutte le distanze fra le specie, formatasi la prima coppia che verrà considerata un'unica entità, l'algoritmo itererà l'operazione fino ad ottenere l'albero cercato.

Per i metodi basati sulla distanza sono pertanto necessari dei modelli per la costruzione delle matrici di distanza, avendo a disposizione un set di dati (caratteri considerati) per i taxa oggetto di studio. Ci sono programmi che sfruttano vari modelli di calcolo, che a seconda del numero di parametri utilizzati sono via via più complessi e forniscono dei risultati sempre più attendibili, nel caso specifico di sequenze nucleotidiche come caratteri considerati, tra i modelli più semplici ricordiamo, il modello di Jukes-Cantor (1969), che tiene conto di un solo parametro, secondo il quale tutte le sostituzioni di basi hanno la stessa probabilità di verificarsi; Kimura (1980) che è un modello di matrice che tiene conto di due parametri, le *transizioni* (ovvero le mutazioni che trasformano una purina in un'altra purina o una pirimidina in un'altra pirimidina, cioè $A \leftrightarrow G$ e $C \leftrightarrow T$) e la *trasversioni* (ovvero le mutazioni che trasformano le purine in pirimidine e viceversa, cioè $A \leftrightarrow C$, $A \leftrightarrow T$, $C \leftrightarrow G$, $C \leftrightarrow T$), transizioni e trasversioni hanno probabilità differenti di verificarsi; fino ad arrivare ai modelli più complessi che tengono in considerazione molti parametri, anche 9 (Lanave et al. 1984) (Saccone et al. 1990)

➤ **Metodi basati sul criterio di ottimizzazione**, si basano sull'applicazione della definizione degli alberi migliori, ciò presuppone la scelta di un criterio di ottimizzazione per trovare l'albero migliore, tra questi ricordiamo:

- **Metodo di massima parsimonia**, si basano sulla ricerca del minor numero di mutazioni possibili tra i nodi dell'albero, e quindi il minor numero di cambiamenti evolutivi. Lo spazio delle soluzioni, ovvero di tutti gli alberi filogenetici pesati, in cui trovare l'albero ottimale, è potenzialmente molto grande.

Il metodo presuppone la scelta dei siti informativi su cui effettuare l'analisi ovvero siti polimorfici dove sono presenti almeno due stati di carattere in almeno due individui. E poi con l'ausilio di algoritmi viene trovato l'albero che richiede il più piccolo numero di cambiamenti, che è quello che spiega meglio i dati.

Quindi è un metodo che non tiene conto delle distanze risultanti da apposite metriche di distanza, ma tiene conto dello stato dei caratteri, cioè della condivisione dei caratteri tra i taxa considerati.

Il metodo di massima parsimonia, nonostante sia abbastanza valido, ha un costo computazionale elevato. Questo spinge molti ricercatori a preferire metodi sulla distanza.

- **Metodo di Massima verosimiglianza (Maximum likelihood)**, si basa sulla ricerca della topologia dell'albero e la relativa lunghezza dei rami che massimizzano la verosimiglianza dei dati osservati con il modello considerato.

Quindi questo metodo parte da un'ipotesi di modello di evoluzione e su questo viene stimato l'albero migliore, che è quello che ha la massima probabilità di essere generato sulla base del modello ipotizzato.

Quindi con l'ausilio di programmi specifici (molto utilizzato per questo tipo di analisi è il modello PHYLIP), il metodo cerca di quantificare la probabilità che ad una certa ipotesi "X", nel nostro caso il modello di albero filogenetico, corrisponda un set di dati a disposizione, ad esempio sequenze nucleotidiche. L'albero che ottiene il massimo valore di probabilità rappresenta la stima di massima verosimiglianza della filogenesi tra le sequenze considerate

È comparativamente il metodo più lento e computazionalmente intensivo, ed è quindi limitato nel numero di specie che può analizzare, ma è considerato accurato e informativo

PARTE SPERIMENTALE

4 MATERIALI E METODI

Per la caratterizzazione ed identificazione a livello genotipico, si sono seguiti i protocolli del “DNA barcoding” suggeriti dal CBOL, Consortium for the Barcode of Life (<http://www.barcodeoflife.org/>), che prevede l’amplificazione di porzioni geniche note, e che soprattutto per le specie spontanee, rappresenta uno strumento unico ed effettivo a disposizione per l’identificazione certa di specie con variabile affinità genetica. L’amplificazione di porzioni note di genoma, e successivo sequenziamento, è l’unica tecnica in grado di fornire risultati univoci, rapidi e precisi, in tempi brevi e con costi adeguati su un ampio numero di specie.

L’intero processo di messa a punto dei protocolli di caratterizzazione molecolare attraverso la metodica del *DNA barcoding* si è articolato in diverse fasi, di seguito descritte.

4.1 Selezione del materiale vegetale

Prima di tutto è stata effettuata una selezione del materiale vegetale su cui condurre le indagini molecolari, il materiale vegetale oggetto di studio è elencato nelle tabelle 3; 4; 5.

Nello specifico si tratta di materiale vegetale già in produzione presso l'azienda vivaistica Faro, sita a Carruba di Giarre (CT), che figura come impresa proponente nell'ambito del progetto PON indicato in precedenza, e di piante autoctone, di interesse ornamentale, prelevate in natura, con possibile impatto commerciale, oggetto di nuova introduzione ed in collezione sia presso la medesima azienda vivaistica Faro sia presso il CRA-SFM di Bagheria, che figura come partner del progetto PON all'interno del quale è stato sviluppata l'attività di dottorato in questione.

Per le specie autoctone di nuova introduzione, prima di procedere all'applicazione dei protocolli di caratterizzazione molecolare, è stato necessario, effettuare preliminarmente, un' accurata caratterizzazione a livello morfologico, poiché come già detto nei capitoli precedenti di questo elaborato, l'identificazione molecolare non deve assolutamente sostituire quella fenotipica ma essere da supporto a quest'ultima nelle fasi di caratterizzazione dei taxa a qualsiasi livello.

Il processo di riconoscimento morfologico è stato condotto da esperti botanici sia dell'università di Catania per quanto riguarda le specie autoctone di nuova introduzione ed in collezione presso l'Azienda Faro, sia dell'università di Palermo per quanto riguarda invece le piante autoctone di nuova introduzione ed in collezione presso il CRA-SFM.

Tab.3 Germoplasma di piante autoctone ed esotiche già in produzione presso l'azienda vivaistica Faro

Famiglia	Specie	ID
Acanthaceae	<i>Acanthus mollis</i> L.	F.A12
Arecaceae	<i>Acoelorrhaphe wrightii</i> H.Wendl. ex Becc.	F.A11
	<i>Arenga engleri</i> Becc.	F.A7
	<i>Caryota urens</i> L.	F.C20
	<i>Chamaerops humilis</i> var. <i>humilis</i> / <i>Chamaerops humilis</i> L.	F.C4
	<i>Chamaerops humilis</i> var. <i>argentea</i> André	F.C5
	<i>Chamaerops humilis</i> L. "Vulcano"	F.C6
	<i>Chamaerops humilis</i> L. "Etna star"	F.C7
	<i>Howea forsteriana</i> (F.Muell.) Becc.	F.H1
	<i>Livistona chinensis</i> (Jacq.) R.Br. ex Mart.	F.L1
	<i>Phoenix canariensis</i> Chabaud	F.P2
	<i>Phoenix dactylifera</i> L.	F.P8
	<i>Phoenix reclinata</i> Jacq.	F.P9
	<i>Phoenix roebelenii</i> O'Brien	F.P7
	<i>Sabal minor</i> (Jacq.) Pers.	F.S14
	<i>Sabal palmetto</i> (Walter) Lodd. ex Schult. & Schult.f.	F.S3
	<i>Trachycarpus fortunei</i> (Hook.) H.Wendl.	F.T1
	<i>Washingtonia robusta</i> H.Wendl.	F.W1
	<i>Washingtonia filifera</i> (Linden ex André) H.Wendl. ex de Bary	F.W2
	<i>Butia capitata</i> (Mart.) Beccari	F.B2
	<i>Bismarckia nobilis</i> Hildebr. & H.Wendl.	F.B3
<i>Brahea armata</i> S. Watson	F.B4	
<i>Brahea edulis</i> H.Wendl. ex S.Watson	F.B5	
<i>Trithrinax campestris</i> (Burmest.) Drude & Griseb.	F.T4	
<i>Arecastrum romanzoffianum</i> (Cham.) Becc../ <i>Syagrus romanzoffiana</i> (Cham.) Glassman	F.A8	
Xanthorrhoeaceae	<i>Aloe arborescens</i> Mill.	FA10
	<i>Aloe vera</i> (L.) Burm.f.	FA9
	<i>Aloe plicatilis</i> (L.) Mill..	FA5
	<i>Aloe</i> × <i>spinosissima</i> Jahand.	FA17

Fabaceae	<i>Spartium junceum</i> L.	F.S9
	<i>Ceratonia siliqua</i> L.	F.C1
Cistaceae	<i>Cistus albidus</i> L.	F.C11
	<i>Cistus salvifolius</i> L.	F.C2
	<i>Cistus x pulverulentus</i> Pourr.	F.C3
	<i>Cistus x skanbergii</i> Lojac.	F.C12
Cycadaceae	<i>Cycas circinalis</i> L.	F.C19
	<i>Cycas revoluta</i> Thunb.	F.C8
Myrtaceae	<i>Myrtus luma</i> Molina	FM1
	<i>Metrosideros excelsa</i> Sol. ex Gaertn.	F.M2
	<i>Myrtus communis</i> L.	F.M5
Lamiaceae	<i>Rosmarinus officinalis</i> L.	F.R4
	<i>Salvia leucantha</i> Cav.	F.S12
	<i>Lavandula angustifolia</i> Mill.	F.L4
	<i>Lavandula stoechas</i> L.	F.L5
	<i>Salvia officinalis</i> L.	F.S11
Ericaceae	<i>Arbutus unedo</i> L.	F.A1
Asteraceae	<i>Helichrysum italicum</i> (Roth) G. Don	F.H2
Iridaceae	<i>Iris germanica</i> L.	F.I5
Strelitziaceae	<i>Strelitzia augusta</i> Thunb	F.S1
	<i>Strelitzia nicolai</i> Regel & K.Koch	F.S2
	<i>Strelitzia reginae</i> Banks	F.S8
Tamaricaceae	<i>Tamarix gallica</i> L.	F.T5

Tab.4 Germoplasma di specie oggetto di nuova introduzione, in collezione presso l'azienda Faro

Famiglia	Specie	ID
Lamiaceae	<i>Sideritis italica</i> (Mill.) Greuter & Burdet	F.S6
Asteraceae	<i>Anthemis cupaniana</i> Tod. ex Nyman	F.A2
	<i>Centaurea sphaerocephala</i> L.	F.C14
	<i>Jacobaea gibbosa</i> (Guss.) B.Nord. & Greuter	F.J1
	<i>Pallenis maritima</i> (L.) Greuter	F.P3
	<i>Ptilostemon greuteri</i> Raimondo & Domina	F.P5
	<i>Senecio candidus</i> (Presl.) DC. / <i>Jacobaea candida</i> (C.Presl) B.Nord. & Greuter	F.S5
	<i>Jacobaea ambigua</i> (Biv.) Pelser & Veldkamp	F.S4
	<i>Anthemis maritima</i> L.	F.A4
Convolvulaceae	<i>Calystegia soldanella</i> (L.) R. Br.	F.C13
Iridaceae	<i>Iris pseudopumila</i> Tineo	F.I1
Amaranthaceae	<i>Diotis maritima</i> (L.) Desf. ex Cass./ <i>Achillea maritima</i> (L.) Ehrend. & YPGuo	F.D1

Tab.5 Germoplasma di specie oggetto di nuova introduzione, in collezione presso il CRA-SFM

Famiglia	Specie	ID
Liliaceae	<i>Tulipa raddii</i> Reboul	T4
Fabaceae	<i>Genista madoniensis</i> Raimondo	G2
	<i>Genista demarcoi</i> Brullo, Scelsi & Siracusa	G4
	<i>Genista tyrrhena</i> Vals.	G1
	<i>Genista cupanii</i> Guss.	G5
	<i>Genista aetnensis</i> (Biv.) DC.	G6
	<i>Genista aristata</i> C.Presl	G3
Ericaceae	<i>Erica sicula</i> Guss.	E7
	<i>Erica peduncularis</i> C.Presl	E9
	<i>Erica multiflora</i> L.	E10
Asteraceae	<i>Helichrysum hyblaenum</i> Brullo	H6
	<i>Helichrysum nebrodense</i> Heldr.	H5
	<i>Helichrysum scandens</i> Guss.	H7
	<i>Hieracium cophanense</i> Lojac.	H2
Brassicaceae	<i>Brassica insularis</i> Moris	B2
	<i>Brassica villosa</i> subsp. <i>tinei</i> (Lojac.) Raimondo & Mazzola	B4
	<i>Brassica rupestris</i> subsp. <i>hispidula</i> Raimondo & Mazzola	B5
Rosaceae	<i>Rosa sicula</i> Tratt.	R6
	<i>Rosa sempervirens</i> L.	R3
	<i>Rosa canina</i> L.	R1
	<i>Rosa corymbifera</i> Borkh.	R10
Caryophyllaceae	<i>Dianthus busambrae</i> Soldano & F. Conti	D2
	<i>Dianthus rupicola</i> subsp. <i>aeolicus</i> (Lojac.) Brullo & Miniss.	D1
	<i>Dianthus rupicola</i> Biv. subsp. <i>rupicola</i>	D3
	<i>Dianthus rupicola</i> subsp. <i>lopadusanum</i> Brullo & Miniss.	D4
	<i>Dianthus siculus</i> C. Presl	D5
	<i>Dianthus rupicola</i> subsp. <i>hermaensis</i> (Coss.) O. Bolòs & Vigo	D6

Euphorbiaceae	<i>Euphorbia ceratocarpa</i> Ten.	Eu1
	<i>Euphorbia characias</i> L.	E2
	<i>Euphorbia dendroides</i> L.	E4
	<i>Euphorbia meuselii</i> Geltman	E5
	<i>Euphorbia myrsinites</i> L.	E6
	<i>Euphorbia helioscopia</i> L.	Eu7
	<i>Euphorbia bivonae</i> Steud.	Eu8
	<i>Euphorbia pithyusa</i> subsp. <i>cupanii</i> (Guss. ex Bertol.) Radcl.-Sm.	Eu9
	<i>Euphorbia amygdaloides</i> L.	Eu11

Il materiale vegetale oggetto di studio è rappresentato da 3 replicati per ciascuna specie (ovvero tre piante distinte per ciascuna specie), nello specifico da ciascuna di esse sono stati prelevati campioni di tessuti, rappresentati da giovane foglie, che costituiscono il materiale vegetale da cui procedere con l'estrazione del DNA.

Ogni singolo campione prelevato è stato catalogato con un codice identificativo (ID) in modo tale da allestire una vera e propria collezione del germoplasma, oltre che per la costituzione di una banca del DNA del germoplasma.



Fig.9 prelievo di alcuni individui in produzione presso l'azienda Faro

4.2 Conservazione del materiale vegetale

La prima operazione eseguita quando i campioni vegetali arrivano in laboratorio, è la liofilizzazione delle giovani foglie che rappresentano il materiale vegetale di partenza da cui effettuare l'estrazione del DNA.

La liofilizzazione è un processo molto importante che consente una lunga conservabilità del materiale vegetale, anche a temperatura ambiente, senza andare incontro a degradazione. E' preferita rispetto alla disidratazione convenzionale, tramite gel di silice o stufe, in quanto non provoca modificazioni irreversibili della struttura del DNA (distorsioni, deterioramenti o rotture). Il materiale da essiccare viene congelato (crio-essiccazione, da fase liquida a solida) e poi per effetto dell'applicazione di un vuoto molto spinto la fase solida passa direttamente a quella di vapore (sublimazione). I campioni di foglie rimangono nel liofilizzatore per circa 48 ore, quindi una parte del materiale liofilizzato è stato utilizzato per le estrazioni di DNA, la restante parte viene

conservata come stock, a cui potere attingere per eventuali altre estrazioni. Tali stock costituiscono una vera e propria banca del germoplasma a tempo indeterminato, da cui originare la rispettiva banca del DNA.

4.3. Estrazione del DNA da materiale vegetale

Una fase fondamentale per la buona riuscita di un saggio molecolare è senza dubbio la fase di estrazione del DNA da materiale vegetale. Lo scopo è quello di ottenere campioni di DNA con una buona concentrazione ed un elevato grado di purezza (assenza di altri composti inibenti come proteine, polifenoli, acidi umici, etc.) che possono influire negativamente sulle successive analisi.

Pertanto in tale ottica si è focalizzata l'attenzione sulla messa a punto di protocolli di estrazione di DNA che permettano di ottenere con i minori costi possibili ed in tempi rapidi, campioni di DNA con le caratteristiche su menzionate. Inoltre si è cercato di ridurre o eliminare la presenza di reagenti tossici.

La fase di estrazione del DNA ha riguardato tutti e tre i campioni vegetali provenienti da tre diversi individui per singola specie.

Per la fase di estrazione del DNA non è stato impiegato alcuna tipologia di kit commerciali, ma è stato messo a punto ed adottato un protocollo completamente manuale che rispetto ai classici kit per estrazione presenti in commercio, risulta sicuramente più economico, e soprattutto ha permesso di ottenere quantitativi di DNA

molto più alti per ogni singola estrazione con un buon livello qualitativo in termini di purezza (assenza di altri composti inibenti come proteine, polifenoli, acidi umici, etc... che possono influire negativamente sulle successive analisi), condizioni essenziali per l'ottenimento di ottime sequenze di DNA ai fini della identificazione molecolare delle specie oggetto di studio, arrivando così alla messa a punto di un protocollo di estrazione che ha permesso di ottenere un ottimo DNA in termini di concentrazione e purezza con costi molto limitati, riducendo nel contempo il più possibile l'impiego di reagenti tossici per la salute umana.

Come materiale di partenza per l'estrazione sono state utilizzate porzioni giovani di foglie, che sono state, come visto, precedentemente sottoposte a liofilizzazione.

Il protocollo di estrazione del DNA prevede cinque fasi fondamentali, che sono: la rottura delle pareti cellulari e lisi delle membrane cellulari, la purificazione del DNA per separare gli acidi nucleici da altri componenti cellulari; la precipitazione del DNA in alcool; il lavaggio del DNA e la successiva sospensione del DNA.

Nello specifico si è fatto riferimento al metodo basato sull'utilizzo del CTAB (esadecil-trimetil-ammonio-bromuro) (Doyle & Doyle, 1987) un detergente ionico che funziona da "anion binding", ovvero le molecole di DNA (che hanno carica negativa) si vanno a legare al CTAB formando un complesso CTAB-DNA che può precipitare; favorendo la sua separazione dai prodotti di scarto.

Il Buffer di estrazione CTAB, messo a punto, oltre a contenere il principio CTAB contiene altri reagenti ognuno dei quali svolge un'azione importante nella estrazione del DNA, quali l'EDTA (acido etilendiamminotetracetico), che è un agente chelante che sequestra i cationi necessari per la stabilizzazione delle membrane cellulari inoltre va ad inibire l'enzima DNasi che idrolizza i legami fosfodiesterici della catena del DNA, danneggiandolo, altro componente importante del buffer è il NaCl, un sale che va a saturare le cariche negative del DNA consentendogli successivamente di precipitare in alcool, e poi il Tris-HCl un composto che funge da tampone che mantiene il pH intorno alla neutralità o leggermente alcalino, condizione fondamentale per la successiva fase di purificazione.

Le fasi in dettaglio:

-Lisi cellulare: l'azione chimica del buffer di estrazione CTAB è stata coadiuvata dall'azione meccanica di un omogeneizzatore, pertanto a 0.04 gr. di materiale vegetale liofilizzato, posti in un tubo da microcentrifuga della capacità di 2,0 ml, sono stati

aggiunti 800 µl di buffer CTAB, biglie di acciaio del diametro di 3mm e piccolissime biglie di vetro del diametro di 1mm.

I Tubi, contenenti il materiale vegetale con il buffer e le biglie, sono stati posti all'interno di un omogeneizzatore programmato per svolgere 3 cicli da 1 minuto ciascuno.

Successivamente i tubini sono stati posti all'interno di un termo block a 65°C per circa 1 ora.

I campioni sono stati successivamente centrifugati a circa 12000 rpm per 5 minuti, per facilitare la separazione della fase liquida da quella solida e con l'ausilio di una pipetta è stata prelevata la sole fase liquida contenete il DNA e posta in un altro tubo da 2ml.

-Purificazione del DNA: sono stati aggiunti alla fase liquida ottenuta in precedenza 700-800 µl di cloroformio (in funzione del volume della fase liquida recuperata in precedenza), il quale completa la denaturazione delle proteine, rimuove i lipidi solubilizzandoli e soprattutto grazie alla sua elevata densità facilita la separazione della fase acquosa (contenente il DNA deproteinizzato) da quella organica, stabilizzando l'interfaccia tra le due fasi, quindi esso, in seguito a centrifugazione a 12000rpm per 5 minuti, permette l'ottenimento di due fasi liquide, una superiore contenente gli acidi nucleici ed una inferiore contenente altri componenti di scarto, separate da una interfaccia solida, rappresentata dalle proteine denaturate.

Dopodiché viene prelevato il surnatante contenente il DNA, e posto in un tubino della capacità di 1.5ml.

-Precipitazione del DNA: sono stati aggiunti al surnatante recuperato nella fase precedente, 700-800 µl di isopropanolo, il DNA è stato lasciato precipitare in alcool per circa 24 ore a -20°C, una volta che il DNA è precipitato, in seguito a centrifugazione a 12000 rpm per circa 20 minuti, si elimina la parte liquida (alcool) e si recupera il pellet di DNA solido (rimasto attaccato sul fondo del tubicino).

-Lavaggio del DNA: ha lo scopo di rimuovere i sali precipitati, in questa fase, sono stati aggiunti al pellet 600 µl di alcool etilico al 70%, ed il tutto è stato nuovamente centrifugato a 12000 rpm per altri 5 minuti, e quindi eliminato nuovamente l'alcool.

-Sospensione del DNA: il DNA è stato sospeso in 50 µl di HPLC

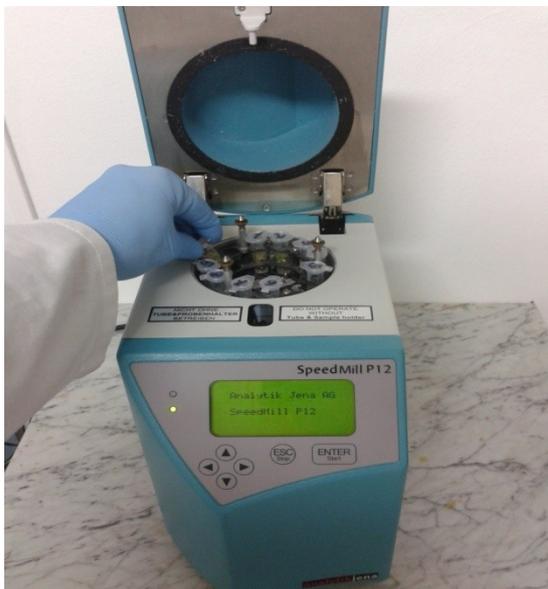


Fig.10 Omogeneizzatore *SpeedMill P12*



Fig.11 Thermo Blok 780

4.4. Valutazione quantitativa e qualitativa del DNA estratto

Una volta che il DNA si è risospeso, prima di procedere alla successiva fase di amplificazione, si effettua una misurazione in termini quantitativi e qualitativi del DNA estratto, in modo tale da essere certi che le successive fasi che caratterizzano il percorso di identificazione molecolare non siano quanto meno compromesse da una cattiva estrazione del DNA.

Questa operazione viene svolta mediante spettrofotometro. Il metodo spettrofotometrico sfrutta la capacità degli acidi nucleici di assorbire la luce a determinate lunghezze d'onda, quindi tale strumento ci da la misura della concentrazione del campione di DNA sulla base dell'assorbimento di luce da parte del campione, ma ci da anche importanti indicazioni su quella che è la purezza del nostro campione attraverso due rapporti:

-**Rapporto A260/A280**, che è l'indice della contaminazione da proteine, per il DNA un rapporto ottimale è intorno a 1,6-1,8 (rapporti superiori indicano contaminazioni da proteine).

-**Rapporto A260/A230**, che è l'indice della contaminazione da carboidrati e fenoli (solventi), un rapporto ottimale è intorno a 2 (rapporti inferiori indicano contaminazioni da solventi).



Fig.12. Spettrofotometro

4.5. Amplificazione tramite PCR

Una volta certi della buona qualità e quantità del DNA estratto, si è passati alla successiva fase di amplificazione del DNA, tramite PCR convenzionale.

Come per la fase di estrazione, anche l'amplificazione ha interessato tutti e tre i campioni vegetali provenienti da tre diversi individui per singola specie, grazie al successo di estrazione di DNA avuto per tutti i singoli campioni di ogni specie oggetto di studio.

Tra i parametri della PCR su cui si è molto lavorato, oltre alla quantità dei singoli componenti facenti parte del mix di reagenti necessari per il processo di amplificazione da porre in soluzione assieme al DNA stampo da riprodurre, vi è sicuramente la temperatura di "annealing" (appaiamento) che è stata valutata e di volta in volta modificata (entro un certo range) per massimizzare l'efficienza di appaiamento tra i primer ed il DNA stampo.

Lavorando su tali parametri si è riusciti a mettere a punto un protocollo di amplificazione (oltre a quello di estrazione) valido e ripetibile, quindi in definitiva un protocollo unico, con delle piccole varianti solo per quanto riguarda la temperatura di annealing in funzione della tipologia di primer utilizzato. In questo modo è stato possibile incrementare il numero dei campioni positivi alle analisi molecolari.

Per l'identificazione molecolare delle piante, sono stati utilizzati per i protocolli di amplificazione, tre regioni del DNA, promosse dal Consortium for the Barcode of Life (CBOL), ovvero *rbcL* (Ribulose-1,5-bisphosphate carboxylase/oxygenase), *matK* (Maturase K) e *trnH-psbA* (Chloroplast intergenic region), in modo tale da avere dalla combinazione della variabilità espressa da ciascun marcatore, un'alta probabilità di successo in termini di identificazione univoca del maggior numero di taxa possibili.

In prima istanza è stata valutata, come da preferenza dell'ampio gruppo di lavoro CBOL, per ciascun taxa esaminato la combinazione dei due "core markers" *rbcL+matK*, a cui è stato affiancato nei casi di insuccesso di discriminazione a livello di specie, il terzo marcatore barcode *trnH-psbA* (flowchart fig 13).

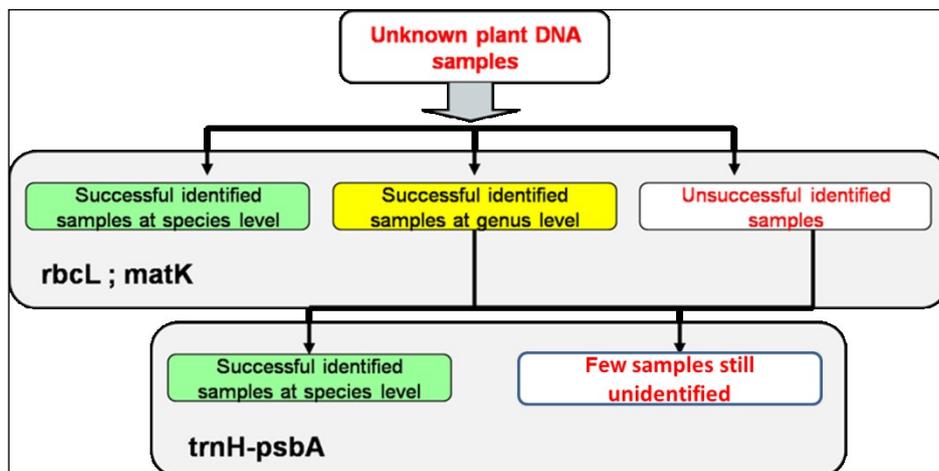


Fig.13 Il flowchart riassume i passi per l'identificazione dei campioni vegetali con l'adozione dei marcatori Barcoding selezionati.

I primers *rbcL*, *matK* e *trnH-psbA* impiegati nel processo di amplificazione tramite PCR e successivo sequenziamento degli amplificati, hanno la seguente sequenza nucleotidica:

- *rbcL*-F ATGTCACCACAAACAGAGACTAAAGC (Levin et al., 2003)
- *rbcL*-R GTAAAATCAAGTCCACCRG (Kress & Erickson, 2007)
- *matK*-3F KIM CGTACAGTACTTTTGTGTTTACGAG (K. J. Kim, pers. com.)
- *matK*-1R KIM ACCCAGTCCATCTGGAAATCTTGGTTC (K. J. Kim, pers. com.)
- *matK* 390 CGATCTATTCATTCAATATTTTC (Cuenoud et al. 2002)
- *matK* MALV R1 TAATGAGAAAGATTTCTGCATAT (Dunning and Savolainen 2010)
- *trnHf_05* CGCGCATGGTGGATTCACAATCC (Tate & Simpson, 2003)
- *psbA3_f* GTTATGCATGAACGTAATGCTC (Sang et al. 1997)

Per le condizioni di PCR si sono seguite le linee suggerite dal CBOL Plant Working Group (Hollingsworth et al., 2009), e le amplificazioni sono state condotte con termociclatore *Gene®Amp PCR System 9700* (Applied Biosystems).

Nelle tabelle seguenti, si riporta il protocollo di amplificazione impiegato tramite PCR convenzionale. Il DNA è stato usato ad una concentrazione di 20ng.

Bf	<i>1X</i>
MgCl₂	<i>1.5mM</i>
dNTP	<i>0.2mM</i>
BSA	<i>0.1mg/ml</i>
Fw	<i>1μM</i>
Rv	<i>1μM</i>
Taq	<i>2u</i>

Tab.6 Mix di reazione per PCR

94 °C x 5 min	
94 °C x 30 sec	
54 °C x 1 min	35 cicli
72 °C x 50 sec	
72 °C x 10 min	

94 °C x 5 min	
94 °C x 30 sec	
53 °C x 30 sec	35 cicli
72 °C x 50 sec	
72 °C x 7 min	

95 °C x 5 min	
94 °C x 30 sec	
55 °C x 1 min	35 cicli
72 °C x 50 sec	
72 °C x 10 min	

Tab.7 Progr. termoc. rbcl

Tab.8 Progr. termoc. matK

Tab.9 Progr. termoc. trnH-psbA



Fig.14 Termociclatore *Gene®Amp PCR System 9700* (Applied Biosystems).

4.6. Elettroforesi

Gli amplificati tramite PCR, sono stati fatti correre su gel di agarosio al 2%, il cui scopo è quello di accertarsi della buona riuscita dell'amplificazione dei segmenti di DNA interessati, tramite i primers *barcoding* utilizzati, mediante l'aggiunta al gel di un colorante fluorescente che si intercala tra le bande del DNA e rende possibile la comparsa delle bande di DNA (che altro non sono che le sequenze nucleotidiche riprodotte in vitro tramite amplificazione, oggetto di studio ai fini della identificazione genetica) su gel di agarosio, a tale scopo è stato utilizzato come colorante il "SYBR safe" in alternativa al bromuro di etidio, molto più rischioso per la salute umana.

I gel ottenuti mediante corsa elettroforetica sono stato visualizzati, mediante un acquirente di immagini "Gel Doc" della BIO-RAD, che permette mediante apposito

software “*Quantity One*”, di visualizzare su PC le bande di DNA amplificate e quindi accertarsi della buona riuscita dell’amplificazione dei segmenti di DNA oggetti di studio ai fini dell’ identificazione molecolare delle piante esaminate.

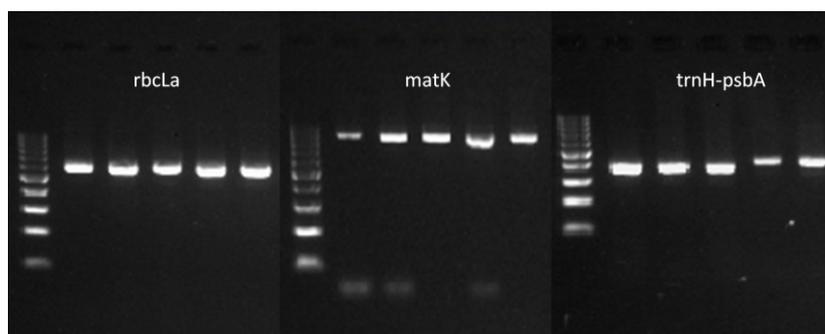


Fig.15 Elettroforesi su gel di agarosio al 2%, di amplificati di *Euphorbia bivonae*, mediante PCR con i primer *rbcLa*, *matK* e *trnH-psbA*



Fig.16 Acquisitore di immagini *Gel Doc* (BIORAD)

4.7. Sequenziamento ed analisi dei dati

Una volta accertata la buona riuscita del processo di amplificazione, i prodotti di PCR sono stati avviati alla procedura di sequenziamento secondo il metodo Sanger (già descritta nella parte introduttiva del presente elaborato).

La fase di sequenziamento, a differenza delle due precedenti, di estrazione ed amplificazione, non ha riguardato tutte le specie oggetto di studio, ma solo quelle che hanno avuto, per almeno uno dei tre campioni per singola specie, un successo di amplificazione tramite PCR.

I prodotti di PCR sono stati purificati e sequenziati seguendo il protocollo del kit commerciale *DYEnamic™ET Terminator Cycle Sequencing Kits* (Amersham Biosciences) mediante sequenziatore automatico *AB3730XL DNA Analyzer* (Applied Biosystems). I frammenti sono stati sequenziati sia in *forward* che in *reverse*, tramite gli stessi primer adottati per le PCR.

Tramite Sequencer software 4.10 (Gene Codes Corporation, USA) gli elettroferogrammi, ottenuti dal sequenziatore, sono stati accuratamente controllati ed eventualmente ripuliti manualmente, *forward* e *reverse* sono stati assemblati per tutte le repliche disponibili in *contigs*.

Le sequenze così ottenute sono state controllate per una prima identità tramite BLAST (NCBI) e successivamente allineate tramite il software MUSCLE, implementato all'interno del programma Mega 6 (Kimura 1980; Saitou and Nei 1987; Tamura et al. 2013) con il quale si è proceduto alla seguente fase di analisi filogenetica.

4.8 Valutazione delle performance dei marcatori barcoding

Sono stati valutati diversi parametri per riuscire a determinare in maniera efficiente il reale potere discriminante dei marcatori Barcoding utilizzati. Si è tenuto conto di due categorie di parametri, parametri che hanno tenuto conto delle performance tecniche e parametri utili a valutare il potere discriminante.

Performance tecniche

-*successo PCR*: è stato calcolato il numero di campioni positivi al test di PCR per ciascun marcatore, sia riferito al numero totale di replicati biologici analizzati, sia al numero di taxa analizzati.

-*successo di sequenziamento*: è stato calcolato il numero di campioni positivi alla procedura di sequenziamento, che ha riguardato solo i campioni positivi al test di PCR per ciascun marcatore, sia riferito al numero totale di replicati biologici analizzati, sia al numero di taxa analizzati.

-*qualità delle sequenze*: la qualità della sequenza è data dalla qualità dei picchi presenti sugli elettroferogrammi ad indicare la precisione ed attendibilità delle sequenze ottenute. Sono state considerate idonee le sequenze con qualità superiore al 70%. Il valore riportato indica la media dei replicati biologici.

-*lunghezza frammento*: è un valore riferito alla media della lunghezza dei frammenti ottenuti per ciascun marcatore, riferiti al totale dei replicati biologici, a seguito delle analisi e pulizia degli elettroferogrammi.

Discriminazione molecolare

- *potere di discriminazione*: il valore di tale parametro è dato dal numero di taxa che sono stati univocamente discriminati a livello di specie (o sottospecie).

Il potere discriminante è stato valutato sia per singolo locus, sia in approccio multi-locus. Il potere di discriminazione di ogni locus è stato valutato mediante analisi filogenetica con Mega 6, condotta tramite confronto tra tutte le sequenze generate in questo studio ed un sottoinsieme di sequenze referenti correlate a ciascun taxa reperite da BOLD Database / GenBank .

-livello di divergenza genetica: questo parametro indica il grado di variabilità tra un gruppo di sequenze, ottenuto dalle matrici di distanza calcolate secondo il parametro K2P% (Kimura, 1980). È stato calcolato all'interno di alcune famiglie considerate più rappresentative per numero di specie.

-Numero di siti variabili: indica il numero di basi soggette a variazioni all'interno del gruppo filogenetico considerato sulla lunghezza totale dei frammenti ottenuti per singolo locus. Come il precedente è stato calcolato all'interno di alcune famiglie considerate più rappresentative per numero di specie.

5 Risultati e discussioni

5.1 Conservazione del materiale vegetale

Per tutte le specie selezionate si è allestita presso il CRA-SFM di Bagheria una banca del materiale vegetale liofilizzato e la rispettiva banca del DNA, come importante stock per la conservazione del materiale vegetale in questione.

Per la categoria delle piante in produzione presso l'Azienda Faro, la collezione del DNA comprende 52 specie, inserite nella tabella 3 a pagina 49.

Per le piante di nuova introduzione ed in collezione presso l'Azienda Faro, la collezione del DNA comprende 12 specie, inserite nella tabella 4 a pagina 51.

Per le piante di nuova introduzione ed in collezione presso il CRA-SFM di Bagheria (PA), la collezione del DNA comprende 36 specie, inserite nella tabella 5.a pagina 52.

Tale materiale è stato identificato a livello morfologico e parte di esso ha avuto un successo di identificazione a livello molecolare ed è disponibile per ulteriori indagini genetiche.

5.2 Estrazione del DNA da materiale vegetale

La prima fase del lavoro si è focalizzata sulla messa a punto di protocolli di estrazione del DNA che potessero garantire un DNA di buona qualità e quantità da matrici vegetali di diversa tipologia, biochimica e strutturale (spessore, rigidità dei tessuti, pelosità, presenza di cere, linfa, eccesso acquoso, spinescenza), e che contemporaneamente potessero soddisfare l'esigenza di rapidità e impiego di sostanze chimiche meno tossiche possibili. Il protocollo riportato in questo studio a pag 55 si basa sul protocollo di estrazione da vegetali di Doyle & Doyle (1987), con delle variazioni che hanno consentito lo snellimento della procedura, partendo da materiale liofilizzato e tritato senza uso di azoto, e la totale eliminazione del fenolo (alto livello di tossicità).

Il protocollo di estrazione del DNA, oltre a risultare abbastanza rapido è risultato anche notevolmente economico. Valori di assorbanza A_{260}/A_{280} in un range di 1,6-2 e assorbanza A_{260}/A_{230} in un range di 1,3-1,7, sono stati registrati nell'80% dei campioni confermando la bontà della procedura utilizzata. Una fase di purificazione su colonna con PVPP è stata utile in presenza di tessuti più rigidi, ricchi in componenti resinose, linfa o eccessi acquosi, come nelle piante grasse (20 % campioni). I valori di concentrazione di DNA sono risultati buoni e di gran lunga superiori a quelli riportati per i principali kit commerciali di estrazione (Zetecche et. al -www.dnabank-network.eu), rientrando in un range di 150-270 ng/μl. Tutte condizioni fondamentali per un buon livello di successo nelle successive fasi che caratterizzano l'identificazione a livello molecolare delle specie vegetali oggetto di studio.

5.3.Valutazione delle performance tecniche dei marcatori barcoding

La valutazione dei marcatori barcoding, basata sulle performance tecniche segue i parametri a pagina 66.

I risultati relativi alle performance tecniche sono riportati in Tab.10. Sul numero totale dei 300 replicati biologici, per rbcL si sono avuti valori del 93% come successo di PCR, del 95% come successo di sequenziamento, con una qualità delle sequenze del 90% ed una lunghezza media dei frammenti di 569 bp, il matK ha fatto registrare un successo di PCR e di sequenziamento, rispettivamente del 70% e 93% ed una qualità delle sequenze dell'80% tutti e tre valori inferiori a rbcL ed una lunghezza media dei frammenti di 766 bp, il marcatore trnH-psbA ha dato un successo di PCR e di sequenziamento, rispettivamente dell'80% e 91% ed una qualità delle sequenze dell'85% con una lunghezza media dei frammenti di 518 bp. Sul numero totale dei 100 taxa testati rbcL ha dato valori superiori rispetto agli altri due marcatori, con successo di PCR del 97% ed un successo di sequenziamento del 99%, per matK i valori registrati sono stati dell'81% per il successo di PCR e 96% per il successo di sequenziamento, mentre il marcatore trnH-psbA ha fatto registrare un successo di PCR e di sequenziamento, rispettivamente dell'89% e 94%.

In relazione ai risultati sopra riportati l'rbcL ha mostrato le migliori performance: il maggiore successo di amplificazione, la migliore resa nel sequenziamento sia in termini di numero di sequenze ottenute sia in termini di qualità delle sequenze ottenute. Il matK, pur avendo registrato maggiori difficoltà di amplificazione, in accordo a numerosi altri studi (Chase et. Al 2007; Chen et. Al 2010; de Vere et al. 2012), ha mostrato un buon successo di sequenziamento e di qualità delle sequenze ottenute, diversamente da quanto indicato in alcuni protocolli che suggeriscono per questa regione la necessità di ulteriori primer da adottare per la fase di sequenziamento (Hollingsworth et. al 2011).

Tab.10 Performance tecniche dei marcatori barcoding: i parametri sono riferiti al totale dei replicati biologici (**a**) e al numero dei taxa testati (**b**)

a

	rbcl	matK	trnH-psbA
Num. di campioni testati*	300	300	222
Successo PCR	(93%) 279/300	(70%) 210/300	(80%) 177/222
Successo sequenziamento (<i>contigs</i>)	(95%) 265/279	(93%) 195/210	(91%) 161/177
Qualità delle sequenze (<i>contigs</i>)	90%	80%	85%
Lunghezza frammento (media bp)	569	766	518

*include tre replicati biologici per ogni taxa testato

b

	rbcl	matK	trnH-psbA
Num. di taxa testati	100	100	74
Successo PCR	(97%) 97/100	(81%) 81/100	(89%) 66/74
Numero di taxa sequenziati con successo	(99%) 96/97	(96%) 78/81	(94%) 62/66

5.4 Identificazione tramite DNA Barcoding

Il percorso di analisi molecolare al fine identificativo dei taxa è stato condotto focalizzando preliminarmente sulla valutazione dei “core markers” (*rbcL* e *matK*). L’impiego del terzo marcatore, l’IGS *trnH-psbA* è stato riservato a quelle situazioni in cui entrambi i core markers hanno presentato difficoltà, o per mancata amplificazione, per fallimento delle sequenze o insufficiente potere discriminante.

I risultati complessivi di identificazione a livello di specie per ciascun taxa testato sono riportati in Tab.11. Su un totale di 100 taxa testati, il 61% dei taxa sono stati identificati con successo a livello di specie con almeno uno dei tre locus, mentre per il 37% rimane l’identificazione a livello di genere. Solo il restante 2% dei taxa è rimasto non determinato a causa del fallimento di tutti e tre i marcatori impiegati.

Considerando i singoli marcatori, l’*rbcL* ha permesso una identificazione univoca a livello di specie di 34 taxa (35%), il *matK* di 34 taxa (44%) ed il *trnH-psbA* di 32 taxa (52%) (Tab.11-12). Il *matK* ha mostrato maggiori valori percentuali di potere risolutivo in termini di discriminazione dei taxa rispetto all’*rbcL*, confermando i trend indicati da altri studi (Chase *et al.* 2007, Chen *et al.* 2010, de Vere *et al.* 2012, Lahaye *et al.* 2008, Khew & Chia 2011, Costion *et al.* 2011).

Laddove vi è stata una mancata discriminazione a livello di specie sia per *rbcL* che per *matK* (Tab.11), il *trnH-psbA* è stato determinante nella risoluzione di ulteriori 14 taxa, permettendo in tal modo di incrementare il numero totale dei taxa discriminati da 47 a 61 taxa. I core markers, impiegati in multi-locus, *rbcL+matK*, hanno permesso l’identificazione univoca a livello di specie di 38 taxa. Le ulteriori combinazioni dei due marcatori *rbcL+trnH-psbA* e *matK+trnH-psbA* hanno consentito la discriminazione rispettivamente di 32 taxa e 25 taxa (Tab.11-12).

L’impiego dell’approccio multi-locus basato sui core markers è apparso quello più efficiente, generando un buon compromesso tra le elevate performance tecniche dell’*rbcL* ed il migliore potere risolutivo supportato dal *matK*. Non appaiono altrettanto efficienti le altre combinazioni testate (Pettengil and Neel 2010; Gere *et al.* 2013).

L’identificazione molecolare dei singoli taxa è stata ottenuta a seguito delle analisi filogenetiche condotte come riportato a pag 67

Per le famiglie più ampie, sono stati generati gli alberi filogenetici Neighbor Joining tree, per singolo marcatore, mostrando il grado di divergenza genetica tra le sequenze dei taxa investigati e le sequenze di riferimento per ciascuna specie presente su BOLD/GenBank. Per tutti i taxa investigati, le distanze genetiche dalle sequenze referenti più affini sono state riportate nelle tabelle riassuntive (Appendice).

Le famiglie dove si è registrato il più alto successo di discriminazione (Tab.11) a livello di specie, sono: Asteraceae con 9 taxa univocamente discriminati su un totale di 13, Caryophyllaceae con 4 taxa su un totale di 6, Fabaceae con 8 taxa su un totale di 8, Euphorbiaceae con 9 taxa su un totale di 9, le Brassicaceae con 3 taxa su un totale di 3, Ericaceae con 4 taxa su un totale di 4.

I minori successi in termini di risoluzione univoca a livello di specie, sono stati riscontrati per le Arecaceae, con 7 taxa discriminati a livello di specie su un totale di 24, e per le Cistaceae dove nessuno dei 4 taxa esaminati ha avuto successo nella discriminazione con tutti e tre i marcatori.

I livelli di divergenza genetica per le famiglie più ampie sono riportati in Tab.13.

Il marcatore rbcL ha fatto registrare i valori più bassi di divergenza genetica per le Arecaceae, con lo 0,7%, ed i valori più alti per le Asteraceae, con il 2,1%, mentre matK ha dato i valori più bassi per le Arecaceae con il 1,5% ed i più alti per le Fabaceae con il 6,4%. Il trnH-psbA ha dato i valori più alti per le Euphorbiaceae con il 9,1% ed i più bassi per le Arecaceae con il 2,9%. Il trnH-psbA ha confermato elevati valori di variabilità e la sua capacità di discriminazione all'interno di gruppi tassonomici molto affini (Chase et. Al 2007; Chen et. Al 2010; Hollingsworth et al 2011; De Mattia 2011).

Nella famiglia delle Arecaceae, che nel nostro studio ha riguardato 24 specie di 15 generi diversi, il marcatore trnH-psbA ha fatto registrare il più alto valore di divergenza genetica con una percentuale del 2,9%, ma stranamente rispetto al suo potenziale conosciuto non ha consentito alcuna discriminazione delle specie testate. I valori più bassi si sono avuti con rbcL con una percentuale dello 0,7%, mentre il matK ha dato valori intermedi ai primi due con un valore di 1,5% (Tab.13). Dagli alberi filogenetici nelle figure 17-18-19 si evince che per le Arecaceae rbcL ha permesso di identificare a livello di specie 2 soli taxa (*Acoelorrhaphe wrightii* e *Caryota urens*). Laddove ha fallito rbcL, il matK è stato determinante per 4 taxa (*Arenga engleri*, *Phoenix roebelenii*, *Sabal minor*, *Bismarckia nobilis*). In ogni caso altri autori indicano l'rbcL ed il matK altamente

risolutivi ed in grado di supportare la filogenesi della famiglia (Asmussen et. al 2006; Asmussen and Chase 2011;).

Solo nel caso di *Washingtonia robusta*, la discriminazione è stata possibile tramite l'impiego di entrambi i core markers, ciò per consentire l'esclusione di specie affini che si sono presentati con entrambi i marcatori.

Per le Fabaceae (8 specie indagate di 3 generi diversi) i valori più bassi di divergenza genetica si sono registrati con rbcL con 1,5% ed i più alti con trnH-psbA con valori di 7,4% (Tab.13), mentre per matK si è avuta una divergenza genetica del 6,4%, comunque particolarmente elevata, discriminando 4 specie su 8. Il matK è stato determinate per 1 taxa (*Genista aristata*), mentre il trnH-psbA è stato determinate per 2 taxa (*Genista tyrrhena*, *Genista demarcoi*). Considerando che *Genista* è stato il genere più rappresentato, con 6 specie, l'rbcL all'interno di questo gruppo ha mostrato un risultato migliore rispetto al matK, discriminando 5 specie (*Spartium junceum*, *Ceratonia siliqua*, *Genista madoniensis*, *Genista cupanii*, *Genista aetnensis*), come anche il trnH-psbA come si evince dagli alberi filogenetici nelle figure 20-21-22. Questo risultato appare in contrasto con il potenziale espresso dal matK all'interno delle Fabaceae in altri studi (Gao et al 2011; Gao et Chen 2009; Madesis et. al 2012), ma il gruppo genista è stato affrontato per la prima volta in questo studio, con ottimi livelli di discriminazione.

Per le Asteraceae (13 specie indagate di 8 generi diversi) il matK ha fatto registrare valori di divergenza genetica del 4,4% e rbcL 2,1%. Per quanto riguarda il trnH-psbA considerata l'eccessiva variabilità mostrata dalle sequenze analizzate, si è preferita una suddivisione in generi: *Anthemis*, *Jacobaeae*, *Helichrysum*. I valori di divergenza genetica più bassi sono stati registrati per *Anthemis* con l'1% ed i più alti per *Jacobaeae* con il 3,3%. (Tab.13).

Per Asteraceae l'albero filogenetico rbcL ha permesso di identificare a livello di specie 4 taxa (*Centaurea sphaerocephala*, *Helichrysum nebrodense*, *Ptilostemon greuteri*, *Pallenis maritima*), mentre il matK ha discriminato 5 taxa risultando determinante in particolare nella discriminazione di 2 specie (*Helichrysum italicum*, *Hieracium cophanense*). Il trnH-psbA è stato determinate nella risoluzione di ulteriori 3 taxa (*Jacobaea gibbosa*, *Jacobaea ambigua*, *Senecio candidus*) (alberi filogenetici figure 23-24-25-26-27). Dalle analisi effettuate *Jacobaea gibbosa* ha mostrato un'ampia variabilità rispetto alle altre specie del genere *Jacobaea*, discostandosi da queste in

tutti e tre i marcatori utilizzati. Pertanto potrebbe costituire un cluster differente all'interno della specie. Per le *Asteraceae* la discriminazione è stata piuttosto elevata in accordo con altri studi sulla famiglia che indicano elevati livelli di successo (Gao et. al 2010).

All'interno della famiglia delle Euphorbiaceae (9 specie indagate di un unico genere) i valori più bassi di divergenza genetica si sono avuti con rbcL con l'1,2% ed i più alti con trnH-psbA con il 9,1% e valori intermedi con il matK con il 4%. (Tab.13)

Nelle figure 28-29-30 sono mostrate le relazioni filogenetiche dei tre marcatori per le Euphorbiaceae. L'rbcL ha identificato 6 taxa (*Euphorbia bivonae*, *Euphorbia ceratocarpa*, *Euphorbia dendroides*, *Euphorbia helioscopia*, *Euphorbia myrsinites*, *Euphorbia pithyusa subsp. cupanii*). Il matK ha identificato correttamente 3 taxa ed è stato determinante per 1 taxa (*Euphorbia amygdaloides*), mentre il trnH-psbA è stato determinante in 2 taxa (*Euphorbia characias*, *Euphorbia meuselii*).

Per i generi *Brassica*, *Erica*, *Cistus*, *Chamaerops*, *Dianthus*, *Euphorbia* e *Genista*, il lavoro di identificazione molecolare si è basato sull'impiego di specie referenti reperite appositamente per questo studio a causa dell'assenza nelle banche dati internazionali di specie affini a quelle selezionate. Pertanto queste specie e le rispettive sequenze rappresentano le referenti utilizzate per il confronto con quelle in collezione e saranno inserite all'interno dei database internazionali.

Tab.11 Risultati di identificazione a livello di specie dei taxa testati con i marcatori *rbcl*, *matK* e *trnH-psbA* (in verde successo di identificazione a livello di specie ed in giallo identificazione a livello di genere)

FAMIGLIA	SPECIE	ID	rbcl	matK	trnH-psbA
ARECACEAE	<i>Chamaerops humilis</i> var. <i>humilis</i>	F.C4			
	<i>Chamaerops humilis</i> var. <i>argentea</i>	F.C5			
	<i>Chamaerops humilis</i> L. "Vulcano"	F.C6		seq. fallita	
	<i>Chamaerops humilis</i> L. "Etna star"	F.C7		seq. fallita	
	<i>Brahea armata</i>	F.B4			seq. fallita
	<i>Brahea edulis</i>	F.B5	PCR fallita	PCR fallita	PCR fallita
	<i>Acoelorrhaphe wrightii</i>	F.A11			non necessario
	<i>Arenga engleri</i>	F.A7			non necessario
	<i>Caryota urens</i>	F.C20			non necessario
	<i>Howea forsteriana</i>	F.H1			
	<i>Livistona chinensis</i>	F.L1			
	<i>Washingtonia robusta</i>	F.W1			non necessario
	<i>Washingtonia filifera</i>	F.W2			seq. fallita
	<i>Trachycarpus fortunei</i>	F.T1			
	<i>Phoenix roebelenii</i>	F.P7			non necessario
	<i>Phoenix canariensis</i>	F.P2			seq. fallita
	<i>Phoenix dactylifera</i>	F.P8			
	<i>Phoenix reclinata</i>	F.P9			
	<i>Sabal minor</i>	F.S14			
	<i>Sabal palmetto</i>	F.S3			
	<i>Bismarckia nobilis</i>	F.B3			non necessario
	<i>Butia capitata</i>	F.B2			PCR fallita
	<i>Syagrus romanzoffiana</i>	F.A8			PCR fallita
	<i>Trithrinax campestris</i>	F.T4			PCR fallita

ASTERACEAE	<i>Anthemis cupaniana</i>	F.A2			
	<i>Anthemis maritima</i>	F.A4	seq. fallita	PCR fallita	
	<i>Centaurea sphaerocephala</i>	F.C14			non necessario
	<i>Helichrysum nebrodense</i>	H5			
	<i>Helichrysum scandens</i>	H7			
	<i>Helichrysum hyblaicum</i>	H6			seq. fallita
	<i>Helichrysum italicum</i>	F.H2	PCR fallita		
	<i>Ptilostemon greuteri</i>	F.P5			non necessario
	<i>Jacobaea gibbosa</i>	F.J1			
	<i>Jacobaea ambigua</i>	F.S4			
	<i>Senecio candidus</i>	F.S5			
	<i>Pallenis maritima</i>	F.P3		PCR fallita	non necessario
	<i>Hieracium cophanense</i>	H2			non necessario
FABACEAE	<i>Spartium junceum</i>	F.S9			non necessario
	<i>Ceratonja siliqua</i>	F.C1			non necessario
	<i>Genista tyrrhena</i>	G1			
	<i>Genista madoniensis</i>	G2			
	<i>Genista aristata</i>	G3			
	<i>Genista demarcoi</i>	G4			
	<i>Genista cupanii</i>	G5			
	<i>Genista aetnensis</i>	G6			non necessario
EUPHORBIACEAE	<i>Euphorbia bivonae</i>	Eu8			
	<i>Euphorbia ceratocarpa</i>	Eu1			
	<i>Euphorbia characias</i>	E2		seq. fallita	
	<i>Euphorbia dendroides</i>	E4		PCR fallita	
	<i>Euphorbia helioscopia</i>	Eu7			
	<i>Euphorbia myrsinites</i>	E6		PCR fallita	
	<i>Euphorbia meuselii</i>	E5		PCR fallita	
	<i>Euphorbia pithyusa</i> subsp. <i>cupanii</i>	Eu9		PCR fallita	
	<i>Euphorbia amygdaloides</i>	Eu11			
ROSACEAE	<i>Rosa sempervirens</i>	R3			
	<i>Rosa canina</i>	R1			
	<i>Rosa sicula</i>	R6			
	<i>Rosa corymbifera</i>	R10			

BRASSICACEAE	<i>Brassica insularis</i>	B2		PCR fallita	
	<i>Brassica villosa</i> subsp. <i>tinei</i>	B4			non necessario
	<i>Brassica rupestris</i> subsp. <i>hispida</i>	B5			non necessario
ACANTHACEAE	<i>Acanthus mollis</i>	F.A12			non necessario
XANTHORRHOEACEAE	<i>Aloe arborescens</i>	FA10			
	<i>Aloe vera</i>	FA9			
	<i>Aloe plicatilis</i>	FA5			
	<i>Aloe x spinosissima</i>	FA17	PCR fallita	PCR fallita	PCR fallita
LILIACEAE	<i>Tulipa radii</i>	T4		PCR fallita	
AMARANTHACEAE	<i>Diotis maritima</i>	F.D1			non necessario
IRIDACEAE	<i>Iris pseudopumila</i>	F.I1			
	<i>Iris germanica</i>	F.I5			
LAMIACEAE	<i>Lavandula angustifolia</i>	F.L4			non necessario
	<i>Lavandula stoechas</i>	F.L5			non necessario
	<i>Salvia leucantha</i>	F.S12			non necessario
	<i>Salvia officinalis</i>	F.S11			
	<i>Rosmarinus officinalis</i>	F.R4			non necessario
	<i>Sideritis italica</i>	F.S6			
STRELITZIACEAE	<i>Strelitzia augusta</i>	F.S1		PCR fallita	PCR fallita
	<i>Strelitzia nicolai</i>	F.S2		PCR fallita	PCR fallita
	<i>Strelitzia reginae</i>	F.S8		PCR fallita	PCR fallita
CISTACEAE	<i>Cistus salviifolius</i>	F.C2		PCR fallita	
	<i>Cistus x pulverulentus</i>	F.C3		PCR fallita	
	<i>Cistus albidus</i>	F.C11		PCR fallita	
	<i>Cistus skanbergii</i>	F.C12		PCR fallita	
MYRTACEAE	<i>Myrtus luma</i>	FM1		PCR fallita	
	<i>Myrtus communis</i>	F.M5			
	<i>Metrosideros excelsa</i>	F.M2			non necessario
CYCADACEAE	<i>Cycas revoluta</i>	F.C8			
	<i>Cycas circinalis</i>	F.C19			non necessario
ERICACEAE	<i>Erica sicula</i>	E7			non necessario
	<i>Erica peduncularis</i>	E9			
	<i>Erica multiflora</i>	E10			
	<i>Arbutus unedo</i>	F.A1			non necessario

CONVOLVULACEAE	<i>Calystegia soldanella</i>	F.C13		PCR fallita	
TAMARICACEAE	<i>Tamarix gallica</i>	F.T5			non necessario
CARYOPHILLACEAE	<i>Dianthus rupicola</i> subsp. <i>aeolicus</i>	D1			
	<i>Dianthus busambrae</i>	D2			
	<i>Dianthus rupicola</i> subsp. <i>rupicola</i>	D3			
	<i>Dianthus rupicola</i> subsp. <i>lopadusanum</i>	D4			
	<i>Dianthus siculus</i>	D5			
	<i>Dianthus rupicola</i> subsp. <i>hermaensis</i>	D6			

Tab.12 Potere discriminante dei marcatori Barcoding

singolo locus	risoluzione a livello di specie		tot. %	
	rbcl*		35%	(34/96)
matK*		44%	(34/78)	
trnH-psbA*		52%	(32/62)	
multi-locus	rbcl + matK**		49%	(38/77)
	rbcl + trnH-psbA**		53%	(32/60)
	matK + trnH-psbA**		54%	(25/46)

I valori tra parentesi sono il numero di taxa univocamente discriminati a livello di specie (o sottospecie) riferito al numero totale di sequenze ottenute per taxa

Tab.13 Divergenze genetiche calcolate con il parametro K2P% (Kimura 1980)

Famiglia	rbcl			matK			trnH-psbA			
	N. seq	siti variabili	GD%	N. seq	siti variabili	GD%	N. seq	sirti variabili	GD%	
Arecaceae	107	24/533	0,7	115	101/770	1,5	36	85/676	2,9	
Fabaceae	25	32/543	1,5	21	181/815	6,4	9	67/329	7,4	
Euphorbiaceae	27	39/540	1,2	14	88/769	4	18	164/736	9,1	
Asteraceae	67	54/563	2,1	83	157/797	4,4	Jacobaea	14	22/416	3,3
							Helichrysum	14	44/533	3,1
							Anthemis	20	8/348	1

Fig.17 Albero filogenetico (Neighbor Joining) *rbcl* Arecaceae

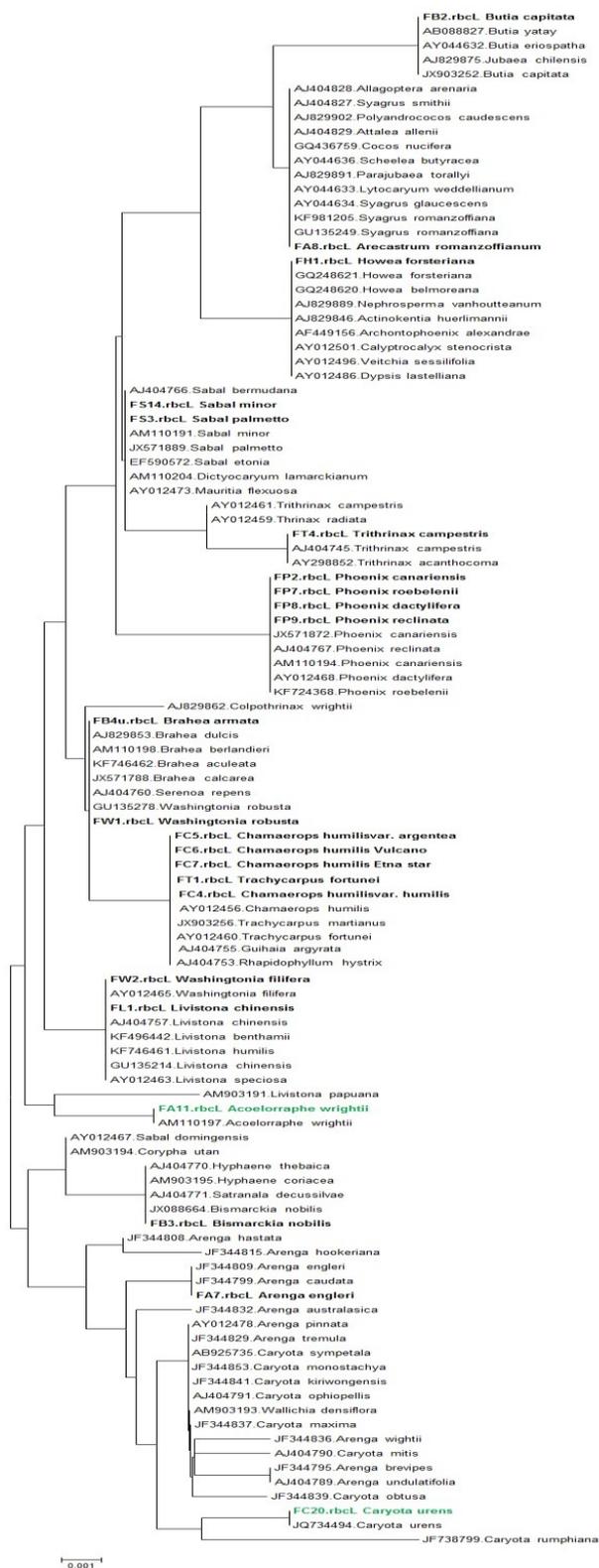


Fig.18 Albero filogenetico (Neighbor Joining) *matK* Arecaceae

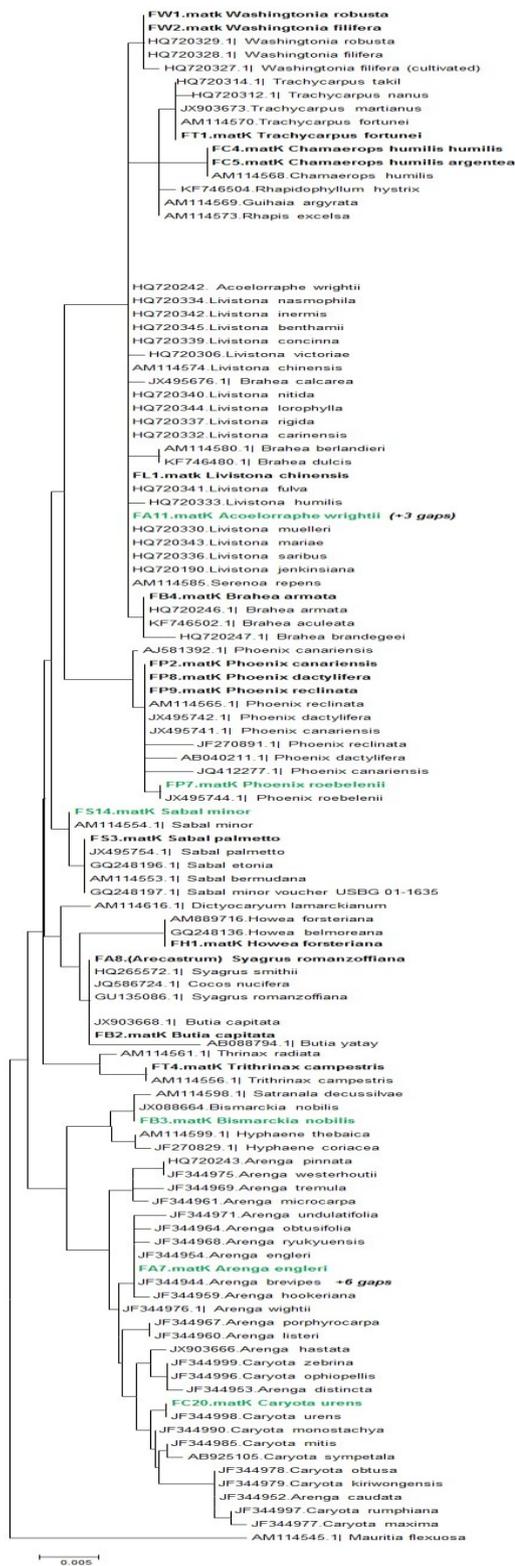


Fig.19 Albero filogenetico (Neighbor Joining) *trnH-psbA* Arecaceae

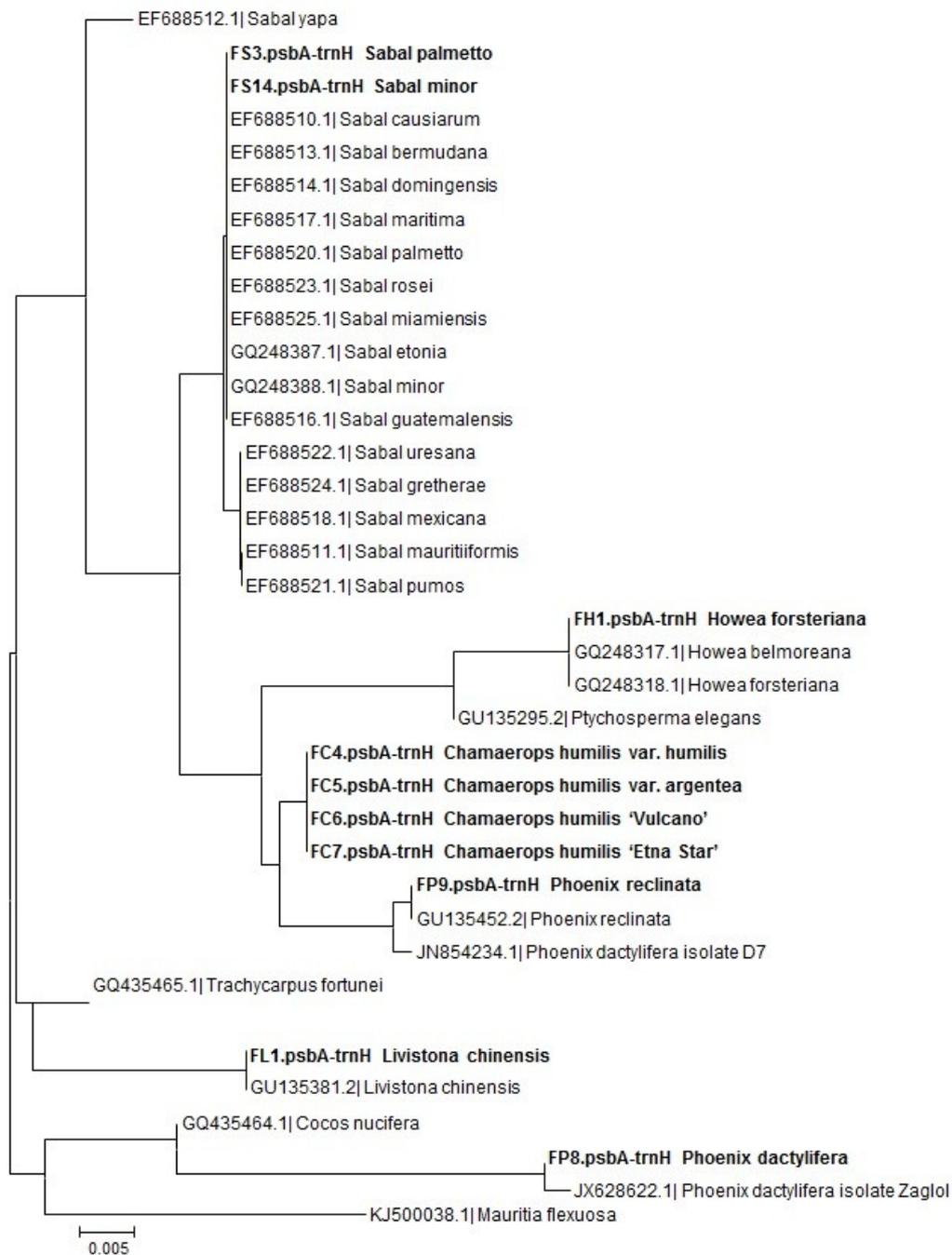


Fig.20 Albero filogenetico (Neighbor Joining) *rbcL* Fabaceae

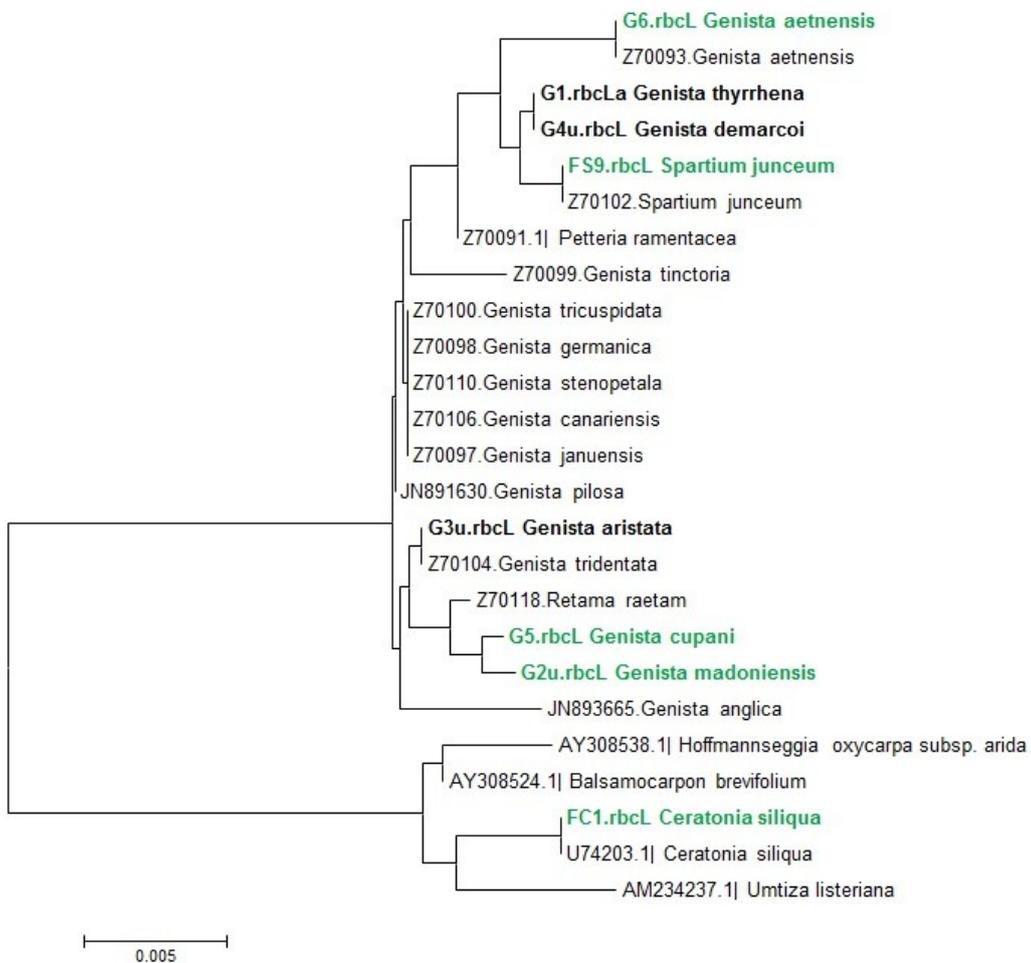


Fig.21 Albero filogenetico (Neighbor Joining) *matK* Fabaceae

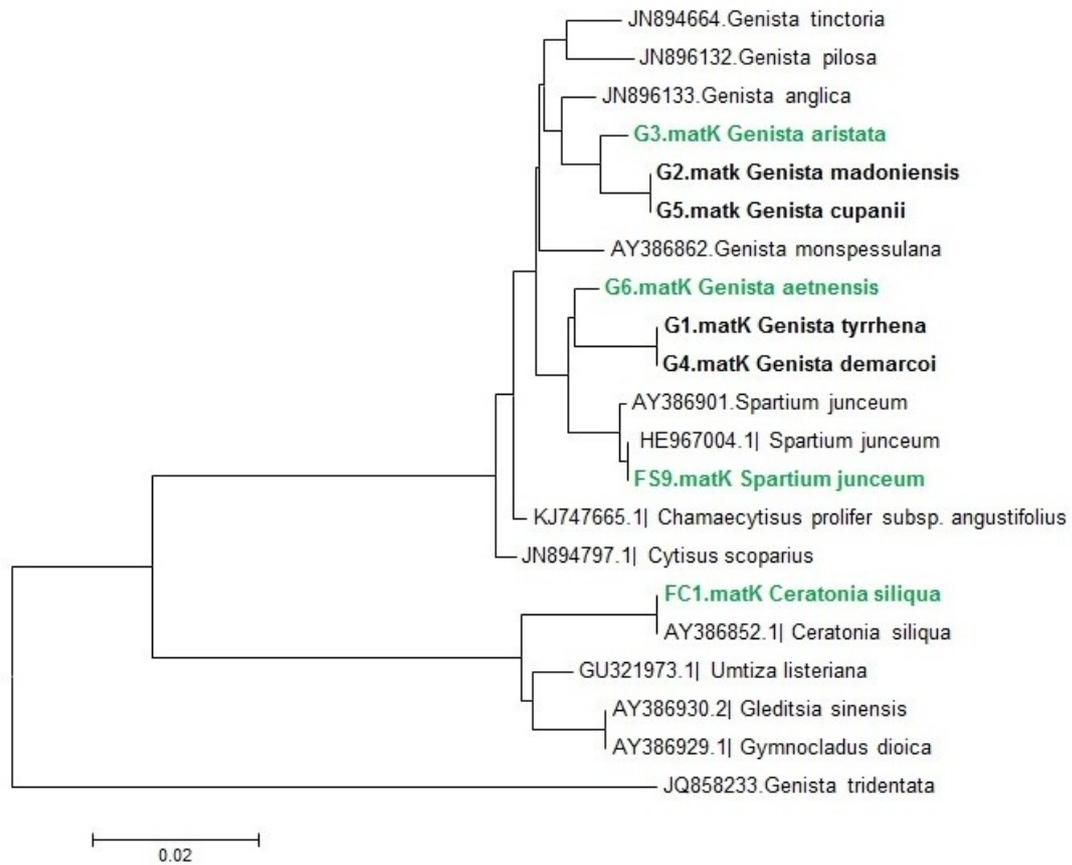


Fig.22 Albero filogenetico (Neighbor Joining) *trnH-psbA* Fabaceae

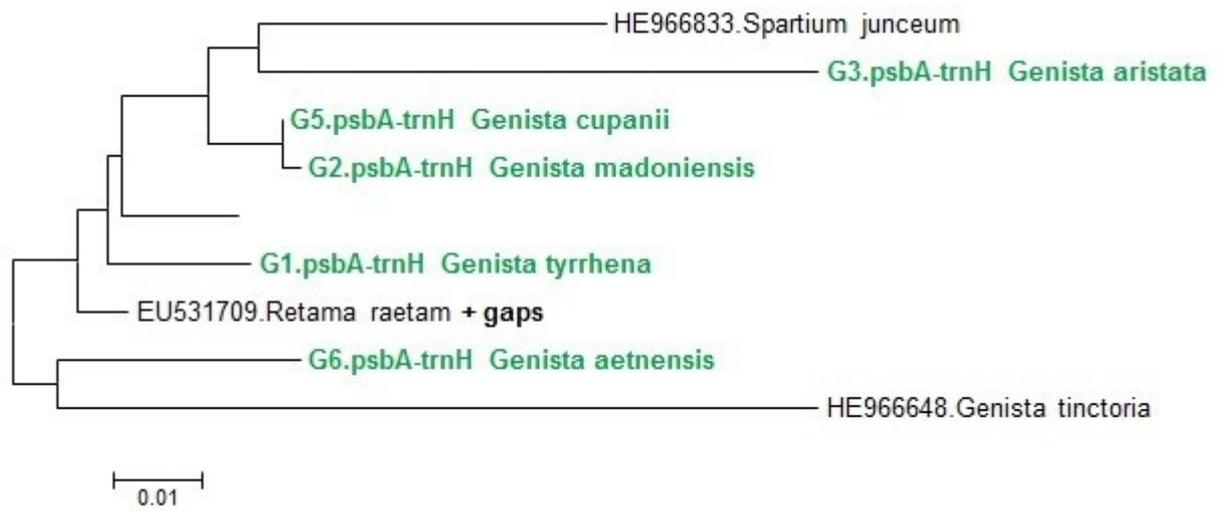


Fig.23 Albero filogenetico (Neighbor Joining) *rbcL* Asteraceae

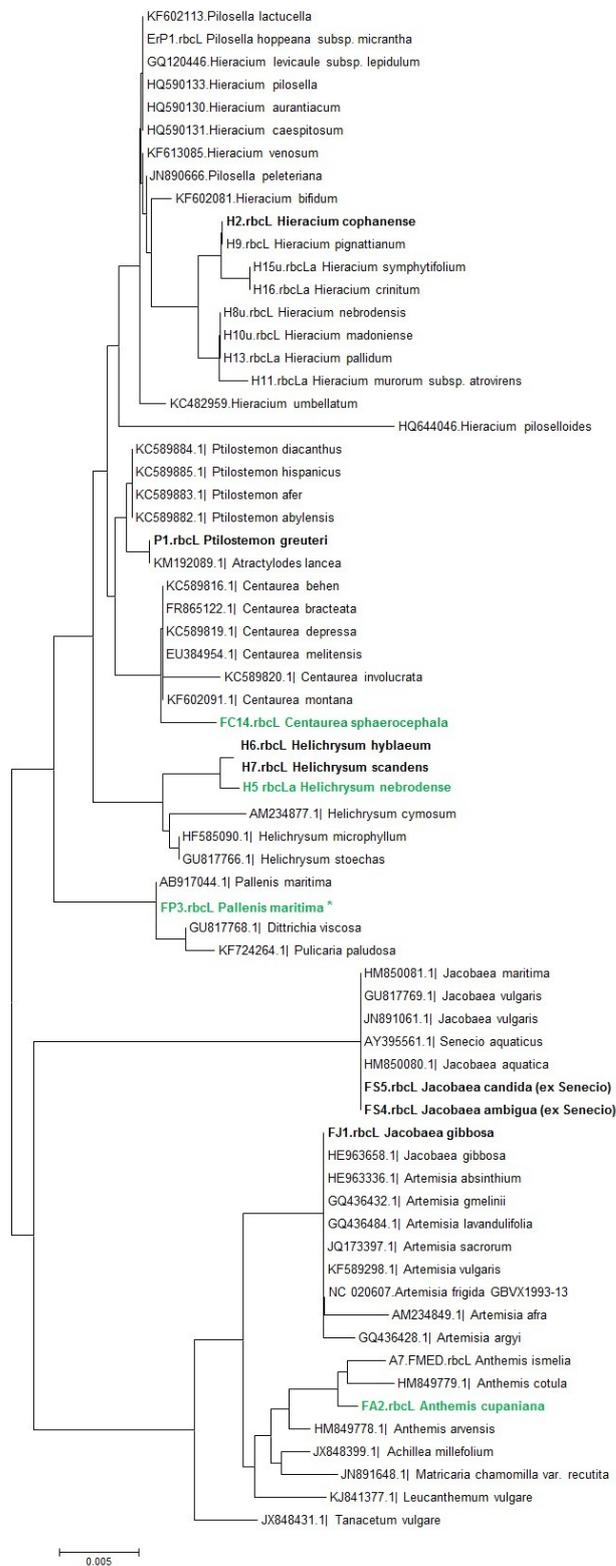


Fig.24 Albero filogenetico (Neighbor Joining) *matK* Asteraceae

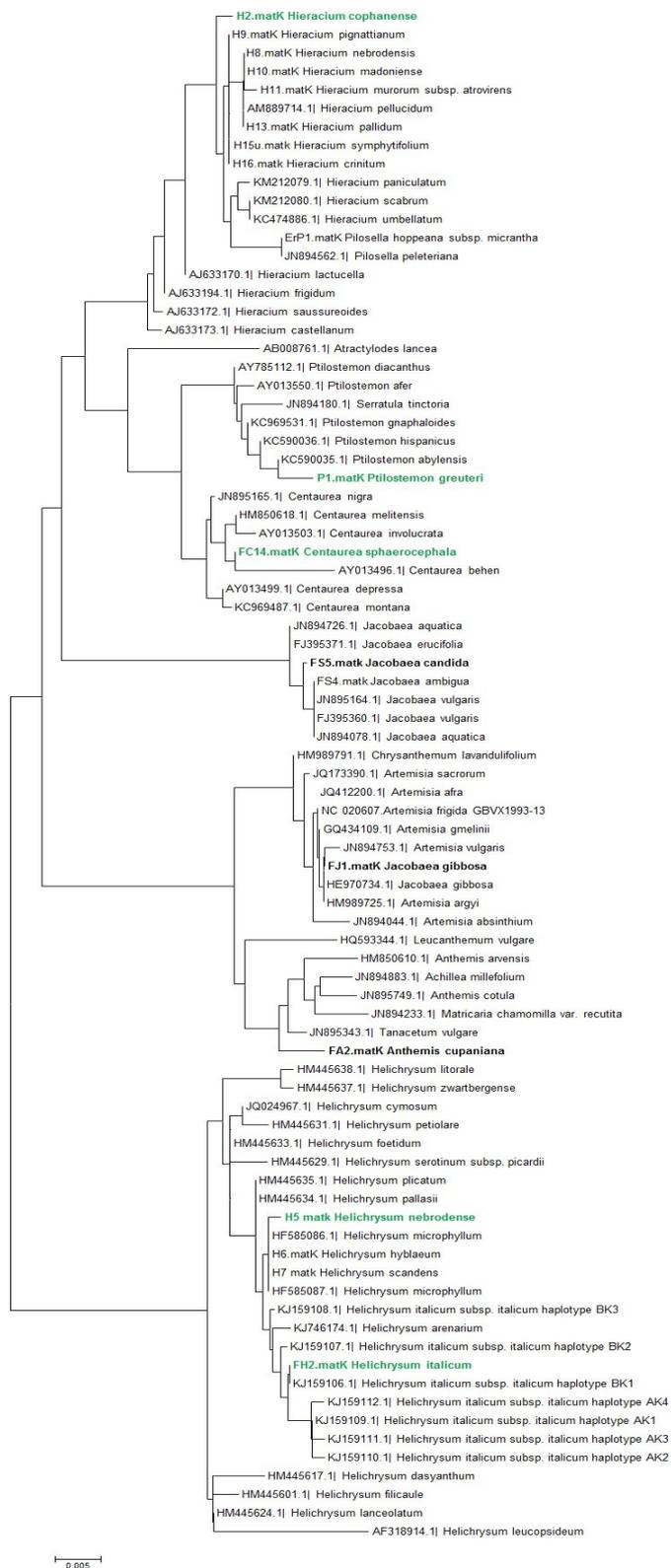


Fig.25 Albero filogenetico (Neighbor Joining) *trnH-psbA* *Jacobaea*

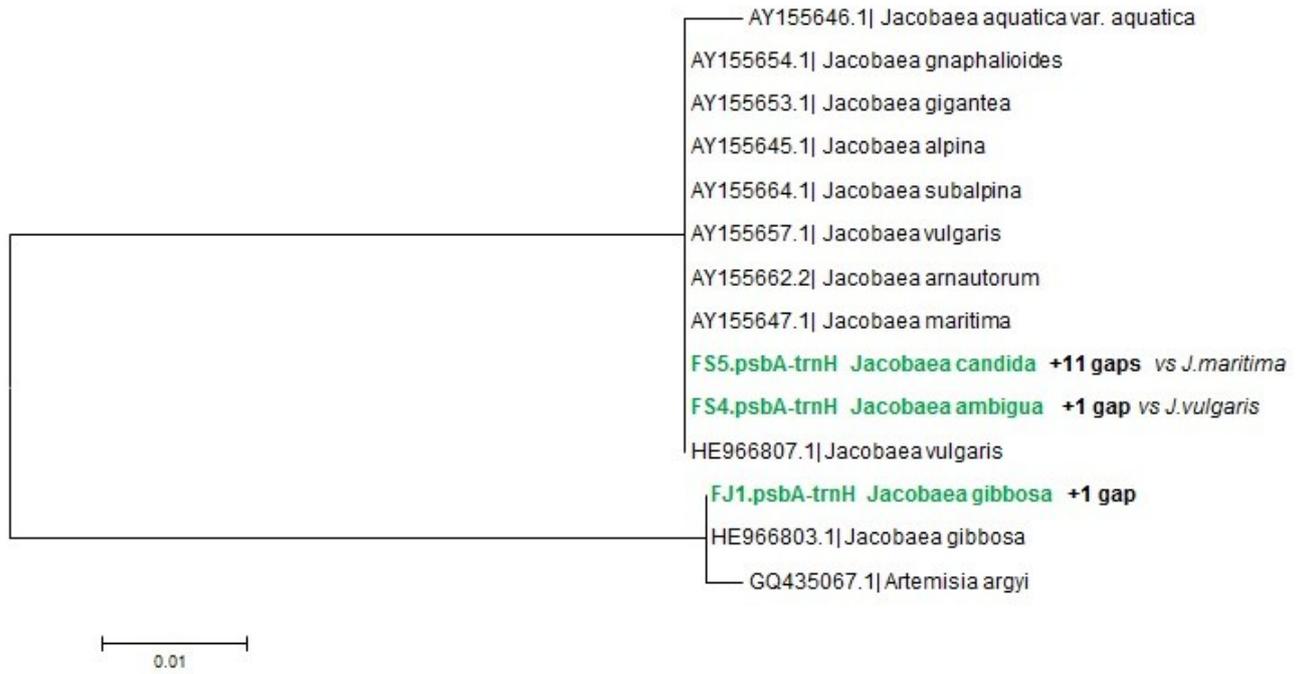


Fig.26 Albero filogenetico (Neighbor Joining) *trnH-psbA Helichrysum*

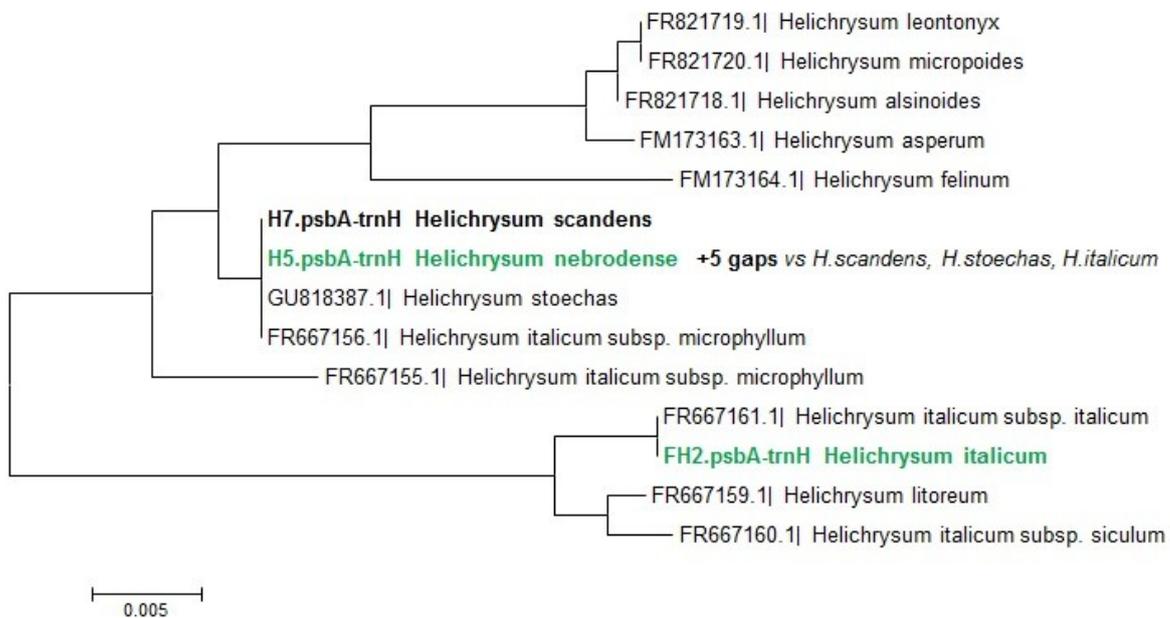
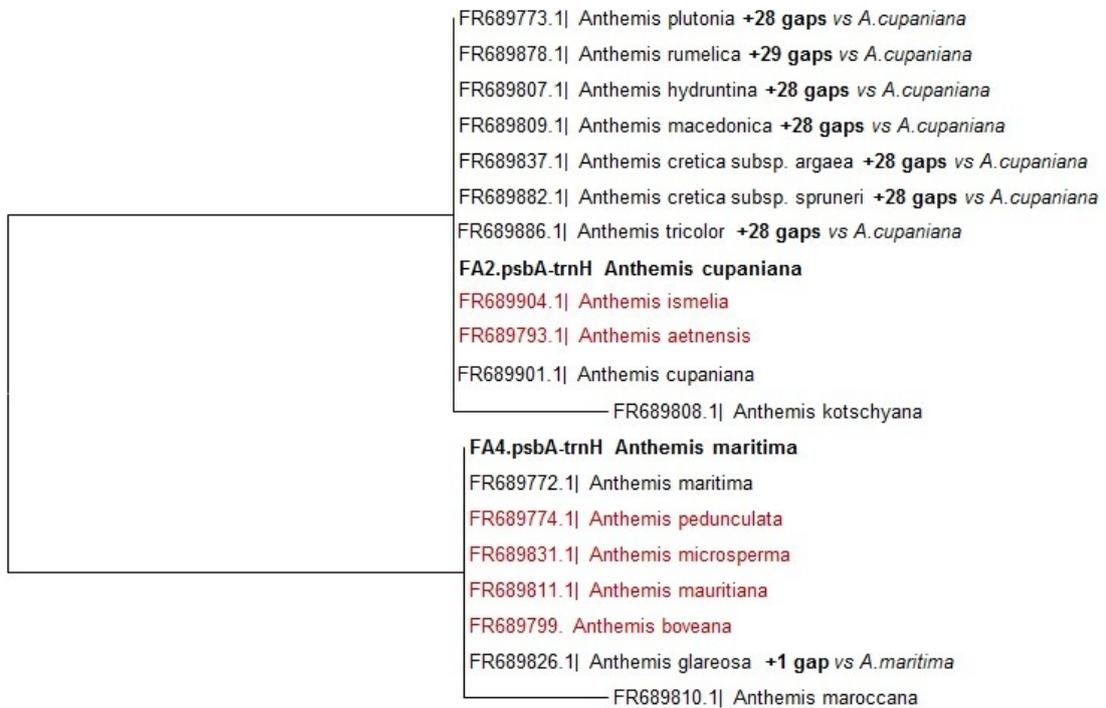


Fig.27 Albero filogenetico (Neighbor Joining) *trnH-psbA* *Anthemis*



0.002

Fig.28 Albero filogenetico (Neighbor Joining) *rbcl* Euphorbiaceae

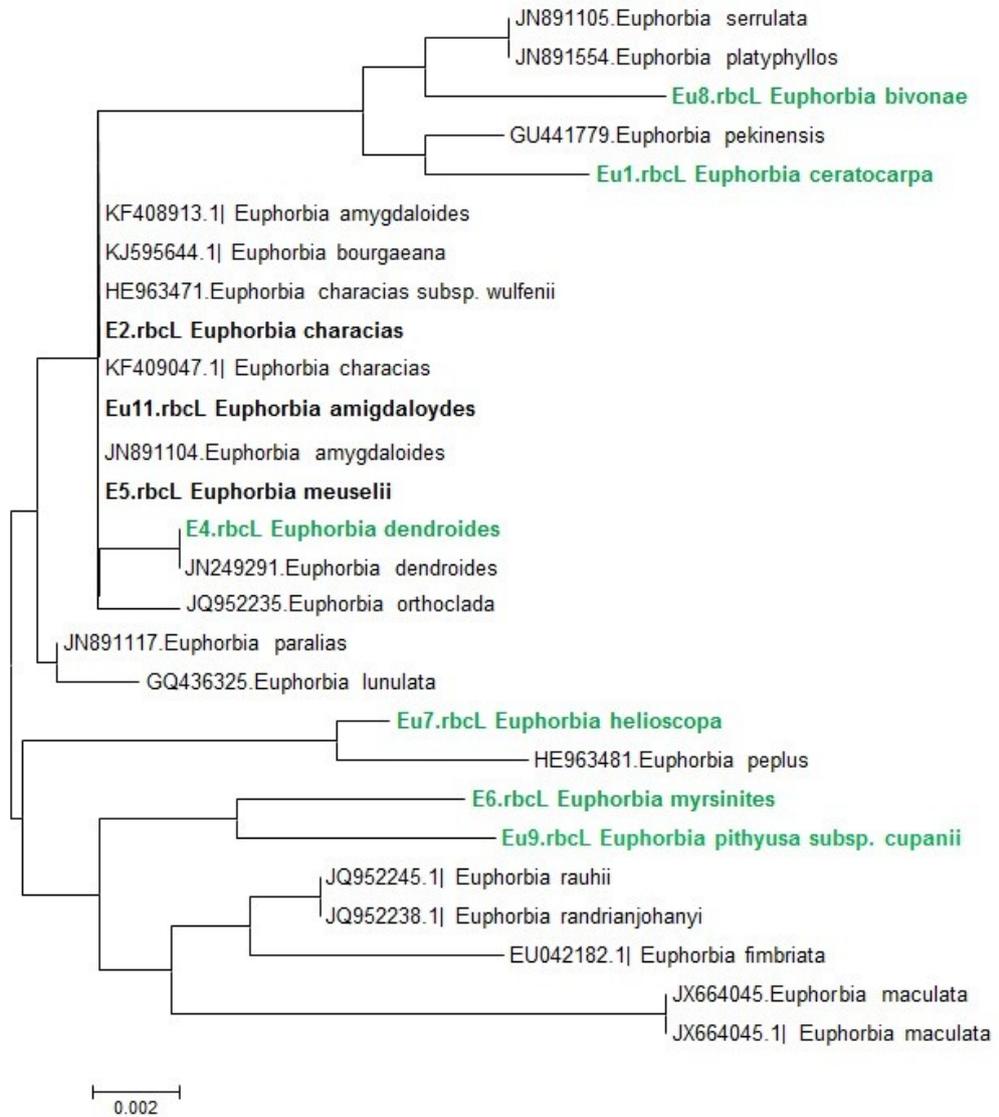


Fig.29 Albero filogenetico (Neighbor Joining) *matK* Euphorbiaceae

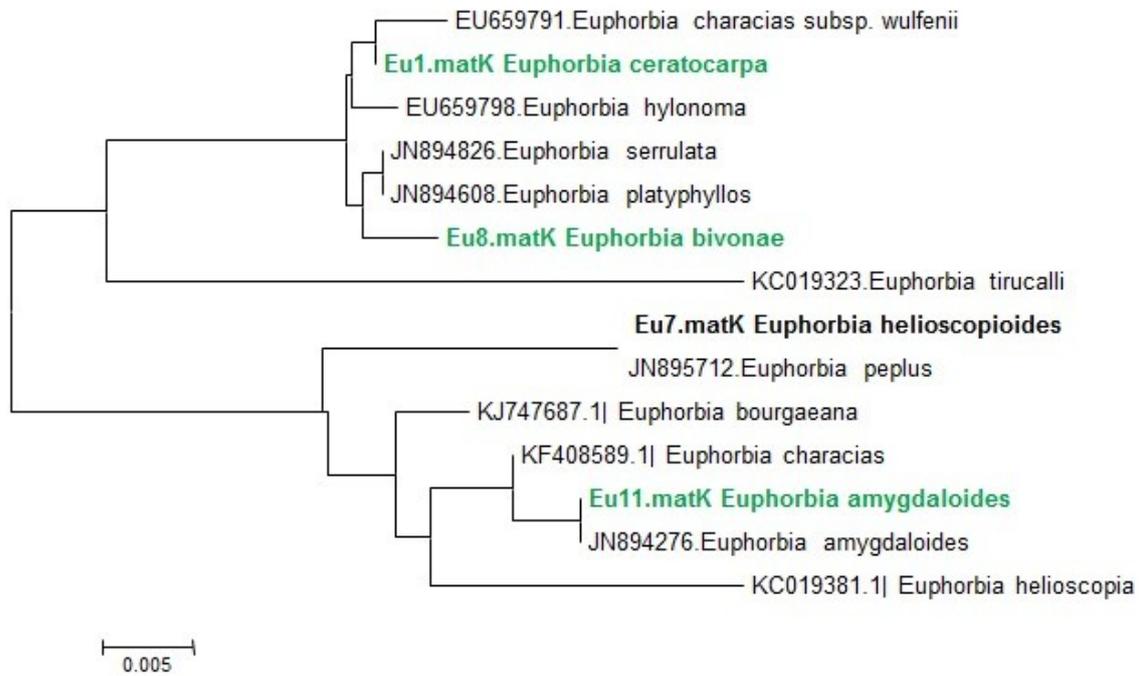
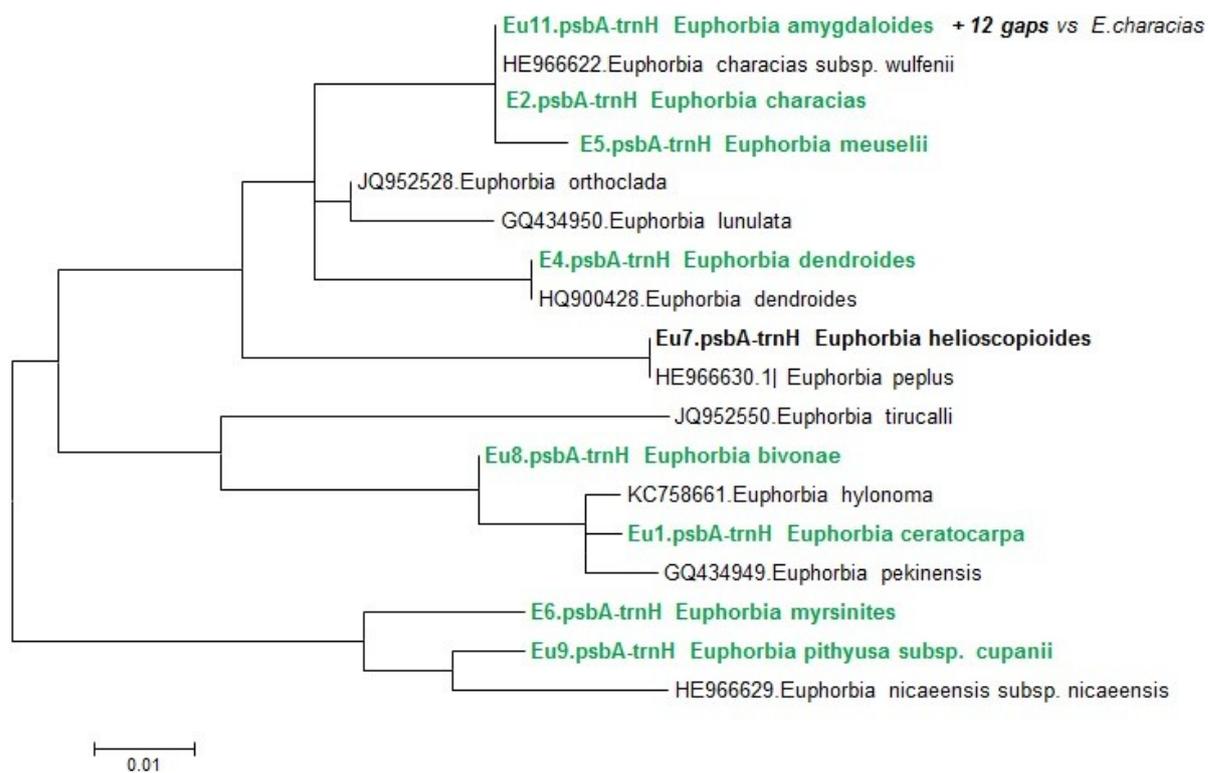


Fig.30 Albero filogenetico (Neighbor Joining) *trnH-psbA* Euphorbiaceae



5.5 Problematiche riscontrate e soluzioni

La principale problematica emersa da questo studio ha riguardato la mancanza di sequenze di riferimento disponibili per la comparazione. Ciò ha determinato l'impossibilità di discriminare alcuni gruppi (vedi Tab.14-15-16 Appendice): *Livistona chinensis*, *Trachycarpus fortunei*, *Phoenix dactylifera*; *Phoenix reclinata*, *Trithrinax campestris*, *Anthemis cupaniana*, *Butia capitata*, *Senecio candidus*, *Aloe arborescens*, *Aloe plicatilis*, *Iris pseudopumila*, *Iris germanica*, *Salvia officinalis*, *Cistus salvifolius*, *Cistus x pulverulentus*, *Cistus albidus*, *Cistus skanbergii*, *Dianthus rupicola* subsp. *aeolicus*, *Dianthus busambrae*. Pertanto emerge la grande importanza di realizzare banche dati molecolari che inglobino la più ampia biodiversità possibile con marcatori universali, ed in particolare l'importanza di realizzare un database dedicato delle principali specie floricole di interesse ornamentale, che possa supportare l'applicazione pratica del protocollo molecolare ai fini della tracciabilità e del monitoraggio da parte di organi di controllo.

Sebbene il DNA Barcoding può raggiungere livelli di risoluzione dell'80-90% i limiti nella discriminazione possono verificarsi in diverse famiglie, tra cui Ericaceae, Lamiaaceae, Orchidaceae, a causa della modesta distanza evolutiva che si può verificare tra specie strettamente correlate o che hanno avuto una divergenza recente, come si è verificato anche in questo studio (Hollingsworth et al., 2009; Ren et al., 2011; De Mattia et al., 2011; Xiang et al., 2011; Maia et al., 2012).

In alcuni casi l'approccio multi-locus può favorire un maggiore successo, tuttavia la valutazione di ulteriori regioni barcoding addizionali, per incrementare il successo della discriminazione, necessita uno studio ed una applicazione che si focalizzi di volta in volta sui singoli gruppi tassonomici di difficile discriminazione (Kress et al., 2005, 2007; Chen et al., 2010; Yao et al. 2010; Hollingsworth et al., 2011).

Il gruppo *Chamaerops*, nel nostro studio rappresentato dalle due varietà ufficialmente accettate, *humilis* ed *argentea*, più ulteriori due varietà commerciali, Vulcano ed Etna-star, sono state discriminate da tutte le specie ed i generi ad esse ritenute affini (*Rhapidophyllum* e *Trachycarpus*) ma non c'è stata discriminazione tra di loro. Solo per queste specie si è ritenuto opportuno iniziare un percorso di identificazione tramite microsatelliti (SSR) specifici, al fine di potere approfondire le relazioni tra individui a

scarsa variabilità genetica (studio in corso). L'adozione di un protocollo di analisi con microsatelliti, per quanto efficace nel discriminare a livelli subspecifici, essendo molto lungo e costoso, risulta infatti applicabile solo a singole specie.

5.6 DNA Barcoding nella tracciabilità delle specie vegetali ornamentali

Per i taxa discriminati a livello di specie con la metodologia del DNA Barcoding, che è stato l'oggetto di questo studio, sarà possibile avviare già un sistema di certificazione genetica e tracciabilità, che segue delle precise fasi all'interno di un protocollo di caratterizzazione molecolare standardizzato e riconosciuto a livello internazionale che segue le linee guida del CBOL Consortium for the barcode of life.

Il sistema di certificazione e tracciabilità, seguirà un percorso ben preciso (fig.31-32-33) che va dall'impiego dei protocolli del DNA Barcoding per la identificazione della specie, fino alla realizzazione di una etichetta dove saranno inclusi, oltre alle indicazioni generiche sulla specie, anche i relativi dati molecolari, che si traducono in un barcode che tramite la scansione con appositi scanner di codici a barre realizzati per questo scopo, permetterà con immediatezza di avere tutte le indicazioni sulla specie.

Fig.31 Schema della procedura di deposito e pubblicazione dei dati Barcode

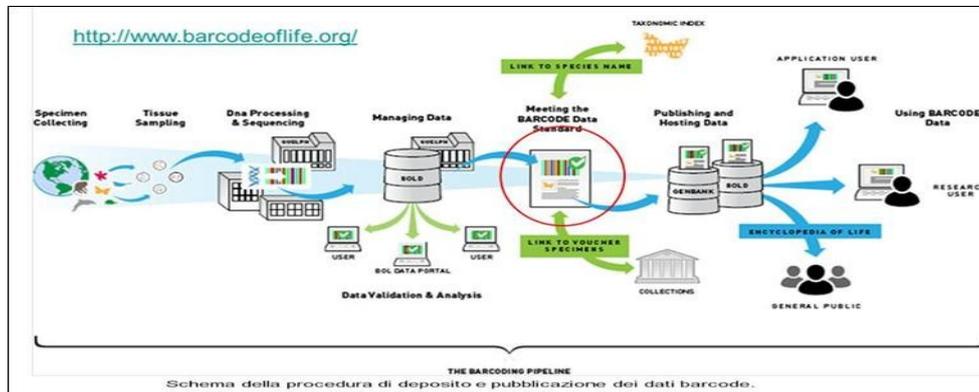


Fig.32 Percorso seguito per la caratterizzazione molecolare delle specie oggetto di studio (Azienda Faro) secondo gli standard internazionali CBOL

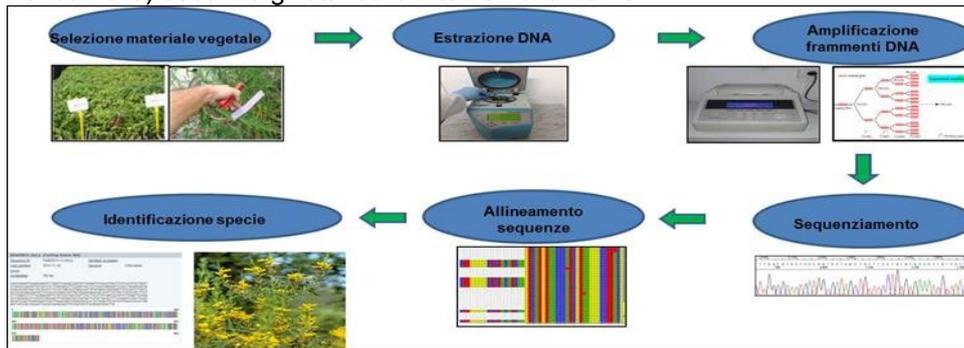


Fig.33 Etichetta per la tracciabilità da proporre alle aziende

<p>Genista madoniensis Raimondo Nome locale: Ginestra delle Madonie</p>		<p>DNA Barcoding Analysis</p>	
<p>Forma biologia : NP Pianta legnosa perenne con gemme poste con gemme poste tra 20 cm e 2m dal suolo</p>		<p>Materiale per DNA: foglie</p>	
<p>Corologia: endemico della Sicilia</p>		<p>Gene: rbcL</p>	
<p>Periodo di fioritura: Aprile-Maggio</p>		<p>SEQUENCE: rbcL (Funding Source: N/A)</p>	
<p>Azienda produttrice.....</p>		<p>Sequence ID: FMED013-12.rbcLa GenBank Accession:</p>	
<p>Lotto n°.....</p>		<p>Last Updated: 2014-11-30 Genome: Chloroplast</p>	
		<p>Locus: rbcL</p>	
		<p>Nucleotide: 582 bp</p>	
		<pre> CACCAGAGACTAAAGCAGTGT766GTCAAGCCG67TAAAGATTAAATGACTTATTACTCTGACT ATAAACCAAGATATCTGATCTTACAGCATCTCCGATACCTCCACCG66HTTCCCTGAGAGACG GTCCCGG67ACTGCGCAATCTCTACTG67ACATGACCACTGTGTGACCGACG66GCTTACCACTCTGATC GCTACAAGAGAGATGCTATCACATCGAGCC67TGTGAGAGAGAAATCAATTTATGCTATGCTATC CTTAGACCTTTTGAAGAG67TCTGTACTAATG7TACTTCCATGTAG67TAAATGTAT766TTCAGG CCCTG66GCTTACTGTGAGATTTGCAATCCCAATGCTATGTAHAATTTCAAG67TCCGCTCAAG GCATCCAGTTGAGAGATAAATGACCAAGTATG66CCTCCCTTATG66ATGCTACTATAAACAATTTG6 G6TTATCCCTAAGATACG67AG66CATTTATGATGTCTACG66GTG66CTG6 </pre>	

6 Conclusioni

In questa ricerca, le prestazioni dei principali marcatori Barcode, *rbcL* e *matK*, sono state valutate su un totale di 100 taxa vegetali, appartenenti a 20 famiglie diverse. I taxa sequenziati con successo per almeno uno dei marcatori considerati sono stati 98 ed il 61% dei taxa valutati sono stati totalmente discriminati a livello di specie o sottospecie. Considerando che il fallimento dei taxa è strettamente legato all'interno del genere, o addirittura al di sotto della specie, con bassissima divergenza evolutiva, questo risultato conferma la potenzialità dell'approccio barcoding per l'analisi rapida di campioni sconosciuti. Gruppi criptici trovati in questo studio hanno messo in evidenza i problemi tecnici già noti, correlati al basso livello di successo di amplificazione e sequenziamento del *matK*, ma, allo stesso tempo, la sua grande potenza di discriminazione rispetto a *rbcL*.

Pertanto, l'adozione dei core markers è apparso un buon compromesso. In alcuni casi l'approccio multi-locus e l'aggiunta del terzo marker *trnH-psbA* può favorire un maggiore successo, come in questo studio.

La valutazione di ulteriori regioni barcoding addizionali, può essere utile per incrementare il successo della discriminazione, ma necessita di uno studio ed una applicazione che focalizzi di volta in volta sui singoli gruppi tassonomici che hanno presentato problematiche con i core markers.

In effetti, se si considera che una delle più grandi sfide del barcoding è la mancanza di sequenze di riferimento disponibili per la comparazione che interessino tutti i marcatori candidati, un ampio campione di referenti relativo al taxa considerato, è necessario per validare l'accuratezza del metodo.

Emerge la grande importanza di realizzare banche dati molecolari che inglobino la più ampia biodiversità possibile con marcatori universali, ed in particolare l'importanza di realizzare un database dedicato delle principali specie floricole di interesse ornamentale, che possa supportare l'applicazione pratica del protocollo molecolare ai fini della tracciabilità e del monitoraggio degli scambi commerciali da parte di organi di controllo.

7 Bibliografia

- Asmussen,C.B., Dransfield,J., Deichmann,V., Barfod,A.S., Pintaud,J.C. and Baker,W.J. (2006) - *A new subfamily classification of the palm family (Arecaceae): Evidence from plastid DNA phylogeny*. Bot. J. Linn. Soc. 151 (1), 15-38
- Asmussen,C.B. and Chase,M.W. (2001) - *Coding and noncoding plastid DNA in palm systematics*. Am. J. Bot. 88 (6), 1103-1117
- Barrett R. and Hebert P.D.N., (2005) - *Identifying spiders through DNA barcodes*. Canadian Journal of Zoology 83: 481–491.
- Bhatramakki D., Rafalski A., (2001) - *Discovery and application of single nucleotide polymorphism markers in plants*. Plant genotyping: the DNA fingerprinting of plants. Edited by Henry RJ. Wallingford: CAB International; p. 179-191.
- Beckmann J.S., Weber J.L., (1992) - *Survey of human and rat microsatellites*. Genomica 12:627-631.
- Chase M.W., Cowan R.S., Hollingsworth P.M., van den Berg C., Madriñán S., Petersen G., Seberg O.,Jørgensen T., Cameron K.M., Carine M., Pedersen N., Hedderson T.A.J., Conrad F., Salazar G.A.,Richardson J.E., Hollingsworth M.L, Barraclough T.G., Kelly L., Wilkinson M., (2007) - *A proposal for a standard protocol to barcode all land plants*. Taxon 56 (2): 295–299.
- Chen S.L., Yao H., Han J.P., Liu C., Song J.Y., Shi L.C., Zhu Y.J., Ma X.Y., Gao T., Pang X.H., Luo K., Li Y.,Li X.W., Jia X.C., Lin Y.L., Leon C., (2010).- *Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species*. PLoS ONE 5: e8613, doi:10.1371/journal.pone.0008613.

- Collins F.S., Brooks L.D. and Chakravarti A., (1998) - *A DNA Polymorphism Discovery Resource for Research on Human Genetic Variation*. *Genome Research* 8:1229–1231
- Costion C., Ford A., Cross H., Crayn D., Harrington M., Lowe A., (2011) - *Plant DNA barcodes can accurately estimate species richness in poorly known floras*. *PLoS ONE* 6 (11): e26841
- De Mattia F., Bruni I., Galimberti A., Cattaneo F., Casiraghi M., Labra M., (2011) - *A comparative study of different DNA barcoding markers for the identification of some members of Lamiaceae*. *Food Res Int* 44: 693–702.
- Doyle J.J., Doyle J.L., (1987) - *A rapid DNA isolation procedure for small quantities of fresh leaf tissue*. *Phytochem Bull* 19:11–15.
- Dunning L.T., Savolainen V., (2010) - *Broad-scale amplification of matK for DNA barcoding plants, a technical note*. *Botanical Journal of the Linnean Society* 164: 1-9
- Dupont S., Wilson K., Obst M., Sköld H., Nakano H. & Thorndyke M.C., (2007) - *Marine ecological genomics: when genomics meets marine ecology*. *Marine ecology progress series* 332: 257-273.
- Fazekas A.J., Burgess K.S., Kesanakurti P.R., Graham S.W., Newmaster S.G., Husband B.C., Percy D.M., Hajibabaei M. & Barrett S.C.H., (2008) - *Multiple Multilocus DNA Barcodes from the Plastid Genome Discriminate Plant Species Equally Well*. *PLoS ONE* 3(7): e2802.
- Fazekas A.J., Kesanakurti P.R., Burgess K.S., Percy D.M., Graham S.W., Barret S.C.H., Newmaster S.G., Hajibabaei M. & Husband B.C., (2009) - *Are plant species inherently harder to discriminate than animal species using DNA barcoding markers*. *Molecular Ecology Resources* 9: 130–139.

- Gao T., Sun Z., Yao H., Song J., et al., (2011) - *Identification of Fabaceae plants using the DNA barcode matK*. *Planta Med.* 77: 92-94.
- Gao T. and Chen S.L., (2009) - *Authentication of the medicinal plants in Fabaceae by DNA barcoding technique*. *Planta Med.* 75: 417.
- Gao T., Yao H., Song J., Zhu Y., Liu C., Chen S., (2010) - *Evaluating the feasibility of using candidate DNA barcodes in discriminating species of the large Asteraceae family*. *BMC Evol. Biol.*, 10: 1-7.
- Hebert P.D.N., Cywinska A., Ball S.L. and deWaard J.R., (2003) - *Biological identifications through DNA barcodes*. *Proceedings of the Royal Society of London* 270: 313-321.
- Hebert P.D.N., Penton E.H., Burns J.M., Janzen D.H. and Hallwachs W., (2004a) - *Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator**. *Proceedings of the National Academy of Sciences USA* 101: 14812–14817.
- Hebert P.D.N., Stoeckle M.Y., Zemplak T.S. and Francis C.M., (2004b) - *Identification of birds through DNA barcodes*. *PLoS Biology* 2: 312.
- Hogg I.D. and Hebert P.D.N., (2005) - *Biological identifications of springtails (Hexapoda: Collembola) from the Canadian arctic, using mitochondrial barcodes*. *Canadian Journal of Zoology* 82: 749–754.
- Hollingsworth P.M., Graham S.W. & Little D.P., (2011) - *Choosing and Using a Plant DNA Barcode*. *PLoS ONE* 6(5): e19254.
- Hollingsworth P.M. et al. CBOL Plant Working Group. 2009. - *A DNA barcode for land plants*. *Proc Nat Acad Sci United States Amer* 106: 12794–12797.

- Janzen D.H., Hajibabaei M., Burns J.M., Hallwachs W., Remigio E. and Hebert P.D.N., (2005) - *Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding*. Philosophical Transactions of the Royal Society of London Series B: Biological Sciences 360: 1835–1845.
- Jukes T.H. and Cantor C.R., (1969) - *Evolution of protein molecules*. In Munro, H.N. (ed), Mammalian Protein Metabolism. Academic Press, New York, NY, pp. 21-132.
- Khew G.S., Chia T.F., (2011) - *Parentage determination of Vanda Miss Joaquim (Orchidaceae) through two chloroplast genes *rbcl* and *matK**. AoB PLANTS. plr018: 1–12
- Kimura M., (1980) - *A simple method for estimating evolutionary rates of base substitution through comparative studies of nucleotide sequences*. J. of Molecular Evolution 16:111-120
- Kress W.J., Wurdack K.J., Zimmer E.A., Weigt L.A., Janzen D.H., (2005) - *Use of DNA barcodes to identify flowering plants*. Proc Natl Acad Sci 102(23): 8369–8374.
- Kress W.J., Erickson D.L., (2007) - *A two-locus global DNA barcode for land plants: The coding *rbcl* gene complements the non-coding *trnH-psbA* spacer region*. PLoS ONE 2:e508.
- Lagercrantz U., Hellegren H., Andersson L. (1993) - *The abundance of various polymorphic microsatellites motifs differs between plants and vertebrates*. Nucleic Acids Res 21(5):1111-1115
- Lahaye R., Bank van der M., Bogarin D., Warner J., Pupulin F., Gigot G., Maurin O., Duthoit S., Barraclough T.G., Savolainen V., (2008) - *DNA barcoding the floras of biodiversity hotspots*. Proc Natl Acad Sci 105(8): 2923-2928.

- Lanave C., Preparata G., Saccone C., Serio G., (1984) - *A new method for calculating evolutionary substitution rates*. J. of Molecular Evolution 20(1):86-93.
- Madesis P., Ganopoulos I., Ralli P. and Tsaftaris A., (2012) - *Barcoding the major Mediterranean leguminous crops by combining universal chloroplast and nuclear DNA sequence targets*. Genet. Mol. Res. 11 (3): 2548-2558
- Maia V.H., Souza da Mata C., Ozório Franco L., Aires Cardoso M., Sodr  Cardoso S.R., Silva Hemerly A., Cavalcanti Gomes Ferreira P., (2012) - *DNA barcoding Bromeliaceae: achievements and pitfalls*. PLoS one 7(1):e2987.
- Meyer C.P. and Paulay G., (2005) - *DNA Barcoding: Error Rates Based on Comprehensive Sampling*. PLoS Biology 3(12): 422.
- Morgante M., Olivieri A.M., (1993) - *PCR-amplified microsatellites as markers in plant genetics*. Plant J. 3(1):175-182
- Pegg G.G., Sinclair B., Briskey L., Aspden W.J., (2006) - *MtDNA barcode identification of fish larvae in the southern Great Barrier Reef, Australia*. Sci. Mar. 70(S2): 7–12.
- Quarta R., Dettori M.T., Verde I., Sartori A., (1998) - *Mappa genomica del pesco e individuazione di marcatori associati a caratteri agronomici utili*. Atti IV Giornate Scientifiche SOI, p. 529.
- Rafalski A., (2002) - *Applications of single nucleotide polymorphisms in crop genetics*. Current Opinion in Plant Biology, Volume 5, Issue 2, Pages 94–100.
- Randi E., C. Tabarroni e S. Rimondi, (2002) - *Genetica forense in applicazione della Convenzione di Washington CITES*. Quad. Cons. Natura, 12, Min. Ambiente - Ist. Naz. Fauna Selvatica

- Ren H., Lu L., Wang H., Li D.Z., (2011) - *DNA barcoding of Gaultheria L. in China (Ericaceae: Vaccinioideae)*. J Syst Evol 49: 411–424.
- Ratnasingham S., Hebert P.D.N., (2007) - *BOLD: The Barcode of Life Datasystem (www.barcodinglife.org)*. Mol Ecol Notes 7: 355–64.
- Saccone C., Lanave C., Pesole G. and Preparata G., (1990) - *The Influence of Base Composition on Quantitative Estimates of Gene Evolution*. Methods in Enzymology 183: 570-583
- Saitou N. and Nei M., (1987) - *The neighbor-joining method: a new method for reconstructing phylogenetic trees*. Mol Biol Evol 4 (4): 406-425.
- Sanger F., Nicklen S., and Coulson A.R., (1977) - *DNA sequencing with chain-terminating inhibitors*. Biochemistry, Proc. Natl. Acad. Sci. USA, Vol 74, 12, pp. 5463-5467
- Saunders G.W., (2005) - *Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications*. Philos Trans R Soc Lond B Biol.Sci 360(1462):1879-1888.
- Saunders G.W. and Kucera H., (2010) - *An evaluation of rbcL, tufA, UPA, LSU and ITS as DNA barcode markers for the marine green macroalgae*. Cryptogamie, Algologie 31(4): 487-528.
- Sneath P.H.A. & Sokal R.R., (1973) - *Numerical taxonomy : The principles and practice of numerical classification*. (Books in Biology). Published by W. H. Freeman, San Francisco. xv + 573 p.
- Tamura K. et al., (2013) - *MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0*. Molecular Biology and Evolution 30: 2725-2729.

- Tautz D., Arctander P., Minelli A., Thomas R.H., Vogler A.P., (2003) - *A plea for DNA taxonomy*. Trends Ecol. Evol. 18(2): 70–74.
- Varshney R.K., Graner A., Sorrells M.E., (2005) - *Genic microsatellite markers in plants: features and applications*. Trends Biotechnol. 23(1):48-55.
- Vences M., Thomas M., Bonett R.M., Vieites D.R., (2005) - *Deciphering amphibian diversity through DNA barcoding: chances and challenges*. Philos T R Soc B 360: 1859-1868
- Vere de N., Rich T.C.G., Ford C.R., Trinder S.A., Long C., Moore C.W., Satterthwaite D., Davies H.,Allainguillaume J., Ronca S., Tatarinova T., Garbett H., Walker K., Wilkinson M.J, (2012 – *DNA barcoding the native flowering plants and conifers of Wales*. PLoS ONE 7 (6). e37945
- Wang Z., Weber J.L., Zhong G., Tanksley S.D., (1994) - *Survey of plant short tandem DNA repeats*. Theoretical and Applied Genetics 88: 1-6
- Ward R.D., Zemlak T.S., Innes B.H., Last P.R. and Hebert P.D.N., (2005) - *DNA barcoding Australia's fish species*. Philosophical Transactions of the Royal Society of London Series B: Biological Sciences 360: 1847–1857.
- Xiang X.G., Hu H., Wang W., Jin X.H, (2011) - *DNA barcoding of the recently evolved genus *Holcoglossum* (Orchidaceae: Aseridinae): a test of DNA barcode candidates*. Mol Ecol Res 11:1012–1021.
- Yao H, Song J, Liu C, Luo K, Han J, Li Y, Pang X, Xu H, Zhu Y, Xiao P, Chen S. (2010) - *Use of ITS2 region as the universal DNA barcode for plants and animals*. PLoS one 5: e13102.

Zetezcche et. al -www.dnabank-network.eu

Legambiente- www.legambiente.it

Consortium for the Barcode of Life – CBOL - www.barcodinglife.org

Catalogue of Life- www.catalogueoflife.org

CITES- www.cites.org

Barcode of Life Data Systems-BOLD- www.boldsystems.org

Decreto Legislativo 19 Maggio 2000 n.151

Decreto n.33043 del 9 Agosto 2000

8 APPENDICE

Tab.14 Percentuale di identità genetica tra i taxa investigati e le sequenze referenti più affini per il marcatore *rbcl*

Famiglia	Taxa	Referente GenBank ID	Identità %	Gap
ARECACEAE	<i>Chamaerops humilis</i> var. <i>humilis</i> (F.C4)	<i>Chamaerops humilis</i> AY012456	100	0
		<i>Rhapidophyllum hystri</i> AJ404753	100	0
		<i>Guihaia argyrata</i> AJ404755	100	0
		<i>Trachycarpus martianus</i> JX90325	100	0
		<i>Trachycarpus fortunei</i> AY012460	100	0
		<i>Chamaerops humilis</i> var. <i>argentea</i> F.C5 (questo studio)	100	0
		<i>Chamaerops humilis</i> "Vulcano" F.C6 (questo studio)	100	0
		<i>Chamaerops humilis</i> "Etna star" F.C7 (questo studio)	100	0
	<i>Chamaerops humilis</i> var. <i>argentea</i> (F.C5)	<i>Chamaerops humilis</i> AY012456	100	0
		<i>Chamaerops humilis</i> var. <i>argentea</i> F.C5 (questo studio)	100	0
		<i>Chamaerops humilis</i> "Vulcano" F.C6 (questo studio)	100	0
		<i>Chamaerops humilis</i> "Etna star" F.C7 (questo studio)	100	0
		<i>Rhapidophyllum hystri</i> AJ404753	100	0
		<i>Guihaia argyrata</i> AJ404755	100	0
		<i>Trachycarpus martianus</i> JX90325	100	0
		<i>Trachycarpus fortunei</i> AY012460	100	0
	<i>Chamaerops humilis</i> "Vulcano" (F.C6)	<i>Chamaerops humilis</i> var. <i>humilis</i> F.C4 (questo studio)	100	0
		<i>Chamaerops humilis</i> var. <i>argentea</i> (F.C5) (questo studio)	100	0
		<i>Chamaerops humilis</i> "Etna star" (F.C7) (questo studio)	100	0
		<i>Rhapidophyllum hystri</i> AJ404753	100	0
		<i>Guihaia argyrata</i> AJ404755	100	0
		<i>Trachycarpus martianus</i> JX90325	100	0
		<i>Trachycarpus fortunei</i> AY012460	100	0
		<i>Chamaerops humilis</i> "Etna star" (F.C7)	<i>Chamaerops humilis</i> var. <i>humilis</i> F.C4 (questo studio)	100
	<i>Chamaerops humilis</i> var. <i>argentea</i> (F.C5) (questo studio)		100	0
	<i>Chamaerops humilis</i> "Vulcano" (F.C6)		100	0
	<i>Rhapidophyllum hystri</i> AJ404753		100	0
	<i>Guihaia argyrata</i> AJ404755		100	0
	<i>Trachycarpus martianus</i> JX90325		100	0
	<i>Trachycarpus fortunei</i> AY012460		100	0
	<i>Brahea armata</i> (F.B4)		<i>Brahea dulcis</i> AJ829853	100
		<i>Serenoa repens</i> AJ404760	100	0
<i>Washingtonia robusta</i> GU135278		100	0	
<i>Brahea berlandieri</i> AM110198		100	0	
<i>Brahea aculeata</i> KF746462		100	0	

	<i>Brahea calcarea</i> JX571788	100	0
<i>Acoelorrhaphe wrightii</i> (F.A11)	<i>Brahea dulcis</i> AJ829853	99	0
<i>Arenga engleri</i> (F.A7)	<i>Arenga caudata</i> JF344799	100	0
	<i>Arenga pinnata</i> AY012478	99	0
	<i>Arenga brevipes</i> JF344795	99	0
	<i>Caryota obtusa</i> JF344839	99	0
<i>Caryota urens</i> (F.C20)	<i>Arenga caudata</i> JF344799	99	0
	<i>Caryota ophiopellis</i> AJ404791	99	0
<i>Howea forsteriana</i> (F.H1)	<i>Howea belmoreana</i> GQ248620	100	0
	Altri 40 ID <i>Howea</i> spp. non distinti	100	0
<i>Livistona chinensis</i> (F.L1)	<i>Livistona speciosa</i> AY012463	100	0
	<i>Livistona benthamii</i> KF496442	100	0
	<i>Livistona humilis</i> KF746461	100	0
	<i>Washingtonia filifera</i> AY012465	100	0
	<i>Colpothrinax wrightii</i> AJ829862	99	0
	<i>Washingtonia robusta</i> GU135278	99	0
	<i>Brahea berlandieri</i> AM110198; <i>Brahea aculeata</i> KF746462; <i>Brahea calcarea</i> JX571788; <i>Brahea dulcis</i> AJ829853	99	0
<i>Washingtonia robusta</i> (F.W1)	<i>Brahea berlandieri</i> AM110198; <i>Brahea aculeata</i> KF746462; <i>Brahea calcarea</i> JX571788; <i>Brahea dulcis</i> AJ829853	100	0
	<i>Brahea armata</i> F.B4 (questo studio)	100	0
	<i>Serenoa repens</i> AJ404760	100	0
	<i>Washingtonia filifera</i> AY012465	99	0
<i>Washingtonia filifera</i> (F.W2)	<i>Livistona chinensis</i> AJ404757	100	0
	<i>Livistona speciosa</i> AY012463	100	0
	<i>Livistona benthamii</i> KF496442	100	0
	<i>Livistona humilis</i> KF746461	100	0
	<i>Washingtonia robusta</i> GU135278	99	0
<i>Trachycarpus fortunei</i> (F.T1)	<i>Trachycarpus martianus</i> JX903256	100	0
	<i>Chamaerops humilis</i> AY012456	100	0
	<i>Rhapidophyllum hystrix</i> AJ404753	100	0
	<i>Guihaia argyrata</i> AJ404755	100	0
<i>Phoenix roebelenii</i> (F.P7)	<i>Phoenix canariensis</i> AM110194	100	0
	<i>Phoenix dactylifera</i> AY012468	100	0
	<i>Phoenix reclinata</i> AJ404767	100	0
	<i>Sabal bermudana</i> AJ404766	99	0
<i>Phoenix canariensis</i> (F.P2)	<i>Phoenix canariensis</i> AM110194	100	0
	<i>Phoenix dactylifera</i> AY012468	100	0
	<i>Phoenix reclinata</i> AJ404767	100	0
	<i>Sabal bermudana</i> AJ404766	99	0
<i>Phoenix dactylifera</i> (F.P8)	<i>Phoenix canariensis</i> AM110194	100	0
	<i>Phoenix dactylifera</i> AY012468	100	0
	<i>Phoenix reclinata</i> AJ404767	100	0
	<i>Sabal bermudana</i> AJ404766	99	0
<i>Phoenix reclinata</i> (F.P9)	<i>Phoenix canariensis</i> AM110194	100	0
	<i>Phoenix dactylifera</i> AY012468	100	0
	<i>Phoenix reclinata</i> AJ404767	100	0
	<i>Sabal bermudana</i> AJ404766	99	0
<i>Sabal minor</i> (F.S14)	<i>Dictyocaryum lamarckianum</i> AM110204	100	0
	<i>Mauritia flexuosa</i> AY012473	100	0
	<i>Sabal palmetto</i> JX571889	100	0
	<i>Sabal etonia</i> EF590572	100	0
	<i>Sabal bermudana</i> AJ404766	100	0

		<i>Brahea dulcis</i> AJ829853; <i>Brahea aculeata</i> KF746462; <i>Serenoa repens</i> AJ404760; <i>Washingtonia robusta</i> GU135278	99	0
	<i>Sabal palmetto</i> (F.S3)	<i>Sabal bermudana</i> AJ404766	100	0
		<i>Sabal etonia</i> EF590572	100	0
		<i>Mauritia flexuosa</i> AY012473	100	0
		<i>Sabal domingensis</i> AY012467	99	0
	<i>Bismarckia nobilis</i> (F.B3)	<i>Hyphaene thebaica</i> AJ404770	100	0
		<i>Satranala decussilvae</i> AJ404771	100	0
		<i>Hyphaene coriacea</i> AM903195	100	0
	<i>Butia capitata</i> (F.B2)	<i>Butia yatay</i> AB088827	100	0
		<i>Butia eriospatha</i> AY044632	100	0
		<i>Jubaea chilensis</i> AJ829875	100	0
		<i>Syagrus romanzoffiana</i> KF981205	99	0
		<i>Cocos nucifera</i> GQ436759	99	0
	<i>Syagrus romanzoffiana</i> / <i>Arecastrum romanzoffiana</i> (F.A8)	<i>Cocos nucifera</i> GQ436759	100	0
		<i>Syagrus glaucescens</i> AY044634	100	0
		<i>Syagrus smithii</i> AJ404827	100	0
		<i>Polyandrococos caudescens</i> AJ829902; <i>Parajubaea torallyi</i> AJ829891; <i>Scheelea butyracea</i> AY044636; <i>Lytocaryum weddellianum</i> AY044633; <i>Attalea allenii</i> AJ404829; <i>Allagoptera arenaria</i> AJ404828	100	0
		<i>Butia capitata</i> JX903252	99	0
		<i>Bismarckia nobilis</i> JX088664	99	0
	<i>Trithrinax campestris</i> (F.T4)	<i>Trithrinax acanthocoma</i> AY298852	100	0
		<i>Trithrinax radiata</i> AY012459	99	0
ASTERACEAE	<i>Anthemis cupaniana</i> (F.A2)	<i>Anthemis arvensis</i> HM849778	99	0
		<i>Anthemis cotula</i> HM849779	99	0
		<i>Anthemis ismelia</i> A7 (questo studio)	99	0
		rbCL assente <i>Anthemis maritima</i> rbCL assente <i>Anthemis aetnensis</i>		
	<i>Centaurea sphaerocephala</i> F.C14	<i>Centaurea behen</i> KC589816	99	0
	<i>Helichrysum hyblaenum</i> H6	<i>Helichrysum scandens</i> H7 (questo studio)	100	0
		<i>Helichrysum nebrodense</i> H5 (questo studio)	99	0
		<i>Helichrysum stoechas</i> GU817766	99	0
		<i>Helichrysum microphyllum</i> HF585090	99	0
	<i>Helichrysum nebrodense</i> H5	<i>Helichrysum hyblaenum</i> H6 (questo studio)	99	0
		<i>Helichrysum scandens</i> H7 (questo studio)	99	0
		<i>Helichrysum microphyllum</i> HF585090	99	0
		<i>Helichrysum stoechas</i> GU817766	99	0
	<i>Helichrysum scandens</i> H7	<i>Helichrysum hyblaenum</i> H6 (questo studio)	100	0

		<i>Helichrysum nebrodense</i> H5 (questo studio)	99	0
		<i>Helichrysum microphyllum</i> HF585090	99	0
		<i>Helichrysum stoechas</i> GU817766	99	0
	<i>Ptilostemon greuteri</i> F.P5	<i>Atractylodes lancea</i> KM192089	100	0
		<i>Ptilostemon diacanthus</i> KC589884	99	0
		<i>Ptilostemon hispanicus</i> KC589885	99	0
		<i>Ptilostemon afer</i> KC589883	99	0
		<i>Ptilostemon abylenis</i> KC589882	99	0
		<i>Jacobaea gibbosa</i> F.J1	<i>Artemisia absinthium</i> HE963336	100
	<i>Artemisia gmelinii</i> GQ436432		100	0
	<i>Artemisia lavandulifolia</i> GQ436484		100	0
	<i>Artemisia sacrorum</i> JQ173397		100	0
	<i>Artemisia vulgaris</i> KF589298		100	0
	<i>Artemisia frigida</i> NC 020607		100	0
	<i>Artemisia argyi</i> GQ436428		99	0
	<i>Jacobaea aquatica</i> HM850080		99	0
	<i>Jacobaea vulgaris</i> JN891061		99	0
	<i>Jacobaea maritima</i> HM850081		99	0
	<i>Pullenia maritima</i> F.P3	<i>Dittrichia viscosa</i> GU817768	99	0
		<i>Pulicaria paludosa</i> KF724264	99	0
	<i>Hieracium cophanense</i> H2	<i>Hieracium pignattianum</i> H9 (questo studio)	100	0
		<i>Hieracium symphytifolium</i> H15 (questo studio)	99	0
		<i>Hieracium crinitum</i> H16 (questo studio)	99	0
		<i>Hieracium venosum</i> KF613085	99	0
		<i>Pilosella peleteriana</i> JN890666	99	0
	<i>Jacobaea ambigua</i> F.S4	<i>Jacobaea vulgaris</i> GU817769	100	0
		<i>Senecio aquaticus</i> AY395561	100	0
		<i>Jacobaea aquatica</i> HM850080	100	0
		<i>Jacobaea maritima</i> HM850081	100	0
	<i>Senecio candidus</i> F.S5	<i>Jacobaea vulgaris</i> GU817769	100	0
		<i>Jacobaea maritima</i> HM850081	100	0
		<i>Jacobaea aquatica</i> HM850080	100	0
		<i>Senecio aquaticus</i> AY395561	99	0
FABACEAE	<i>Spartium junceum</i> F.S9	<i>Genista thyrrhena</i> G1 (questo studio)	99	0
		<i>Genista demarcoi</i> G4 (questo studio)	99	0
		<i>Petteria ramentacea</i> Z70091	99	0
	<i>Ceratonia siliqua</i> F.C1	<i>Balsamocarpon brevifolium</i> AY308524	99	0
	<i>Genista thyrrhena</i> G1	<i>Genista demarcoi</i> G4 (questo studio)	100	0
	<i>Genista madoniensis</i> G2	<i>Genista cupani</i> G5 (questo studio)	99	0
	<i>Genista aristata</i> G3	<i>Genista tridentata</i> Z70104	100	0
	<i>Genista demarcoi</i> G4	<i>Genista thyrrhena</i> G1 (questo studio)	100	0
	<i>Genista cupanii</i> G5	<i>Genista madoniensis</i> G2u (questo studio)	99	0
<i>Genista aetnensis</i> G6	<i>Genista thyrrhena</i> G1 (questo studio)	99	0	
		<i>Genista demarcoi</i> G4 (questo studio)	99	0
EUPHORBIACEAE	<i>Euphorbia bivonae</i> Eu8	<i>Euphorbia pekinensis</i> GU441779	99	0
	<i>Euphorbia ceratocarpa</i> Eu1	<i>Euphorbia pekinensis</i> GU441779	99	0
	<i>Euphorbia characias</i> E2	<i>Euphorbia characias</i> subsp. <i>wulfenii</i> HE963471	100	0
		<i>Euphorbia amygdaloides</i> KF408913	100	0
	<i>Euphorbia dendroides</i> E4	<i>Euphorbia dendroides</i> JN249291	100	0
<i>Euphorbia amygdaloides</i> KF408913		99	0	

	<i>Euphorbia helioscopia</i> Eu7	<i>Euphorbia peplus</i> HE963481	99	0
	<i>Euphorbia myrsinites</i> E6	<i>Euphorbia pithyusa</i> subsp. <i>cupanii</i> Eu9 (questo studio)	99	0
	<i>Euphorbia meuselii</i> E5	<i>Euphorbia amygdaloides</i> KF408913	100	0
		<i>Euphorbia bourgaeana</i> KJ595644	100	0
		<i>Euphorbia characias</i> KF409047	100	0
	<i>Euphorbia pithyusa</i> subsp. <i>cupanii</i> Eu9	<i>Euphorbia rauhii</i> JQ952245	99	0
	<i>Euphorbia amygdaloides</i> Eu11	<i>Euphorbia meuselii</i> E5 (questo studio); <i>Euphorbia characias</i> KF409047; <i>Euphorbia bourgaeana</i> KJ595644	100	0
ROSACEAE	<i>Rosa sempervirens</i> R3	<i>Rosa laevigata</i> JN407360	100	0
		<i>Rosa multiflora</i> var. <i>cathayensis</i> JQ436573	100	
		<i>Rosa chinensis</i> GU363820	100	0
		<i>Rosa odorata</i> var. <i>gigantea</i> KF753637	100	
	<i>Rosa canina</i> R1	<i>Rosa corymbifera</i> R10 (questo studio)	100	0
		<i>Rosa arvensis</i> JN893237.	100	0
		<i>Rosa stylosa</i> JN891155.	100	0
		<i>Rosa obtusifolia</i> JN893206.	100	0
		<i>Rosa spinosissima</i> JN893205	99	0
		<i>Rosa sicula</i> R6	<i>Rosa agrestis</i> HE963632.	99
		<i>Rosa micrantha</i> JN891902	99	0
	<i>Rosa corymbifera</i> R10	<i>Rosa canina</i> FN689381	100	0
		<i>Rosa canina</i> R1 (questo studio)	100	0
		<i>Rosa arvensis</i> JN893237.	100	0
<i>Rosa stylosa</i> JN891155.		100	0	
<i>Rosa obtusifolia</i> JN893206.		100	0	
	<i>Rosa spinosissima</i> JN893205	99	0	
BRASSICACEAE	<i>Brassica insularis</i> B2	<i>Brassica napus</i> JF807907	99	0
	<i>Brassica villosa</i> subsp. <i>tinei</i> B4	<i>Brassica napus</i> JF807907	100	0
		<i>Brassica oleracea</i> M88342	100	0
	<i>Brassica rupestris</i> subsp. <i>hispida</i> B5	<i>Brassica napus</i> JF807907	100	0
<i>Brassica oleracea</i> M88342		100	0	
ACANTHACEAE	<i>Acanthus mollis</i> F.A12	<i>Acanthus montanus</i> L12592	98	0
XANTHORRHOEACEAE	<i>Aloe plicatilis</i> F.A5	<i>Aloe x spinosissima</i> KF733455	100	0
	<i>Aloe vera</i> F.A9	<i>Aloe x spinosissima</i> KF733455	99	0
	<i>Aloe arborescens</i> F.A10	<i>Aloe comosa</i> JQ024498	100	0
		<i>Aloe x spinosissima</i> KF733455	100	0
LILIACEAE	<i>Tulipa radii</i> T4	<i>Tulipa lanata</i> JN417437	99	0
AMARANTHACEAE	<i>Diotis maritima</i> (<i>Achillea maritima</i>) F.D1	<i>Achillea millefolium</i> L13641	99	0
IRIDACEAE	<i>Iris pseudopumila</i> F.11	<i>Iris germanica</i> JF942052	99	0
		<i>Iris subflora</i> rbcl assente		
	<i>Iris germanica</i> F.15	<i>Iris pseudopumila</i>	99	0
		<i>Iris pallida</i> rbcl assente		
LAMIACEAE	<i>Lavandula angustifolia</i> F.L4	<i>Lavandula stoechas</i> HQ902787	99	0
	<i>Lavandula stoechas</i> F.L5	<i>Lavandula angustifolia</i> HE963537	99	0
	<i>Salvia leucantha</i> F.S12	<i>Salvia officinalis</i> HE963643	99	0
		<i>Salvia farinacea</i> JQ933993	99	0
	<i>Salvia officinalis</i> F.S11	<i>Salvia sclarea</i> KC473332; <i>Salvia verbenaca</i> HM850333; <i>Salvia fruticosa</i> M590078; <i>Salvia x superba</i> HM590074; <i>Salvia staminea</i> AY570449; <i>Salvia texana</i> AY570447; <i>Salvia summa</i> AY570444; <i>Salvia roemeriana</i> AY570438; <i>Salvia pentstemonoides</i> AY570434; <i>Salvia palaestina</i> AY570433; <i>Salvia lyrata</i> AY570426; <i>Salvia henryi</i> AY570420;	100	0

		<i>S.indica</i> Z37443; <i>S.canariensis</i> Z37439; <i>S.bucharica</i> Z37438		
		<i>Perovskia atriplicifolia</i> AY570391	100	0
		<i>Salvia viscosa</i> AY570451	99	0
	<i>Rosmarinus officinalis</i> F.R4	<i>Salvia officinalis</i> HE963643	99	0
	<i>Sideritis italica</i> F.S6	<i>Sideritis syriaca</i> AF501999	99	0
STRELITZIACEAE	<i>Strelitzia augusta</i> F.S1	<i>Strelitzia nicolai</i> F.S2 (questo studio)	100	0
		<i>Strelitzia nicolai</i> JX572998	100	0
		<i>Strelitzia reginae</i> AM235157	99	0
	<i>Strelitzia nicolai</i> F.S2	<i>Strelitzia augusta</i> F.S1 (questo studio)	100	0
		<i>Strelitzia nicolai</i> JX572998	100	0
	<i>Strelitzia reginae</i> F.S8	<i>Strelitzia nicolai</i> JX572998	99	0
		<i>Strelitzia augusta</i> F.S1 (questo studio)	99	0
CISTACEAE	<i>Cistus salviifolius</i> F.C2	<i>Cistus clusii</i> FJ225879; <i>Cistus libanotis</i> FJ225870, <i>Cistus munbyi</i> FJ225877; <i>Cistus albidus</i> F.C11(questo studio)	100	0
	<i>Cistus x pulverulentus</i> F.C3	<i>Cistus albidus</i> FJ225860; <i>Cistus crispus</i> FJ225882; <i>Cistus heterophyllus</i> FJ225868; <i>Cistus chinamadensis</i> subsp. <i>gomeræ</i> KJ595624; <i>Cistus symphytifolius</i> FJ225864; <i>Cistus skanbergii</i> F.C12 (questo studio)	100	0
	<i>Cistus albidus</i> F.C11	<i>Cistus clusii</i> FJ225879; <i>Cistus libanotis</i> FJ225870; <i>Cistus munbyi</i> FJ225877	100	0
	<i>Cistus skanbergii</i> F.C12	<i>Cistus albidus</i> FJ225860; <i>Cistus crispus</i> FJ225882; <i>Cistus heterophyllus</i> FJ225868; <i>Cistus chinamadensis</i> subsp. <i>gomeræ</i> KJ595624; <i>Cistus symphytifolius</i> FJ225864	100	0
MYRTACEAE	<i>Myrtus luma</i> FM1	<i>Myrtus communis</i> HF585092	100	0
		<i>Myrtus communis</i> AF294254	99	0
		<i>Myrtus communis</i> subsp. <i>Communis</i> HE963567	99	0
	<i>Myrtus communis</i> F.M5	<i>Myrtus luma</i> FM1 (questo studio)	100	0
		<i>Siphoneugena guilfoyleana</i> KF981269	99	0
<i>Metrosideros excelsa</i> F.M2	<i>Metrosideros angustifolia</i> JX572770	99	0	
CYCADACEAE	<i>Cycas revoluta</i> F.C8	<i>Cycas taitungensis</i> AP009339	100	0
		<i>Cycas micholitzii</i> GQ436680	100	0
		<i>Cycas wadei</i> AF394341	100	0
		<i>Cycas panzhihuaensis</i> AB917042	100	0
	<i>Cycas circinalis</i> F.C19	<i>Cycas micronesica</i> EU016864	99	0
		<i>Cycas revoluta</i> JN867588	99	0
ERICACEAE	<i>Erica sicula</i> E7	<i>Erica peduncularis</i> E9 (questo studio)	99	0
		<i>Erica multiflora</i> E10 (questo studio)	99	0
		<i>Erica sicula</i> AF419823	99	0
	<i>Erica peduncularis</i> E9	<i>Erica multiflora</i> E10 (questo studio)	100	0
	<i>Erica multiflora</i> E10	<i>Erica peduncularis</i> E9	100	0
<i>Arbutus unedo</i> F.A1	<i>Arbutus canariensis</i> L12597	99	0	
CONVOLVULACEAE	<i>Calystegia soldanella</i> F.C13	<i>Calystegia sepium</i> AY100992	100	0
TAMARICACEAE	<i>Tamarix gallica</i> F.T5	<i>Tamarix aphylla</i> KC505173	99	0
CARYOPHILLACEAE	<i>Dianthus rupicola</i> subsp. <i>aeolicus</i> D1	<i>Dianthus caryophyllus</i> M77699	100	0
		<i>Dianthus deltoides</i> JN892692	100	0
		<i>Dianthus armeria</i> JN892039	100	0
		D2 (questo studio); D3 (q.studio); D4 (q.studio); D5 (q.studio); D6 (q.studio)	100	0

	<i>Dianthus seguieri</i> rbcl assente		
	<i>Dianthus furcatus</i> rbcl assente		
<i>Dianthus busambræ</i> D2	<i>Dianthus caryophyllus</i> M77699	100	0
	<i>Dianthus deltoides</i> JN892692	100	0
	<i>Dianthus armeria</i> JN892039	100	0
	D1 (questo studio); D3 (q.studio);D4 (q.studio);D5 (q.studio);D6 (q.studio)	100	0
	<i>Dianthus seguieri</i> rbcl assente		
	<i>Dianthus furcatus</i> rbcl assente		
<i>Dianthus rupicola</i> subsp. <i>rupicola</i> D3	<i>Dianthus caryophyllus</i> M77699	100	0
	<i>Dianthus deltoides</i> JN892692	100	0
	<i>Dianthus armeria</i> JN892039	100	0
	D1 (questo studio); D2 (q.studio);D4 (q.studio);D5 (q.studio);D6 (q.studio)	100	0
	<i>Dianthus seguieri</i> rbcl assente		
	<i>Dianthus furcatus</i> rbcl assente		
<i>Dianthus rupicola</i> subsp. <i>Lopadusanum</i> D4	<i>Dianthus caryophyllus</i> M77699	100	0
	<i>Dianthus deltoides</i> JN892692	100	0
	<i>Dianthus armeria</i> JN892039	100	0
	D1 (questo studio); D2 (q.studio);D3 (q.studio);D5 (q.studio);D6 (q.studio)	100	0
	<i>Dianthus seguieri</i> rbcl assente		
	<i>Dianthus furcatus</i> rbcl assente		
<i>Dianthus siculus</i> D5	<i>Dianthus caryophyllus</i> M77699	100	0
	<i>Dianthus deltoides</i> JN892692	100	0
	<i>Dianthus armeria</i> JN892039	100	0
	D1 (questo studio); D2 (q.studio);D3 (q.studio);D4 (q.studio);D6 (q.studio)	100	0
	<i>Dianthus seguieri</i> rbcl assente		
	<i>Dianthus furcatus</i> rbcl assente		
<i>Dianthus rupicola</i> subsp. <i>hermaensis</i> D6	<i>Dianthus caryophyllus</i> M77699	100	0
	<i>Dianthus deltoides</i> JN892692	100	0
	<i>Dianthus armeria</i> JN892039	100	0
	D1 (questo studio); D2 (q.studio);D3 (q.studio);D4 (q.studio);D5 (q.studio)	100	0
	<i>Dianthus seguieri</i> rbcl assente		
	<i>Dianthus furcatus</i> rbcl assente		

Tab.15 Percentuale di identità genetica tra i taxa investigati e le sequenze referenti più affini per il marcatore *matK*

Famiglia	Taxa	matK		
		Referente GenBank ID	Identità %	Gap
ARECACEAE	<i>Chamaerops humilis</i> var. <i>humilis</i> (F.C4)	<i>Chamaerops humilis</i> var. <i>argentea</i> F.C5 (questo studio)	100	0
		<i>Rhapidothymus hystrix</i> KF746504	99	0
		<i>Guihaia argyrata</i> AM114569	99	0
		<i>Trachycarpus martianus</i> JX903673	99	0
		<i>Trachycarpus fortunei</i> AM114570	99	0
	<i>Chamaerops humilis</i> var. <i>argentea</i> (F.C5)	<i>Chamaerops humilis</i> var. <i>humilis</i> F.C4 (questo studio)	100	0
		<i>Rhapidothymus hystrix</i> KF746504	99	0
		<i>Guihaia argyrata</i> AM114569	99	0
		<i>Trachycarpus martianus</i> JX903673	99	0
		<i>Trachycarpus fortunei</i> AM114570	99	0
	<i>Brahea armata</i> (F.B4)	<i>Brahea aculeata</i> KF746502	100	0
		<i>Serenoa repens</i> AM114585	99	0
		<i>Washingtonia robusta</i> HQ720329	99	0
		<i>Brahea berlandieri</i> AM114580	99	0
		<i>Brahea dulcis</i> KF746480	99	0
		<i>Brahea calcarea</i> JX495676	99	0
	<i>Acoelorrhaphe wrightii</i> (F.A11)	<i>Serenoa repens</i> AM114585	99	3
		<i>Livistona</i> spp. (16 ID veddi App.2)	99	3
	<i>Arenga engleri</i> (F.A7)	<i>Arenga ryukyuensis</i> JF344968	99	0
		<i>Arenga caudata</i> JF344952	98	0
		<i>Arenga brevipes</i> JF344944	99	6
	<i>Caryota urens</i> (F.C20)	<i>Caryota monostachya</i> JF344990	99	0
	<i>Howea forsteriana</i> (F.H1)	<i>Howea belmoreana</i> GQ248136	100	0
	<i>Livistona chinensis</i> (F.L1)	<i>Livistona</i> spp. Non distinti 13 ID (vedi albero)	100	0
		<i>Serenoa repens</i> AM114585	100	0
		<i>Colpothrinax wrightii</i> HQ720256	100	0
		<i>Colpothrinax cookii</i> HQ720254	100	0
<i>Colpothrinax aphanopetala</i> HQ720253		100	0	
<i>Licuala tanycola</i> HQ720293		100	0	

	<i>Acoelorrhaphe wrightii</i> HQ720242	100	0
	<i>Livistona humilis</i> HQ720333	99	0
	<i>Washingtonia filifera</i> HQ720328	99	0
<i>Washingtonia robusta</i> (F.W1)	<i>Washingtonia filifera</i> HQ720328	100	0
	<i>Brahea aculeata</i> KF746502; <i>Brahea armata</i> HQ720246; <i>Brahea calcarea</i> JX495676; <i>Brahea berlandieri</i> AM114580; <i>Brahea dulcis</i> KF746480	99	0
	<i>Serenoa repens</i> AM114585	99	0
<i>Washingtonia filifera</i> (F.W2)	<i>Washingtonia robusta</i> HQ720329	100	0
	<i>Livistona humilis</i> HQ720333	99	0
	<i>Livistona chinensis</i> AM114574	99	0
	<i>Livistona benthamii</i> HQ720345	99	0
	<i>Livistona speciosa</i> matK assente		
<i>Trachycarpus fortunei</i> (F.T1)	<i>Trachycarpus martianus</i> JX903673	100	0
	<i>Trachycarpus takil</i> HQ720314	100	0
	<i>Rhapidophyllum hystrix</i> KF746504	99	0
	<i>Guihaia argyrata</i> AM114569	99	0
	<i>Chamaerops humilis</i> AM114568	99	0
<i>Phoenix roebelenii</i> (F.P7)	<i>Phoenix roebelenii</i> JX495744	100	0
	<i>Phoenix canariensis</i> JX495741	99	0
	<i>Phoenix reclinata</i> AM114565	99	0
	<i>Phoenix dactylifera</i> JX495742	99	0
<i>Phoenix canariensis</i> (F.P2)	<i>Phoenix reclinata</i> AM114565	100	0
	<i>Phoenix dactylifera</i> JX495742	100	0
	<i>Phoenix roebelenii</i> JX495744	99	0
<i>Phoenix dactylifera</i> (F.P8)	<i>Phoenix canariensis</i> JX495741	100	0
	<i>Phoenix reclinata</i> AM114565	100	0
	<i>Phoenix roebelenii</i> JX495744	99	0
<i>Phoenix reclinata</i> (F.P9)	<i>Phoenix canariensis</i> JX495741	100	0
	<i>Phoenix dactylifera</i> JX495742	100	0
	<i>Phoenix roebelenii</i> JX495744	99	0
<i>Sabal minor</i> (F.S14)	<i>Sabal palmetto</i> JX495754	99	0
	<i>Sabal bermudana</i> AM114553	99	0
	<i>Sabal etonia</i> GQ248196	99	0
	<i>Dictyocaryum lamarckianum</i> AM114616	99	0
	<i>Mauritia flexuosa</i> AM114545	98	0

	<i>Sabal palmetto</i> (F.S3)	<i>Sabal bermudana</i> AM114553	100	0
		<i>Sabal etonia</i> GQ248196	100	0
		<i>Sabal minor</i> AM114554	99	0
		<i>Mauritia flexuosa</i> AM114545	96	0
	<i>Bismarckia nobilis</i> (F.B3)	<i>Bismarckia nobilis</i> JX088664	100	0
		<i>Satranala decussilvae</i> AM114598	99	0
		<i>Hyphaene thebaica</i> AM114599	99	0
	<i>Butia capitata</i> (F.B2)	<i>Jubaea chilensis</i> EU004869	100	0
		<i>Syagrus romanzoffiana</i> GU135086	100	0
		<i>Cocos nucifera</i> JQ586724	100	0
		<i>Syagrus smithii</i> HQ265572	100	0
		<i>Butia yatay</i> AB088794	99	0
		mat assente <i>Butia eriospatha</i>		
	(Arecastrum)/ <i>Syagrus romanzoffiana</i> (F.A8)	<i>Syagrus smithii</i> HQ265572	100	0
		<i>Cocos nucifera</i> JQ586724	100	0
		<i>Butia capitata</i> JX903668	100	0
		<i>Jubaea chilensis</i> EU004869	100	0
		<i>Syagrus glaucescens</i> matK assente		
	<i>Trithrinax campestris</i> (F.T4)	<i>Trithrinax radiata</i> AM114561	99	0
		<i>Trithrinax acanthocoma</i> matK assente		
ASTERACEAE	<i>Anthemis cupaniana</i> (F.A2)	<i>Anthemis cotula</i> JN895749	99	0
		<i>Anthemis arvensis</i> HM850610	97	0
		matK assente <i>Anthemis maritima</i>		
		matK assente <i>Anthemis aetnensis</i>		
	<i>Centaurea sphaerocephala</i> F.C14	<i>Centaurea melitensis</i> HM850618	99	0
		<i>Centaurea nigra</i> JN895165	99	0
	<i>Helichrysum italicum</i> F.H2	<i>Helichrysum italicum</i> subsp. <i>italicum</i> KJ159106	100	0
		<i>Helichrysum microphyllum</i> HF585087	99	0
	<i>Helichrysum hyblaicum</i> H6	<i>Helichrysum scandens</i> H7 (questo studio)	100	0
		<i>Helichrysum microphyllum</i> HF585086	100	0
		<i>Helichrysum plicatum</i> HM445635	99	0
		<i>Helichrysum pallasii</i> HM445634	99	0
		<i>Helichrysum foetidum</i> HM445633	99	0
	<i>Helichrysum nebrodense</i> H5 (questo studio)	99	0	

		<i>Helichrysum italicum</i> subsp. <i>italicum</i> KJ159106	99	0
<i>Helichrysum nebrodense</i> H5		<i>Helichrysum hyblaicum</i> H6 (questo studio)	99	0
		<i>Helichrysum plicatum</i> HM445635	99	0
		<i>Helichrysum scandens</i> H7 (questo studio)	99	0
<i>Helichrysum scandens</i> H7		<i>Helichrysum hyblaicum</i> H6 (questo studio)	100	0
		<i>Helichrysum microphyllum</i> HF585086	100	0
		<i>Helichrysum nebrodense</i> H5 (questo studio)	99	0
<i>Ptilostemon greuteri</i> F.P5		<i>Atractylodes lancea</i> AB008761	97	0
		<i>Ptilostemon abylenis</i> KC590035	99	0
		<i>Ptilostemon hispanicus</i> KC590036	99	0
		<i>Ptilostemon gnaphaloides</i> KC969531	99	0
		<i>Ptilostemon afer</i> AY013550	99	0
		<i>Ptilostemon diacanthus</i> AY785112	99	0
<i>Jacobaea gibbosa</i> F.J1		<i>Artemisia argyi</i> HM989725	100	0
		<i>Artemisia vulgaris</i> JN894753	99	0
		<i>Artemisia afra</i> JQ412200	99	0
		<i>Artemisia sacrorum</i> JQ173390	98	9
		<i>Artemisia absinthium</i> JN894044	98	9
		<i>Artemisia gmelinii</i> GQ434109	99	9
		matK assente <i>Artemisia lavandulifolia</i>		
		FJ395360.1 <i>Jacobaea vulgaris</i>	99	0
		JN894078.1 <i>Jacobaea aquatica</i> ;	99	0
<i>Hieracium cophanense</i> H2		<i>Hieracium pignattianum</i> H9 (questo studio)	99	0
		<i>Hieracium pellucidum</i> AM889714	99	0
		<i>Hieracium umbellatum</i> KC474886	99	0
<i>Jacobaea ambigua</i> F.S4		<i>Jacobaea vulgaris</i> JN895164	100	0
		<i>Jacobaea aquatica</i> JN894078	100	0
		<i>Jacobaea aquatica</i> JN894726	99	0
<i>Senecio candidus</i> F.S5		<i>Jacobaea vulgaris</i> JN895164	99	0
		<i>Jacobaea aquatica</i> JN894078	99	0
		<i>Jacobaea ambigua</i> FS4 (questo studio)	99	0
		matk assente <i>Jacobaea maritima</i>		
FABACEAE	<i>Spartium junceum</i> F.S9	<i>Chamaecytisus prolifer</i> subsp. <i>angustifolius</i> KJ747665	99	0
	<i>Ceratonia siliqua</i> F.C1	<i>Umtiza listeriana</i> GU321973	98	0

	<i>Genista tyrrhena</i> G1	<i>Genista demarcoi</i> G4 (questo studio)	100	0
	<i>Genista madoniensis</i> G2	<i>Genista cupanii</i> G5 (questo studio)	100	0
	<i>Genista aristata</i> G3	<i>Genista tridentata</i> JQ858233	95	0
	<i>Genista demarcoi</i> G4	<i>Genista tyrrhena</i> G1 (questo studio)	100	0
	<i>Genista cupanii</i> G5	<i>Genista madoniensis</i> G2 (questo studio)	100	0
	<i>Genista aetnensis</i> G6	<i>Genista tyrrhena</i> G1 (questo studio)	99	0
EUPHORBIACEAE	<i>Euphorbia bivonae</i> Eu8	<i>Euphorbia serrulata</i> JN894826	99	0
	<i>Euphorbia ceratocarpa</i> Eu1	<i>Euphorbia hylonoma</i> EU659798	99	0
	<i>Euphorbia helioscopia</i> Eu7	<i>Euphorbia peplus</i> JN895712	100	0
	<i>Euphorbia amygdaloides</i> Eu11	<i>Euphorbia characias</i> KF408589	99	0
ROSACEAE	<i>Rosa sempervirens</i> R3	<i>Rosa laevigata</i> JN407194	99	0
		<i>Rosa micrantha</i> HE966983	100	0
		<i>Rosa multiflora</i> var. <i>cathayensis</i> GU363758	99	0
		<i>Rosa odorata</i> var. <i>gigantea</i> AB011994	99	0
		<i>Rosa chinensis</i> JN566105	99	0
		<i>Rosa agrestis</i> JN894864	99	0
	<i>Rosa canina</i> R1	<i>Rosa obtusifolia</i> JN894868	100	0
		<i>Rosa stylosa</i> JN894318	100	0
		<i>Rosa corymbifera</i> R10 (questo studio)	100	0
		<i>Rosa agrestis</i> JN894864	99	0
	<i>Rosa sicula</i> R6	<i>Rosa agrestis</i> JN894864	100	0
		<i>Rosa mollis</i> JN894862	100	0
		<i>Rosa gallica</i> AB011978	100	0
		<i>Rosa moschata</i> AB011990	99	0
	<i>Rosa corymbifera</i> R10	<i>Rosa canina</i> AB011980	100	0
		<i>Rosa obtusifolia</i> JN894868	100	0
<i>Rosa stylosa</i> JN894318		100	0	
<i>Rosa canina</i> R1 (questo studio)		100	0	
<i>Rosa sempervirens</i> AB048600		99	0	
BRASSICACEAE	<i>Brassica villosa</i> subsp. <i>tinei</i> B4	<i>Brassica napus</i> JF926672	99	0
		<i>Brassica oleracea</i> JF926675	98	0

	<i>Brassica rupestris</i> subsp. <i>hispida</i> B5	<i>Brassica napus</i> JF926672	99	0
		<i>Brassica oleracea</i> JF926675	99	0
ACANTHACEAE	<i>Acanthus mollis</i> F.A12	<i>Acanthus longifolius</i> AJ429326	99	0
XANTHORRHOEACEAE	<i>Aloe plicatilis</i> F.A5	<i>Aloe dichotoma</i> subsp. <i>pillansii</i> JQ024127	99	0
		<i>Aloe x spinosissima</i> matK assente		
	<i>Aloe vera</i> F.A9	<i>Aloe dyeri</i> JQ024129	99	0
		<i>Aloe comosa</i> JQ024124	99	0
<i>Aloe arborescens</i> F.A10	<i>Aloe x spinosissima</i> matK assente			
AMARANTHACEAE	<i>Diotis maritima</i> / <i>Achillea maritima</i> F.D1	<i>Achillea millefolium</i> JN894883	99	0
IRIDACEAE	<i>Iris pseudopumila</i> F.I1	<i>Iris subbiflora</i> FJ197301	100	0
	<i>Iris germanica</i> F.I5	<i>Iris pallida</i> var. <i>pallida</i> HM574689	100	0
LAMIACEAE	<i>Lavandula angustifolia</i> F.L4	<i>Lavandula stoechas</i> HQ902704	97	0
	<i>Lavandula stoechas</i> F.L5	<i>Lavandula angustifolia</i> HE967430	97	0
	<i>Sideritis italica</i> F.S6	<i>Erigeron acris</i> JN896122	99	0
	<i>Salvia leucantha</i> F.S12	<i>Salvia farinacea</i> JQ934057	99	0
	<i>Salvia officinalis</i> F.S11	<i>Salvia fruticosa</i> HQ902726; <i>Salvia sclarea</i> JQ934094; <i>Salvia verbenaca</i> HM850803	99	0
		matk assente: <i>Salvia x superba</i> , <i>Salvia staminea</i> , <i>Salvia texana</i> , <i>Salvia summa</i> , <i>Salvia roemeriana</i> , <i>Salvia pentstemonoides</i> , <i>Salvia palaestina</i> , <i>Salvia lyrata</i> , <i>Salvia henryi</i> , <i>S.indica</i> , <i>S.canariensis</i> , <i>S.bucharica</i> , <i>S.nemorosa</i> , <i>S.lavandulaefolia</i>		
<i>Rosmarinus officinalis</i> F.R4	<i>Salvia officinalis</i> FR719098	98	0	
MYRTACEAE	<i>Myrtus communis</i> F.M5	<i>Neomyrtus pedunculata</i> KM065097	98	0
		matk assente <i>Myrtus luma</i> F.M1 (questo studio)		
	<i>Metrosideros excelsa</i> F.M2	<i>Metrosideros robusta</i> KM065260	99	0
CYCADACEAE	<i>Cycas revoluta</i> F.C8	<i>Cycas taitungensis</i> AF410159	100	0
		<i>Cycas panzihuaensis</i> AF410158; <i>Cycas wadei</i> AF410163; <i>Cycas micholitzii</i> AB116591	99	0

	<i>Cycas circinalis</i> F.C19	<i>Cycas micronesica</i> GQ203833	100	0
ERICACEAE	<i>Erica sicula</i> E7	<i>Erica peduncularis</i> E9 (questo studio)	99	0
		<i>Erica multiflora</i> E10 (questo studio)	99	0
	<i>Erica peduncularis</i> E9	<i>Erica multiflora</i> E10 (questo studio)	100	0
		<i>Erica peduncularis</i> E9	100	0
	<i>Arbutus unedo</i> F.A1	<i>Arbutus canariensis</i> U61345	99	0
TAMARICACEAE	<i>Tamarix gallica</i> F.T5	<i>Tamarix pakistanica</i> KC840664	99	0
CARYOPHILLACEAE	<i>Dianthus rupicola</i> subsp. <i>aeolicus</i> D1	<i>Dianthus busambræ</i> D2 (questo studio)	100	0
		<i>Dianthus caryophyllus</i> JQ412232; <i>Dianthus sylvestris</i> subsp. <i>Tergestinus</i> HE967399; <i>Dianthus seguieri</i> AY936321; <i>Dianthus furcatus</i> AY936320	100	0
		<i>Dianthus deltoides</i> JN895428; <i>Dianthus armeria</i> FJ404832	99	0
	<i>Dianthus busambræ</i> D2	<i>Dianthus rupicola</i> subsp. <i>aeolicus</i> D1 (questo studio)	100	0
		<i>Dianthus caryophyllus</i> JQ412232; <i>Dianthus sylvestris</i> subsp. <i>Tergestinus</i> HE967399; <i>Dianthus seguieri</i> AY936321; <i>Dianthus furcatus</i> AY936320	100	0
		<i>Dianthus deltoides</i> JN895428; <i>Dianthus armeria</i> FJ404832	99	0
	<i>Dianthus rupicola</i> subsp. <i>hermaensis</i> D6	<i>Dianthus rupicola</i> subsp. <i>hermaensis</i> D6 (questo studio)	100	0
		<i>Dianthus deltoides</i> JN895428; <i>Dianthus armeria</i> FJ404832; <i>Dianthus seguieri</i> AY936321; <i>Dianthus furcatus</i> AY936320; <i>Dianthus caryophyllus</i> JQ412232; <i>Dianthus sylvestris</i> subsp. <i>tergestinus</i> HE967399	99	0
	<i>Dianthus rupicola</i> subsp. <i>Lopadusanum</i> D4	D1 (questo studio); D2 (q. studio); D3 (q.studio); D5 (q. studio); D6 (q. studio)	99	0
	<i>Dianthus siculus</i> D5	D1 (questo studio); D2 (q. studio); D3 (q.studio); D4 (q. studio); D6 (q. studio)	99	0
<i>Dianthus rupicola</i> subsp. <i>hermaensis</i> D6	<i>Dianthus rupicola</i> subsp. <i>rupicola</i> D3 (questo studio)	100	0	
	<i>Dianthus deltoides</i> JN895428; <i>Dianthus armeria</i> FJ404832; <i>Dianthus seguieri</i> AY936321; <i>Dianthus furcatus</i> AY936320; <i>Dianthus caryophyllus</i> JQ412232; <i>Dianthus sylvestris</i> subsp. <i>tergestinus</i> HE967399	99	0	

Tab.16 Percentuale di identità genetica tra i taxa investigati e le sequenze referenti più affini per il marcatore *trnH-psbA*

trnH-psbA				
Famiglia	Taxa	Referente GenBank ID	Identità %	Gap
ARECACEAE	<i>Chamaerops humilis</i> var. <i>humilis</i> (F.C4)	<i>Phoenix dactylifera</i> JN854234	99	0
		<i>Phoenix reclinata</i> GU135452	99	0
		<i>Chamaerops humilis</i> var. <i>argentea</i> F.C5 (questo studio)	100	0
		<i>Chamaerops humilis</i> "Vulcano" F.C6 (questo studio)	100	0
		<i>Chamaerops humilis</i> "Etna star" F.C7 (questo studio)	100	0
	<i>Chamaerops humilis</i> var. <i>argentea</i> (F.C5)	<i>Phoenix dactylifera</i> JN854234	99	0
		<i>Phoenix reclinata</i> GU135452	99	0
		<i>Chamaerops humilis</i> var. <i>humilis</i> F.C4 (questo studio)	100	0
		<i>Chamaerops humilis</i> "Vulcano" F.C6 (questo studio)	100	0
		<i>Chamaerops humilis</i> "Etna star" F.C7 (questo studio)	100	0
	<i>Chamaerops humilis</i> "Vulcano" (F.C6)	<i>Phoenix dactylifera</i> JN854234	99	0
		<i>Phoenix reclinata</i> GU135452	99	0
		<i>Chamaerops humilis</i> var. <i>argentea</i> F.C5 (questo studio)	100	0
		<i>Chamaerops humilis</i> "Vulcano" F.C6 (questo studio)	100	0
		<i>Chamaerops humilis</i> "Etna star" F.C7 (questo studio)	100	0
	<i>Chamaerops humilis</i> "Etna star" (F.C7)	JN854234.1 <i>Phoenix dactylifera</i>	99	0
		GU135452.2 <i>Phoenix reclinata</i>	99	0
		<i>Chamaerops humilis</i> L. var. <i>argentea</i> F.C5 (questo studio)	100	0
		<i>Chamaerops humilis</i> L. "Vulcano" F.C6 (questo studio)	100	0
		<i>Chamaerops humilis</i> L. "Etna star" F.C7 (questo studio)	100	0
	<i>Howea forsteriana</i> (F.H1)	<i>Howea belmoreana</i> GQ248317	100	0
		<i>Ptychosperma elegans</i> GU135295	99	0
		6 ID rimangono non distinti trnH assente		
	<i>Livistona chinensis</i> (F.L1)	<i>Sabal pumos</i> EF688521	99	
trnH assente <i>Livistona</i> spp. (19 ID)				
trnH assente <i>Serenoa repens</i>				

		trnH assente <i>Colpothrinax</i>		
		trnH assente <i>Licuala tanycola</i>		
	<i>Trachycarpus fortunei</i> (F.T1)	trnH assente <i>Trachycarpus martianus</i>		
	<i>Phoenix dactylifera</i> (F.P8)	<i>Phoenix reclinata</i> GU135452	89	0
		trnH assente <i>Phoenix canariensis</i>		
	<i>Phoenix reclinata</i> (F.P9)	<i>Phoenix dactylifera</i> JN854234	98	0
		trnH assente <i>Phoenix canariensis</i>		
	<i>Sabal minor</i> (F.S14)	<i>Sabal etonia</i> GQ248387; <i>Sabal bermudana</i> EF688513; <i>Sabal palmetto</i> EF688520; <i>Sabal causiarum</i> EF688510; <i>Sabal domingensis</i> EF688514; <i>Sabal guatemalensis</i> EF688516; <i>Sabal maritima</i> EF688517; <i>Sabal rosei</i> EF688523; <i>Sabal miamiensis</i> EF688525	100	0
	<i>Sabal palmetto</i> (F.S3)	<i>Sabal etonia</i> GQ248387; <i>Sabal bermudana</i> EF688513; <i>Sabal minor</i> GQ248388; <i>Sabal causiarum</i> EF688510; <i>Sabal domingensis</i> EF688514; <i>Sabal guatemalensis</i> EF688516; <i>Sabal maritima</i> EF688517; <i>Sabal rosei</i> EF688523; <i>Sabal miamiensis</i> EF688525	100	0
ASTERACEAE	<i>Anthemis cupaniana</i> F.A2	<i>Anthemis ismelia</i> FR689904	100	0
		<i>Anthemis aetnensis</i> FR689793	100	0
		<i>Anthemis kotschyana</i> FR689808	90	28
		<i>Anthemis maritima</i> FR689772	90	28
	<i>Anthemis maritima</i> F.A4	<i>Anthemis maritima</i> FR689772	100	0
		<i>Anthemis pedunculata</i> FR689774	100	0
		<i>Anthemis microsperma</i> FR689831	100	0
		<i>Anthemis mauritiana</i> FR689811	100	0
		<i>Anthemis boveana</i> FR689799	100	0
		<i>Anthemis glareosa</i> FR689826	99	1
		<i>Anthemis maroccana</i> FR689810	99	0
		<i>Anthemis cupaniana</i> FR689901	90	28
		<i>Helichrysum italicum</i> F.H2	<i>Helichrysum italicum</i> subsp. <i>siculum</i> FR667160	99
<i>Helichrysum nebrodense</i> H5	<i>Helichrysum scandens</i> H7 (questo studio)	99	5	

		<i>Helichrysum italicum</i> subsp. <i>microphyllum</i> FR667156	99	5
		<i>Helichrysum stoechas</i> GU818387	99	5
	<i>Helichrysum scandens</i> H7	<i>Helichrysum stoechas</i> GU818387	100	0
		<i>Helichrysum italicum</i> subsp. <i>microphyllum</i> FR667156	100	0
		<i>Helichrysum felinum</i> FM173164	96	0
		trnH assente <i>Helichrysum hyblaenum</i>		
	Jacobaea gibbosa F.J1	<i>Artemisia argyi</i> GQ435067	99	0
		trnH assente <i>Artemisia lavandulifolia</i>		
		<i>Jacobaea</i> spp. (vedi albero)	99	0
	<i>Jacobaea ambigua</i> F.S4	<i>Jacobaea vulgaris</i> HE966807	99	1
		<i>Jacobaea gigantea</i> AY155653	99	2
		<i>Jacobaea aquatica</i> AY155646	98	8
	<i>Senecio candidus</i> F.S5	<i>Jacobaea maritima</i> AY155647	96	11
		<i>Jacobaea vulgaris</i> AY155657	95	12
		<i>Jacobaea ambigua</i> F.S4 (questo studio)	95	13
		<i>Jacobaea aquatica</i> AY155646	99	0
FABACEAE	<i>Genista tyrrhena</i> G1	<i>Genista demarcoi</i> G4 (questo studio)	98	0
	<i>Genista madoniensis</i> G2	<i>Genista cupanii</i> G5 (questo studio)	99	0
	<i>Genista aristata</i> G3	<i>Genista cupanii</i> G5 (questo studio)	93	0
	<i>Genista cupanii</i> G5	<i>Genista madoniensis</i> G2 (questo studio)	99	0
	<i>Genista demarcoi</i> G4	<i>Genista tyrrhena</i> G1 (questo studio)	98	0
EUPHORBIACEAE	<i>Euphorbia bivonae</i> Eu8	<i>Euphorbia hylonoma</i> KC758661	95	16
	<i>Euphorbia ceratocarpa</i> Eu1	<i>Euphorbia pekinensis</i> GQ434949	96	16
	<i>Euphorbia characias</i> E2	<i>Euphorbia characias</i> subsp. <i>wulfenii</i> HE966622	100	0
		<i>Euphorbia amygdaloides</i> Eu11 (questo studio)	98	12
	<i>Euphorbia dendroides</i> E4	<i>Euphorbia dendroides</i> HQ900428	100	0
		<i>Euphorbia characias</i> E2 (questo studio)	98	12
	<i>Euphorbia helioscopia</i> Eu7	<i>Euphorbia peplus</i> HE966630	100	0
	<i>Euphorbia myrsinites</i> E6	<i>Euphorbia nicaeensis</i> subsp. <i>niccaensis</i> HE966629	93	0
<i>Euphorbia meuselii</i> E5	<i>Euphorbia characias</i> subsp. <i>wulfenii</i> HE966622	99	2	

	<i>Euphorbia pithyusa</i> subsp. <i>cupanii</i> Eu9	<i>Euphorbia nicaeensis</i> subsp. <i>Nicaeensis</i> HE966629	94	16
	<i>Euphorbia amygdaloides</i> Eu11	<i>Euphorbia characias</i> subsp. <i>wulfenii</i> HE966622	98	11
ROSACEAE	<i>Rosa sempervirens</i> R3	<i>Rosa corymbifera</i> R10 (questo studio)	99	0
		<i>Rosa canina</i> R1	99	0
		<i>Rosa micrantha</i> HE966778	91	0
		<i>Rosa Canina</i> HE966777	99	0
		<i>Rosa phoenicia</i> AB043944	99	0
	<i>Rosa canina</i> R1	<i>Rosa corymbifera</i> R10 (questi studio)	100	0
		<i>Rosa sempervirens</i> R3 (questo studio)	100	0
	<i>Rosa sicula</i> R6	<i>Rosa agrestis</i> DQ778744	99	0
		<i>Rosa rubiginosa</i> DQ778769	99	0
	<i>Rosa corymbifera</i> R10	<i>Rosa sempervirens</i> R3	100	0
<i>Rosa canina</i> R1		100	0	
BRASSICACEAE	<i>Brassica insularis</i> B2	<i>Brassica oleracea</i> var. <i>capitata</i> GQ184356	99	0
XANTHORRHOEACEAE	<i>Aloe plicatilis</i> F.A5	<i>Aloe haemanthifolia</i> KC960554	99	0
		<i>Aloe x spinosissima</i> trnH assente		
	<i>Aloe vera</i> F.A9	<i>Aloe spicata</i> JQ039282	98	0
	<i>Aloe arborescens</i> F.A10	<i>Aloe haworthioides</i> JQ039265	99	0
<i>Aloe x spinosissima</i> trnH assente				
LILIACEAE	<i>Tulipa radii</i> T4	<i>Tulipa cretica</i> AM049257	96	0
IRIDACEAE	<i>Iris pseudopumila</i> F.11	<i>Iris subbiflora</i> trnH Assente		
	<i>Iris germanica</i> F.15	<i>Iris pallida</i> trnH assente		
CISTACEAE	<i>Cistus salviifolius</i> F.C2	<i>Cistus x pulverulentus</i> F.C3 (questo studio); <i>Cistus albidus</i> F.C11 (q. studio); <i>Cistus skanbergii</i> F.C12 (q. studio)	99	0
		mancano ref. Altri cistus in banca dati		
	<i>Cistus x pulverulentus</i> F.C3	<i>Cistus skanbergii</i> F.C12 (q. studio); <i>Cistus albidus</i> F.C11 (q. studio); <i>Cistus salviifolius</i> F.C2 (q.studio)	99	0
		mancano ref. Altri cistus in banca dati		
	<i>Cistus albidus</i> F.C11	<i>Cistus salviifolius</i> F.C2 (q.studio); <i>Cistus x pulverulentus</i> F.C3 (q. studio); <i>Cistus skanbergii</i> F.C12 (q. studio)	99	0
		mancano ref. Altri cistus in banca dati		

	<i>Cistus skanbergii</i> F.C12	<i>Cistus salviifolius</i> F.C2 (q.studio); <i>Cistus x pulverulentus</i> F.C3 (q.studio); <i>Cistus albidus</i> F.C11 (q.studio)	99	0
		mancano ref. Altri cistus in banca dati		
LAMIACEAE	<i>Sideritis italica</i> F.S6	<i>Sideritis trojana</i> GU230757	95	0
	<i>Salvia officinalis</i> F.S11	<i>Perovskia atriplicifolia</i> DQ667341	99	0
		trnH assente: <i>Salvia staminea</i> , <i>Salvia texana</i> , <i>Salvia summa</i> , <i>Salvia roemeriana</i> , <i>Salvia pentstemonoides</i> , <i>Salvia palaestina</i> , <i>Salvia lyrata</i> , <i>Salvia henryi</i> , <i>S.indica</i> , <i>S.bucharica</i> , <i>S.nemorosa</i> , <i>S.lavandulaefolia</i>		
MYRTACEAE	<i>Myrtus luma</i> FM1	<i>Myrtus communis</i> subsp. <i>communis</i> HE966716	99	0
		<i>Myrtus communis</i> AM489872	100	0
		<i>Myrtus communis</i> HF585096	100	0
	<i>Myrtus communis</i> F.M5	<i>Myrtus luma</i> F.M1 (questo studio)	100	0
CYCADACEAE	<i>Cycas revoluta</i> F.C8	<i>Cycas taitungensis</i> AP009339	99	0
ERICACEAE	<i>Erica peduncularis</i> E9	<i>Erica multiflora</i> E10 (questo studio)	99	0
	<i>Erica multiflora</i> E10	<i>Erica peduncularis</i> E9	99	0
CONVOLVULACEAE	<i>Calystegia soldanella</i> F.C13	<i>Calystegia sepium</i> FJ395547	96	0
CARYOPHILLACEAE	<i>Dianthus rupicola</i> subsp. <i>aeolicus</i> D1	<i>Dianthus busambrae</i> D2 (questo studio)	98	8
		<i>Dianthus furcatus</i> trnH assente		
		<i>Dianthus caryophyllus</i> trnH assente		
	<i>Dianthus busambrae</i> D2	<i>Dianthus rupicola</i> subsp. <i>aeolicus</i> D1 (questo studio)	98	8
		<i>Dianthus furcatus</i> trnH assente		
		<i>Dianthus caryophyllus</i> trnH assente		
	<i>Dianthus rupicola</i> subsp. <i>rupicola</i> D3	<i>Dianthus rupicola</i> subsp. <i>hermaensis</i> D6 (questo studio)	99	2
	<i>Dianthus rupicola</i> subsp. <i>Lopadusanum</i> D4	<i>Dianthus siculus</i> D5 (questo studio)	98	17
<i>Dianthus siculus</i> D5	<i>Dianthus rupicola</i> subsp. <i>Lopadusanum</i> D4 (questo studio)	98	17	
<i>Dianthus rupicola</i> subsp. <i>hermaensis</i> D6	<i>Dianthus rupicola</i> subsp. <i>rupicola</i> D3 (questo studio)	99	2	