



UNIVERSITÀ DEGLI STUDI DELLA TUSCIA DI VITERBO

DIPARTIMENTO DI TECNOLOGIE, INGEGNERIA E SCIENZE DELL'AMBIENTE E
DELLE FORESTE

CORSO DI DOTTORATO DI RICERCA IN
SCIENZE E TECNOLOGIE PER LA GESTIONE FORESTALE E AMBIENTALE
XXIII Ciclo

**STIMA DELLA BIODIVERSITÀ DELLA DENDROFLORA ITALIANA
ATTRAVERSO IL METODO DEL DNA BARCODING: IL GENERE
QUERCUS**

Settore scientifico-disciplinare: AGR/05

Coordinatore: Prof. Gianluca Piovesan

Tutor: Prof. Bartolomeo Schirone

Co-Tutor: Prof. Marcella Attimonelli

Dottoranda: Roberta Piredda

a mia madre Andreina e mio padre Gesuino

Indice

1 Introduzione

◆ *Il DNA Barcoding*

Le origini

Il DNA Barcoding negli animali

◆ *Il DNA barcoding nelle piante*

Premessa

Il genoma cloroplastico

Il candidato barcode

◆ *Concetto di specie e DNA barcoding*

La classificazione

La specie

L'analisi dei dati

2 Il genere Quercus

Introduzione

Le problematiche del genere Quercus

Le classificazioni subgeneriche

La situazione tassonomica delle querce italiane

Il subgenere Quercus

Il subgenere Cerris

Il subgenere Sclerophyllodris

Querce di dubbia collocazione

3 Il disegno sperimentale della tesi

Strategia di campionamento

La scelta dei marcatori

4 Materiali e Metodi

- ◆ *Il materiale vegetale*
- ◆ *I marcatori molecolari*
 - matK*
 - rbcL*
 - rpoC1*
 - trnH-psbA*
- ◆ *Il protocollo molecolare*
 - L'estrazione del DNA*
 - L'amplificazione*
 - La purificazione*
 - Il sequenziamento*
- ◆ *L'analisi dei dati*

5 Risultati

- Campionamento floristico*
- Campionamento taxon-based*

6 Discussione

1 IL DNA BARCODING

Le origini

Il termine “DNA barcode” è stato introdotto per la prima volta dal genetista di popolazione P. Hebert per indicare una ben definita regione di DNA capace di discriminare tutte le specie.

In effetti già in precedenza sequenze nucleotidiche erano state utilizzate per classificare ed esplorare le specie, tra questi ricordiamo il primo storico lavoro di Carl Woese che attraverso l’uso del gene della subunità ribosomale 16S-rDNA evidenziò l’esistenza del dominio degli Archea (Woese and Fox, 1977). Tuttavia, quello che rese questa proposta innovativa fu il concetto di “standardizzazione” ossia utilizzare lo stesso gene (o parte di esso) per tutte le specie, possibilmente ottenendo tale frammento attraverso una coppia di primers universali. Quindi, supponendo di attribuire quattro colori diversi alle basi nucleotidiche *ATCG*, le differenze lungo la sequenza porterebbero alla generazione di differenti “codici a barre” specifici per ogni specie.

Il progetto Barcode riesce a decollare grazie al sostegno economico della Alfred P. Sloan Foundation di New York che consente la nascita, nel duemilaquattro, del Consortium Barcode of Life (CBOL) con sede ufficiale presso il Museo Nazionale di Storia Naturale “Smithsonian Institution” a Washington. Numerose istituzioni sia pubbliche che private hanno aderito al CBOL che resta attualmente la struttura di riferimento a livello mondiale, con la funzione di promuovere il progetto Barcode attraverso varie iniziative quali conferenze, workshops, networks, gruppi di lavoro.

Proprio nell’anno della fondazione il CBOL elabora un primo documento illustrato di quattro pagine “Ten Reason for Barcoding Life” con il quale brevemente il progetto viene presentato alla comunità scientifica e non solo. Il documento definisce il concetto di Barcode e riassume in dieci punti i vantaggi e gli aspetti innovativi di tale approccio. Nelle pagine seguenti si riporta il documento e qui di seguito la traduzione dei concetti più significativi.

1. Lavorare con frammenti. Il Barcode sarà in grado di identificare una specie anche a partire da pezzi e frammenti. Questo consentirà di identificare animali indesiderati o materiale vegetale in prodotti alimentari trasformati e di individuare i prodotti commerciali derivati da specie regolamentate. Barcode aiuterà anche a ricostruire i cicli alimentari individuando frammenti di cibo dagli stomaci.

2. Lavorare con diversi stadi di sviluppo. Il barcode sarà in grado di identificare una specie nelle sue molteplici forme, dalle uova, seme, attraverso le larve e piante, sia in forme adulte che giovanili.

3. Smaschera i “simili”. Il barcode sarà in grado di distinguere tra le specie che si assomigliano, scoprendo organismi pericolosi camuffati da innocui e consentendo una visione più precisa della biodiversità.

4. Riduce le ambiguità. Scritto come una sequenza di quattro nucleotidi discreti - CATG – e proveniente da una regione uniforme sui genomi, il barcode offrirà una identificazione digitale della vita. Una libreria di codici a barre digitali fornirà un riferimento inequivocabile, che faciliterà l'identificazione delle specie.

5. Potenziare le competenze umane. Il cervello umano possiede dei limiti, ed è stato appurato che un esperto tassonomista non sia in grado di classificare oltre un certo numero di specie. Prevedendo una numero di specie totali di circa cento milioni, gli scienziati potrebbero utilizzare il barcode per velocizzare l'identificazione degli organismi conosciuti e dedicare più tempo per l'identificazione di nuove specie.

6. Democratizzazione della conoscenza. Una libreria di barcode consentirà a molte più persone non esperte di chiamare per nome le specie che ci circondano. Renderà possibile identificazione di specie abbondante o rare, native o invasive, valorizzando la biodiversità a livello locale e globale.

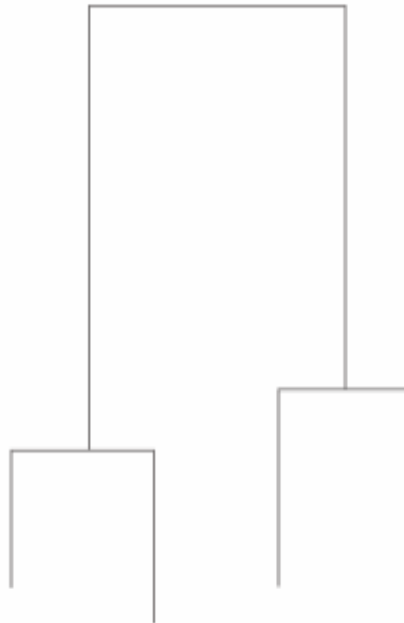
7. Progettazione di una guida elettronica palmare da campo “Barcode Life”. E’ in fase di progettazione la realizzazione di un dispositivo da campo che consentirà in tempo reale di ottenere la sequenza del nostro campione ignoto e collegandosi ad un database centrale fornire il nome della specie in esame. Immaginate le potenzialità di uno scolaro con un Barcode in mano che potrebbe leggere la biodiversità naturale, o la praticità che avrebbe per un ecologo, o la velocità delle operazioni di sicurezza svolte da un ispettore in un porto con un Barcode collegato a un computer centrale!

8. Germogli nuovi sull'albero della vita. L'albero della vita di Darwin, si arricchirà di nuove specie, e la ricchezza delle informazioni genetiche acquisite attraverso il barcode forse consentirà una classificazione più naturale e completa di quella basata su somiglianze e differenze.

9. Potenziamento del valore delle collezioni. La costruzione del database di codici a barre inizia con i milioni di esemplari collezionati in musei, erbari, zoo e giardini dei quali sarà fatto il barcode. Questa azione senza dubbio contribuirà a preservare la biodiversità della Terra.

10. Veloce compilazione dell'enciclopedia della vita. La creazione di una libreria di codici a barre legati al voucher di campioni correttamente classificati faciliterà l'accesso pubblico alla conoscenza biologica, contribuendo a creare una enciclopedia on-line della vita sulla Terra, con una pagina web per ogni specie di piante e animali.

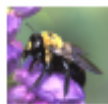
Differences in DNA barcodes help create trees showing genetic distances between species.



Hermit Thrush



American Robin



Bumblebee



Honey Bee

More information is available at:

<http://barcoding.si.edu/>

and

<http://phe.rockefeller.edu/BarcodeConference/index.html>

BARCODING LIFE: TEN REASONS

IDENTIFYING SPECIES BY DNA



CONSORTIUM FOR THE BARCODE OF LIFE



CONSORTIUM FOR THE BARCODE OF LIFE (CBOL) is an international collaboration of natural history museums, herbaria, biological repositories, and biodiversity inventory sites, together with academic and commercial experts in genomics, taxonomy, electronics, and computer science. The initial organizational support for CBOL is provided by a 2.5 year grant from the Sloan Foundation.

The mission of **CBOL** is to rapidly accelerate compiling of DNA barcodes of known and newly discovered plant and animal species, establish a public library of sequences linked to named specimens, and promote development of portable devices for DNA barcoding.

Mark Stoeckle The Rockefeller University

Paul Waggoner Connecticut Agricultural Experiment Station

Jesse Ausubel Alfred P. Sloan Foundation

v3.0 September 18, 2004

BARCODE OF LIFE: A short DNA sequence, from a uniform locality on the genome, used for identifying species.

DNA sequences from a uniform locality on genomes can be a *barcode of life* for identifying species, always the front line in discovery, monitoring and research. Since Linnaeus, biologists have used distinguishing features in taxonomic keys to apply binomial species names, such as *Homo sapiens*. Then, as a master key opens all the rooms in a building, the binomial species name accesses all knowledge about a species.

From insects to birds, evidence now shows that short DNA sequences from a uniform locality on genomes can also be a distinguishing feature. As a Linnaean binomial is an abbreviated label for the morphology of a species, the short sequence is an abbreviated label for the genome of the species.

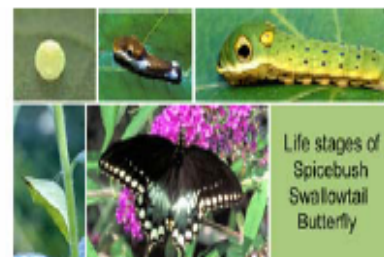
The *barcode of life* thus provides an additional master key to knowledge about a species. Compiling a public library of sequences linked to named specimens, plus faster and cheaper sequencing, will make this new barcode key increasingly practical and useful.

What additional power does barcoding offer?

1. Works with fragments. Barcoding can identify a species from bits and pieces. When established, barcoding will quickly identify undesirable animal or plant material in processed foodstuffs and detect commercial products derived from regulated species. Barcoding will help reconstruct food cycles by identifying fragments in stomachs and assist plant science by identifying roots sampled from soil layers.



2. Works for all stages of life. Barcoding can identify a species in its many forms, from eggs and seed, through larvae and seedlings, to adults and flowers.

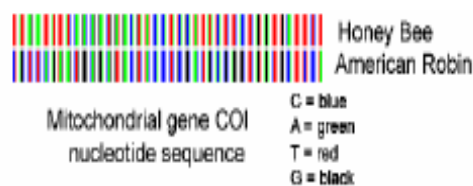


3. Unmasks look-alikes. Barcoding can distinguish among species that look alike, uncovering dangerous organisms masquerading as harmless ones and enabling a more accurate view of biodiversity.



Anopheles mosquitoes are vectors for human malaria. Only a few of the 430 known species transmit infection.

4. Reduces ambiguity. Written as a sequence of four discrete nucleotides - CATG - along a uniform locality on genomes, a barcode of life provides a digital identifying feature, supplementing the more analog gradations of words, shapes and colors. A library of digital barcodes will provide an unambiguous reference that will facilitate identifying species invading and retreating across the globe and through centuries.



5. Makes expertise go further. The bewildering diversity of about 2 million species already known confines even an expert to morphological identification of only a small part of the plant and animal kingdoms. Foreseeing millions more species to go, scientists can equip themselves with barcoding to speed identification of known organisms and facilitate rapid recognition of new species.



Known species:

>150,000	>250,000	>300,000	>30,000
flies, mosquitoes	flowering plants	beetles	crabs, lobsters

6. Democratizes access. A standardized library of barcodes will empower many more people to call by name the species around them. It will make possible identification of species whether abundant or rare, native or invasive, engendering appreciation of biodiversity locally and globally.

7. Opens the way for an electronic handheld field guide, the Life Barcode.

Barcoding links biological identification to advancing frontiers in DNA sequencing, miniaturization in electronics, and computerized information storage. Integrating those links will lead to portable desktop devices and ultimately to hand-held barcoders. Imagine the promise of a schoolchild with a barcoder in hand learning to read wild biodiversity, the power granted to a field ecologist surveying with a barcoder and global positioning system, or the security imparted by a port inspector with a barcoder linked to a central computer!



A handheld barcoder, such as the one envisioned here, would have many uses. Promoting technology development of portable devices for field use is a major goal of this initiative.

8. Sprouts new leaves on the tree of life.

Since Darwin, biologists seeking a natural system of classification have drawn genealogical trees to represent evolutionary history. Barcoding the similarities and differences among the nearly 2 million species already named will provide a wealth of genetic detail, helping to draw the tree of life on Earth. Barcoding newly discovered species will help show where they belong among known species, sprouting new leaves on the tree of life.

9. Demonstrates value of collections.

Compiling the library of barcodes begins with the multimillions of specimens in museums, herbaria, zoos and gardens, and other biological repositories. The spotlight that barcoding shines on these institutions and their collections will strengthen their ongoing efforts to preserve Earth's biodiversity.

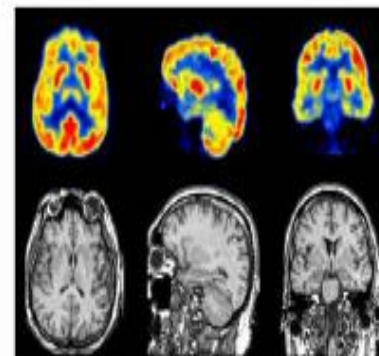


The National Museum of Natural History has over 600,000 bird specimens representing more than 80% of the world's known species.

10. Speeds writing the encyclopedia of life. Compiling a library of barcodes linked to vouchered specimens and their binomial names will enhance public access to biological knowledge, helping to create an on-line encyclopedia of life on Earth, with a web page for every species of plant and animal.



Earth is home to an estimated 10 million species of plants and animals.



The human brain can learn to identify a few thousand species, a small fraction of life's diversity.

Il DNA barcoding negli animali

Descritte le ragioni, i motivi e i vantaggi del progetto Barcode, tuttavia restano da definire alcuni aspetti di fondamentale importanza.

Esiste un frammento genomico con caratteristiche tali da funzionare bene in tutte le specie?

Quali sono queste caratteristiche?

Senza dubbio il genoma mitocondriale possiede alcune peculiarità che lo rendono un candidato migliore rispetto al genoma nucleare:

- elevato numero di copie per cellula (da 100 a 10000) di mtDNA rispetto alle 2 copie di quello nucleare
- modalità di eredità citoplasmatica
- assenza di ricombinazione
- tasso evolutivo maggiore (circa 10 volte) rispetto al genoma nucleare, a causa di un sistema di riparo assente o con ridotta efficienza, quindi maggior potenziale informativo a parità di lunghezza di sequenza

Proprio grazie a queste proprietà infatti, diversi frammenti mitocondriali sono stati da sempre utilizzati in studi di vario tipo (filogenetico, popolazionale, etc.), ma quale poteva essere il più adatto per svolgere al meglio il ruolo di Barcode?

A parte poche eccezioni il DNA mitocondriale negli animali contiene lo stesso set di geni: 13 proteine codificanti, 2 RNA ribosomali, 22 transfer RNA, tuttavia l'ordine e la polarità dei geni è abbastanza variabile. Per questo motivo la scelta poteva cadere solo all'interno di un gene.

Come riportato in una tabella disponibile nel Hebert Lab (www.dnabarcoding.ca/primer/Prior_mtDNA.html) e mostrata in figura1 , il database NCBI conteneva nell'agosto 2002 circa 90000 entries di geni mitocondriali animali, per la maggior parte appartenenti al citocromo b (cytb) e alla subunità 1 il citocromo C ossidasi (COI) , ma mentre nel caso del cytb le sequenze appartengono principalmente ai Cordati, i dati relativi alla COI sono distribuiti nei vari gruppi animali.

Mitochondrial Gene Sequences in NCBI (July 31st, 2002).				
PROTEIN	TOTAL	CHORDATA	ARTHROPODA	OTHERS
Cyt <i>B</i>	26193	22610	1604	1979
COI	12174	1598	7021	3555
ATP6	4601	3881	70	644
COII	4060	648	3177	235
ND2	2776	2452	224	100
ND5	2370	565	1725	80
ND1	2046	1137	638	271
ND4	1957	1426	162	369
ATP8	997	430	54	513
ND3	987	820	69	98
ND6	906	767	60	79
ND4L	852	677	75	100
COIII	795	534	134	127
RIBOSOMAL	TOTAL	CHORDATA	ARTHROPODA	OTHERS
16S	15453	6658	5668	3127
12S	10717	7850	2242	625

Figura 1: Entries di geni mitocondriali animali

Sulla base dei dati allora presenti tutti i geni vennero valutati prendendo in considerazione vari aspetti, quali lunghezza del gene, tasso evolutivo, presenza di inserzioni e delezioni (indels) e la possibilità di costruire robusti primers universali. Un frammento di 648 nucleotidi in prossimità dell'estremità 5' del gene COI venne considerato il miglior candidato, dando il via alla prima fase del progetto Barcode rivolta a dimostrare la flessibilità del frammento COI nei vari gruppi animali. Lo stesso P. Hebert (Hebert et al. 2003a, 2003b) esplorò le potenzialità di questa regione a partire da dati disponibili su NCBI dimostrando un'alta variabilità interspecie, che nel 98% dei casi sarebbe maggiore del 2% , e una bassa variabilità intraspecie in media meno del 0,3% (generata secondo l'autore dalla presenza di "selective sweep"). Vengono inoltre prodotte circa 200 nuove sequenze COI di Lepidotteri (correttamente assegnate a livello di specie), e utilizzato il set di primers universali (Folmer et al. 1994) LCO1490: 5'-GGTCAACAAATCATAAAGATATTGG-3' e HCO2198: 5'-TAAACTTCAGGGTGACCAAAAAATCA-3' , ancora oggi utilizzati.

Gettate le basi teorico-pratiche del metodo, saranno ora i vari gruppi di ricerca che, attraverso la raccolta e l'analisi di dati sperimentali relativi ai diversi gruppi tassonomici, metteranno in luce le potenzialità e i limiti della tecnica barcode.

Tuttavia all'interno della comunità scientifica non tutti furono entusiasti di questo nuovo approccio per lo studio della biodiversità (Wheeler, 2004; Will and Rubinoff, 2004; Ebach and Holdredge, 2005; Will et al., 2005) sono un piccolo esempio. Una grossa fetta di tassonomisti infatti, vide il DNA barcode come un nemico che in qualche modo sminuiva le loro competenze: come poteva un solo gene essere il principale mezzo per identificare una specie

già nota, per non parlare poi di quelle nuove? Come poteva rappresentare in modo corretto le relazioni tra le specie?

Queste ed altre critiche vennero fatte da più parti, tanto che P. Hebert in un articolo, scritto quasi sotto forma di intervista (Hebert and Gregory 2005), cercò di placare il dibattito specificando tra le altre cose, un concetto chiave:

lo scopo del barcode non è stabilire relazioni filogenetiche ma esclusivamente identificare le specie conosciute ed aiutare ad identificare quelle nuove, fornendo un primo segnale molecolare che, successivamente, approfonditi studi morfologici ed ecologici potranno confermare o meno.

Inoltre esistono altri limiti all'applicazione della tecnica, ma essi restano, secondo l'autore, dei casi talmente circoscritti da non influenzare la globalità del progetto. Essi riguardano quattro aspetti:

- *Gruppi con bassa divergenza di sequenza.* Potrebbero esistere dei gruppi in cui la divergenza tra specie vicine risulti inferiore alla soglia del 2% rendendo problematica la loro risoluzione come già evidenziato nel phylum Cnidaria dallo stesso Hebert.
- *Specie giovani.* Le specie che si fossero separate da un tempo inferiore al milione di anni potrebbero non aver accumulato un numero sufficiente di mutazioni per poterle distinguere. Infatti assumendo un tasso di divergenza del 2% per milione di anni, su un frammento di 600bp si saranno generati solo 12 siti diagnostici.
- *Ibridizzazione.* La presenza di specie differenti che si incrocino tra loro non consentirebbe, ovviamente, in nessun caso la corretta assegnazione di specie.
- *Presenza di pseudogeni.* Nel caso in cui il materiale di partenza fosse molto degradato potrebbe avvenire la coamplificazione di frammenti di mitocondri presenti nel nucleo (NumtS).

Nel febbraio 2005 si è svolta a Londra la prima conferenza ufficiale del progetto “ The First International Barcoding of Life Conference” per una prima valutazione dei risultati ottenuti per i diversi taxa. A questa seguirono il congresso a Taipei nel 2007 e Città del Messico nel 2009 e la Conferenza Europea a Braga nel giugno 2010. Nel corso di questi anni sebbene siano emerse difficoltà per alcuni gruppi tassonomici (per esempio Whitworth et al. 2007; Shearer & Coffroth 2008), è stato tuttavia possibile affermare il successo del frammento COI per i metazoi, e sono stati istituiti numerosi progetti dedicati all'esplorazione di specifici gruppi tassonomici, o biomi particolari. Ne sono la conferma (www.barcodingbirds.org , www.fishbol.org , www.lepbarcoding.org , www.polarbarcoding.org , www.barcoding.si.edu/major_projects.html).

Il DNA barcoding nelle piante

Premessa

Il genoma mitocondriale delle piante possiede caratteristiche diverse dagli animali, tali che non lo rendono adatto come barcode:

- basso tasso evolutivo e presenza di ricombinazione intramolecolare (Mower et al. 2007)
- presenza di introni
- frequente trasferimento di geni verso il nucleo

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:

- sufficiente variabilità interspecie per discriminare ma basso livello di variabilità intraspecie (barcoding-gap)
- possibilità di utilizzare primers universali
- facilità tecnica di sequenziamento e quindi assenza di regioni omopolimeriche o con ripetizioni
- lunghezza tale da permettere l'amplificazione con una sola reazione
- facilità di analisi, cioè assenza o un numero molto limitato di inserzioni o delezioni

In un primo momento la regione spaziatrice interna dei geni ribosomiali nucleari (ITS) fu presa in considerazione in quanto molto variabile nelle angiosperme (Kress et al. 2005). Tuttavia essa presenta delle caratteristiche evolutive e tecniche che la rendono problematica (Felinier et al. 2007). Infatti, sebbene questi geni rappresentino una famiglia genica che, sottoposta alla cosiddetta "evoluzione concertata", dovrebbe presentare uniformità all'interno dell'individuo e della specie, è tuttavia nota la presenza di una certa variabilità anche all'interno degli individui. Quindi non è possibile escludere, soprattutto in presenza di materiale degradato, ma non solo, l'amplificazione di geni paraloghi. Per questo motivo tale regione venne, in un primo momento, esclusa. Ciononostante, vedremo come parecchio tempo dopo, essa sarà parzialmente riconsiderata in particolare per quei generi che si presentano ostici.

A questo punto il genoma cloroplastico, le cui caratteristiche in qualche modo lo rendono più simile al mitocondrio dei metazoi, restò l'unico candidato e cominciò quindi una fase iniziale nella quale molte regioni cloroplastiche sia codificanti che non codificanti furono esplorate per poterne valutare la potenzialità. Apparve chiaro fin da subito che, a causa del tasso evolutivo del cloroplasto che è decisamente minore del mitocondrio animale, una sola regione non sarebbe stata sufficiente per discriminare le specie vegetali, tanto che la ricerca di un singolo marcatore come barcode venne paragonata alla ricerca del Santo Graal (Rubinoff *et al.*, 2006).

Il genoma cloroplastico

I cloroplasti delle piante verdi sono organelli citoplasmatici che ospitano vari pigmenti ed enzimi coinvolti nella funzionalità dell'apparato fotosintetico.

Fin dal secolo scorso si osservò che alcune pigmentazioni venivano ereditate per via materna, sollevando l'ipotesi che i cloroplasti potessero contenere e quindi trasportare geni, tuttavia la presenza di DNA fu dimostrata solo nel 1962.

Il genoma cloroplastico detto plastoma, si è conservato nel corso dell'evoluzione, infatti la struttura e l'organizzazione è nel complesso la stessa sia nelle piante vascolari che nelle alghe verdi (Sugiura, 1989).

Esso si è originato da un cianobatterio ancestrale penetrato per endosimbiosi nella cellula eucariotica (Curtis and Clegg, 1984; Delwiche et al. 1995; Barbrook et al., 1998; Turmel et al. 1999), tuttavia il genoma dei cianobatteri contiene più di tremila potenziali geni codificanti, mentre il plastoma odierno ne contiene circa centoventicinque. Dunque la maggior parte dei geni dell'ancestro endosimbiontico sono stati persi o trasferiti nel nucleo (Martin and Hermann, 1998; Race et al., 1999). Infatti esperimenti condotti su *Arabidopsis thaliana* mostrano come siano presenti sul cromosoma 4 numerose proteine altamente simili a quelle presenti nel cianobatterio *Synochocystis sp.*. Poiché la struttura e l'organizzazione del DNA di cloroplasto (cpDNA) è, come già detto, molto uniforme si ipotizza che il trasferimento del materiale genetico verso il nucleo deve sia avvenuto molto precocemente nell'evoluzione, anche se diversi studi provano che questo processo sia ancora in atto (Baldauf et al. 1990).

Ciascun plastidio contiene copie multiple di genoma cloroplastico; questa quantità varia a seconda dello stadio di sviluppo della foglia e del cloroplasto stesso. Un cloroplasto maturo contiene in genere 200-300 molecole che sono organizzate in zone discrete dette *nucleoidi* (Figura 2).

I CLOROPLASTI

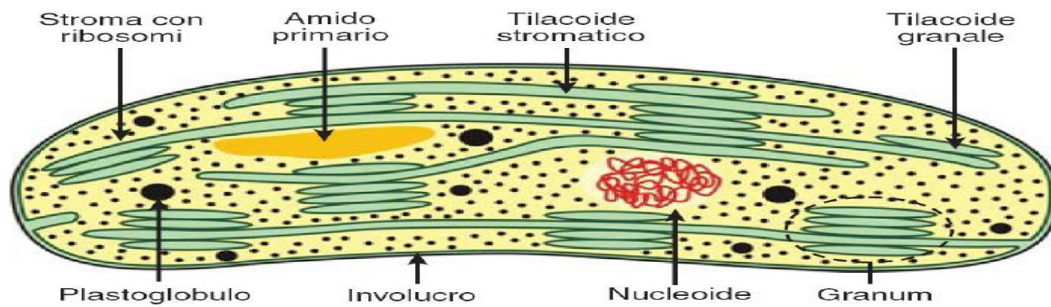


Figura 2. Organizzazione generale dei cloroplasti

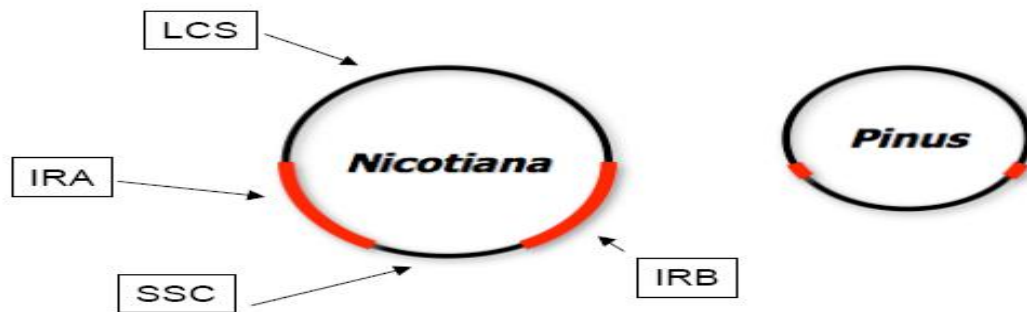
Il plastoma è una molecola circolare con un range generale che oscilla tra le 70 Kb (Epifagus-pianta parassita non fotosintetica) e 2000 Kb (Acetabularia), mentre nelle Angiosperme troviamo valori compresi tra 120 e 160 kb (Figura 3).

La sua organizzazione tipica nelle Angiosperme, consiste in due ***inverted repeats (IRa e IRb)*** che contengono sequenze identiche ripetute ma orientate in senso inverso, separate da una ***large single-copy (LSC)*** e una ***small single-copy (SSC)***. Le principali variazioni nella taglia dei genomi nei vari taxa sono generate dalla diversa lunghezza delle regioni IR (Palmer, 1990; 1991) e non dalla presenza di geni addizionali, tuttavia la riduzione estrema o la perdita di tali regioni è un evento molto raro osservato in Pinus e pochi membri delle Fabaceae (Wakasugi et al., 1994).(Figura 4 A B)

Taxa	Genome size (In kb)	Inverse duplication (In kb)
Angiospermae		
<i>Nicotiana tabacum</i>	156	25
<i>Spinacia oleracea</i>	150	24
<i>Pelargonium hortorum</i>	217	76
<i>Pisum sativum</i>	120	Not present
<i>Epifragus virginiana</i>	70	22
<i>Oryza sativa</i>	134	21
Gymnospermae		
<i>Pinus</i>	120	Not present
<i>Ginkgo biloba</i>	158	17
Pteridophyta		
<i>Osmunda cinnamomea</i>	144	10
Bryophyta		
<i>Marchantia polymorpha</i>	121	10
Chlorophyta		
<i>Codium fragile</i>	85	Not present
<i>Chlamydomonas reinhardtii</i>	195	22
<i>Chlamydomonas moewusii</i>	292	41
Rhodophyta		
<i>Cyanophora paradoxa</i>	127	10
Chromophyta		
<i>Dictyota dichotoma</i>	123	5

Figura 3- Dimensioni del genoma cloroplastico nei diversi taxa

(A)



(B)

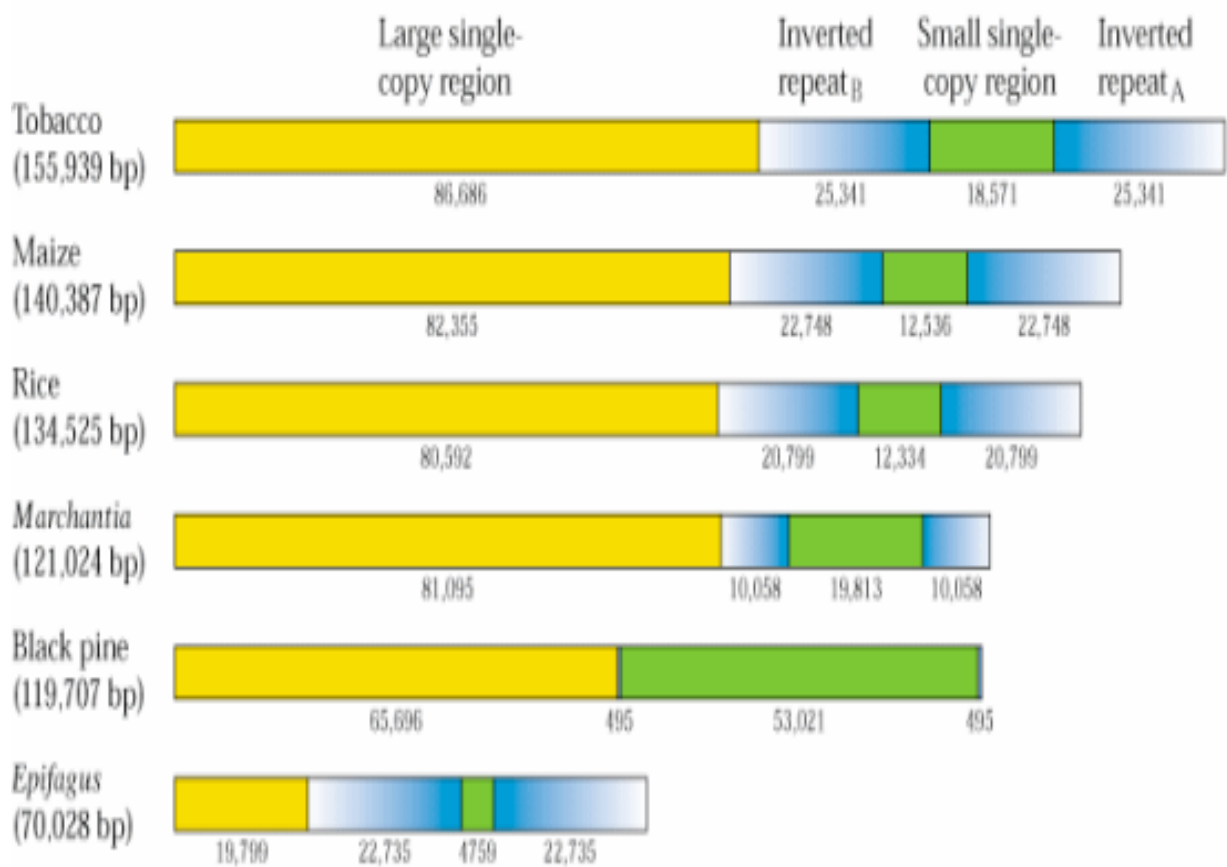


Figura 4 (A) (B): Organizzazione delle regioni genomiche *inverted repeats* (*IRa* e *IRb*), *large single-copy* (*LSC*) e *small single-copy* (*SSC*) nei diversi taxa.

Le regioni IRa e IRb contengono sempre i geni rRNA (indicati con *rrn* 16S, 23S, 4,5S e 5S) che quindi esistono in duplice copia, ma possono essere presenti anche altri tipi di geni che risulteranno anch'essi duplicati. In linea generale i geni codificanti possono contenere introni e appartengono a due principali categorie funzionali:

- geni coinvolti nell'apparato genetico (replicazione, trascrizione, traduzione)
- geni implicati nella funzione fotosintetica

Esistono comunque altri geni che codificano per la degradazione di proteine, la sintesi acidi grassi, la respirazione e alcune ORF con funzione ancora sconosciuta indicate come *ycfs* (hypothetical chloroplast frames). La nomenclatura dei geni si basa sul sistema utilizzato dai batteri, dove si usano le lettere minuscole con la sigla descrittiva della probabile funzione, inoltre se il prodotto proteico fa parte di un complesso di subunità, viene utilizzata una lettera maiuscola dell'alfabeto per distinguerla (es. *psbA* -> *gene fotosistema II subunità A*). (figura 5)

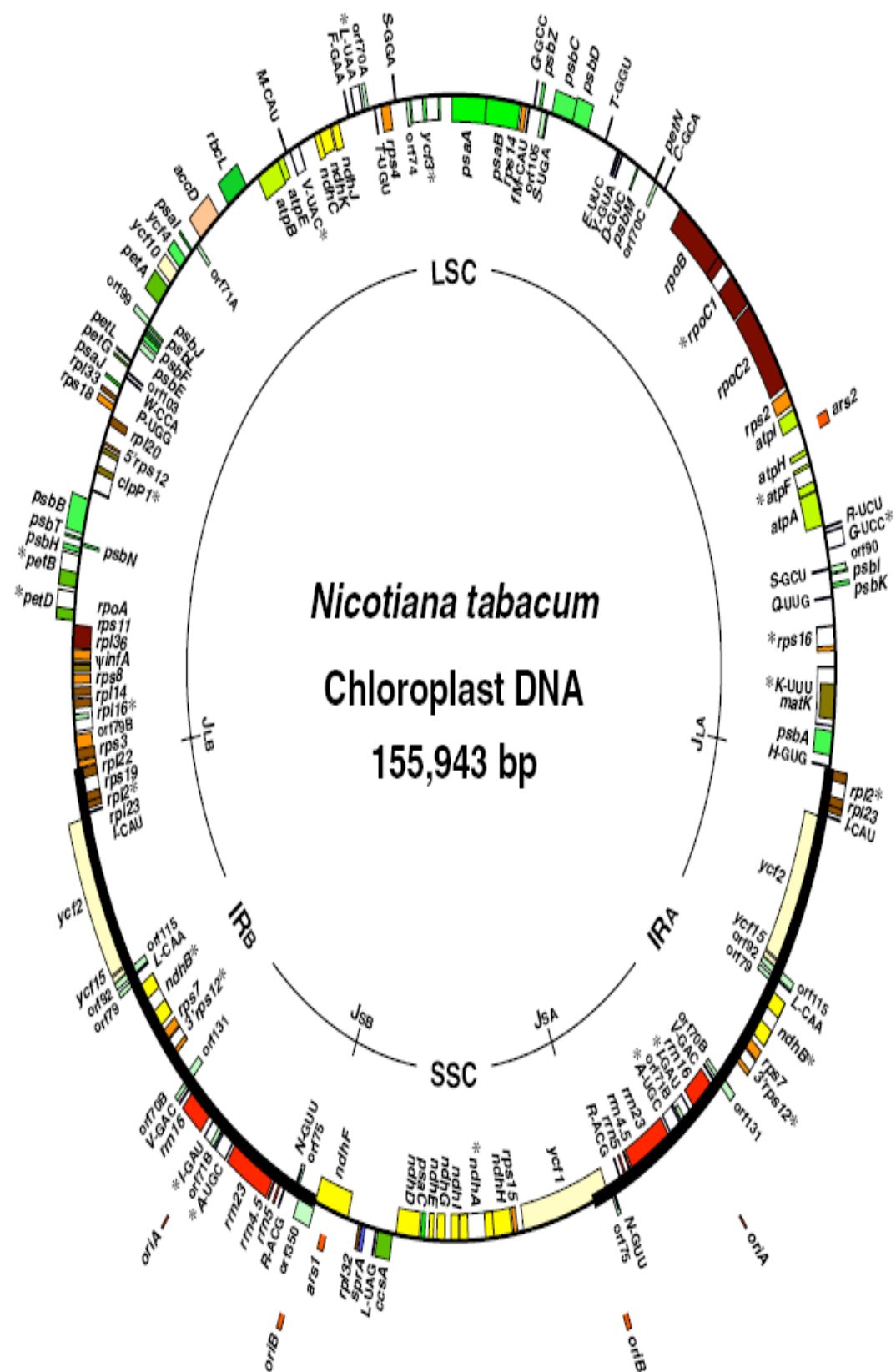


Fig. 5. Organizzazione generale del genoma cloroplastico

Il candidato barcode

Naturalmente, all'interno del CBOL fu creato un gruppo di lavoro specifico per le piante (Plant Working Group-PWG) formato da esperti nel settore e diretto da Peter Hollingsworth del Royal Botanic Garden di Edimburgo.

Ci fu quindi un periodo di sperimentazione nel quale cominciarono ad essere esplorate più regioni cloroplastiche in vari gruppi tassonomici e con vari approcci di studio, e come accennato in precedenza, già i primi studi evidenziarono come nelle piante (qualunque fosse il marcatore utilizzato) il potere discriminante di un singolo marcatore fosse nettamente inferiore a quello osservato nel regno animale (Kress and Erickson, 2007).

Interessantemente, in una valutazione del potere discriminante a livello di specie attraverso la monofilia, è stato osservato che a parità di PICs (siti informativi secondo il criterio evolutivo della massima parsimonia), negli animali, utilizzando come Barcode la COI, il successo era di oltre il 90%, mentre nelle piante del 46% con un singolo gene e circa il 70% con più marcatori uniti (Fazekas *et al.*, 2009).

Infatti, nel corso della Seconda Conferenza Internazionale del Barcoding tenutasi a Taipei nel 2007, alcuni gruppi cominciarono a proporre alcune combinazioni di marcatori come potenziali barcode (Pennisi, 2007):

1. Chase *et al.* (consortium led by Royal Botanic Garden Kew, UK)

rpoC1+rpoB+matK* OR *rpoC1+matK+trnH-psbA

2. Kim *et al.* (Korea University, Korea)

matK+atpF-H+psbK-I* OR *matK+atpF-H+trnH-psbA

3. Kress and Erickson (Smithsonian Institute, USA)

rbcL+trnH-psbA

Queste proposte ricalcano in qualche modo il cosiddetto “two-tiered approach” (Newmaster *et al.* 2006), che consiste nell’associazione di una regione codificante, la cosiddetta “anchor” capace di evidenziare relazioni tassonomiche più profonde come famiglia o genere, ed una non codificante, con tasso evolutivo più veloce che fornisca la risoluzione a livello di specie.

Tuttavia è importante sottolineare come talvolta i diversi marcatori siano caratterizzati da performance differenti nei diversi generi, rendendo ancora più complessa la situazione (Fazekas *et al.*, 2008).

Nonostante queste proposte, anche altre regioni continuarono ad essere utilizzate aumentando la quantità di informazioni a disposizione della comunità.

In un articolo del duemilanove (Hollingsworth ML *et al*, 2009) si è cercato di riassumere i risultati dei lavori più importanti sottolineando come spesso la comparazione dei dati non sia semplice. La figura 6 riporta una tabella dall' articolo sopra citato, dove sono schematizzate non solo le regioni utilizzate, ma anche le strategie di campionamento, successo di amplificazione, i dati di divergenza o variazione dei locus (come analizzati dagli autori) e infine il barcode proposto. Appare evidente come sia le strategie di campionamento che quelle di analisi siano differenti, rendendo difficile una valutazione critica e complessiva. In particolare, le principali strategie di campionamento possono essere ricondotte a tre tipi principali:

- **Floristic** questo tipo di approccio consiste nel campionamento di un solo individuo per specie a partire da una comunità vegetale caratteristica di una determinata area geografica; ci consente principalmente di valutare l'universalità dei primers e problematiche biomolecolari, ma non fornisce nessun dato sulla variabilità intra specie.
- **Species pairs** in questo caso vengono presi generi anche tassonomicamente lontani e all'interno di essi campionate coppie di specie; anche qui sarà principalmente valutata l'efficacia dei primers ma non saranno fornite informazioni sulla variabilità intraspecie, e saranno insufficienti quelle all'interno del genere.
- **Taxon-based** può essere considerato come l'unico che sia veramente in grado di fornire informazioni sull'efficacia del marcatore a livello di specie; consiste nella scelta di un certo gruppo tassonomico (spesso un genere) all'interno del quale, per ogni specie, vengono campionati più individui provenienti da diverse aree geografiche.

Nonostante queste ed altre discrepanze tra i dati generati, gli autori dell'articolo evidenziano sette regioni le cui caratteristiche sembrano essere promettenti come candidati barcode. Queste regioni sono: *rpoB*, *rpoC1*, *rbcL*, (codificanti molto conservate e facilmente allineabili), *matK* (una regione codificante con tasso evolutivo elevato), e *trnH-psbA*, *atpF-atpH*, *psbK-psbI* (regioni inter-geniche ad evoluzione molto rapida).

Nel novembre duemilanove, il sito del CBOL rende noto un comunicato ufficiale sul Barcode delle piante. Nei due mesi precedenti, infatti erano state presentate al CBOL due proposte ufficiali, entrambe ben supportate, da parte di due gruppi diversi che proponevano 2 set differenti di marcatori. La prima proposta (Agosto 2009) è del Working Plant Group (WPG) che riassume quanto da loro pubblicato (CBOL Plant Working Group 2009) e consiste in 2 loci

rbcL* e *matK; in particolare viene sottolineato che, sebbene *matK* presenti difficoltà di amplificazione e necessità dello sviluppo di ulteriori primers, viene preferito alla regione non-codificante *trnH-psbA* che pur avendo forti potenzialità presenta problemi tecnici generati dalla presenza di omopolimeri che rendono impossibile l'assemblaggio automatico della sequenza.

La seconda proposta (Settembre 2009) proviene da un gruppo coordinato da Kress ed Erickson (Smithsonian Institute, USA) e propone l'adozione di 3 loci: ***rbcL* , *matK*, e *trnH-psbA*** evidenziando come 3 loci sarebbero più efficaci di due. Inoltre, le caratteristiche di *matK* e *trnH-psbA* vengono valutate nel dettaglio concludendo che sia il problema dei primers legato a *matK* sia quello dell'analisi di sequenza per *trnH-psbA* dovranno essere risolte. Tuttavia viene sottolineato come le problematiche legate a *trnH-psbA* non possono essere considerate insuperabili (o comunque non più gravi di quelle di *matK*) e tali da escluderlo.

Per decidere tra queste due proposte il CBOI nomina tre revisori indipendenti per la valutazione delle proposte e nel novembre arriva la decisione che ufficialmente approva la proposta del WPG con i due marcatori ***rbcL* e *matK***. Tuttavia esiste un' importante postilla: l'Esecutivo del CBOI accoglie la raccomandazione dei revisori esterni i quali consigliavano una finestra temporale di diciotto mesi per raccogliere ulteriori dati su *matK* *trnH-psbA* ed anche altre potenziali regioni non codificanti. Questo nella speranza che nuovi dati possano, forse, rendere più semplice la decisione definitiva, anche perché l'attuale valore di successo (70% circa), viene ritenuto troppo basso e quindi assolutamente da incrementare.

I diciotto mesi scadranno nel maggio 2011, attendiamo quindi per quella data un nuovo capitolo del dibattito.

Fig. 6: Schema dei risultati ottenuti nei principali studi barcode. (Hollingsworth ML *et al*, 2009)

Study	Regions compared	Sampling strategy	Universality (% success)	Sequence divergence /variation	Barcode recommendation
Kress <i>et al.</i> 2005	<i>atpB-rbcL</i> , ITS, <i>psbM-trnD</i> , <i>trnC-ycf6</i> , <i>trnH-psbA</i> , <i>trnL-F</i> , <i>trnK-rps16</i> , <i>trnV-atpE</i> , <i>rpl36-rps8</i> , <i>ycf6-psbM</i>	19 individuals/19 species from 7 angiosperm families	<i>trnH-psbA</i> , <i>rpl136-rpf8</i> , <i>trnL-F</i> = 100% <i>trnC-ycf6</i> , <i>ycf6-psbM</i> = 90% Other regions = 73–80%	% sequence divergences: ITS (2.81%) <i>trnH-psbA</i> (1.24%) <i>trnH-psbA</i> had ~ 2 × sequence divergence of other plastid regions <i>trnH-psbA</i> >> <i>rbcL</i>	ITS and <i>trnH-psbA</i>
	ITS, <i>rbcL</i> *, <i>trnH-psbA</i>	83 individuals/83 species from 50 families	<i>trnH-psbA</i> = 100% <i>rbcL</i> = 95% ITS ≤ 88%		
Kress & Erickson 2007	<i>accD</i> , ITS1, <i>ndhJ</i> , <i>matK</i> , <i>trnH-psbA</i> , <i>rbcL</i> , <i>rpoB</i> , <i>rpoC1</i> , <i>ycf5</i>	96 individuals/96 species from 43 families of land plants	<i>trnH-psbA</i> , <i>trnK</i> = 95% <i>trpoC1</i> = 90% <i>accD</i> , <i>rpoB</i> = 80% <i>ndhJ</i> = 70%, ITS1 = 60% <i>ycf5</i> = 50% <i>matK</i> = 40%	ITS (5.7%) <i>trnH-psbA</i> (2.69%) <i>rpoB</i> (2.05%) <i>ycf5</i> (1.55%) <i>rpoC1</i> (1.38%) <i>rbcL</i> (1.29%) <i>accD</i> (1.2%) <i>matK</i> (1.13%) <i>ndhJ</i> (0.2%)	<i>rbcL</i> and <i>trnH-psbA</i>

Fig.6

Sass <i>et al.</i> 2007	<i>accD</i> , ITS, <i>ndhJ</i> , <i>matK</i> , <i>trnH-psbA</i> , <i>rpoB</i> , <i>rpoC1</i> , <i>ycf5</i>	21 individuals/21 species from 10 genera of cycads, more individuals for some regions (up to 96)	<i>rpoC1</i> , <i>ycf5</i> = 100% ITS = 100%?† <i>accD</i> = 96% <i>ndhJ</i> = 57% <i>rpoB</i> = 33% <i>matK</i> = 24% <i>trnH-psbA</i> double- banded in most samples	ITS most variable; quantitative figures on relative sequence divergence of other regions not given, other than c. 10% of bases variable in each region	ITS (considered most promising)
Newmaster <i>et al.</i> 2008	<i>accD</i> , <i>matK</i> , <i>trnH- psbA</i> , <i>rbcL</i> , <i>rpoB</i> , <i>rpoC1</i> , UPA	40 individuals/8 species in Myristicaceae	<i>trnH-psbA</i> , <i>rpoC1</i> , UPA = 100% <i>rpoB</i> , <i>accD</i> , ≥ 95% <i>rbcL</i> = 90% <i>matK</i> required primer redesign, then = 98%	Uncorrected interspecific p- distances: <i>trnH-psbA</i> (0.060) <i>matK</i> (0.042) <i>accD</i> (0.003) <i>rpoC1</i> (0.002) <i>rbcL</i> (0.002) <i>rpoB</i> (0.001) UPA (0.001)	<i>matK</i> and <i>trnH- psbA</i>
Lahaye <i>et al.</i> 2006a	<i>accD</i> , <i>ndhJ</i> , <i>matK</i> , <i>trnH-psbA</i> , <i>rbcL</i> , <i>rpoB</i> , <i>rpoC1</i> , <i>ycf5</i>	172 individuals/86 species total (consisting of 71 individuals/48 orchid species + 101 individuals/38 species from 13 angiosperm families)	In the orchids, <i>ycf5</i> did not amplify and <i>ndhJ</i> amplification was patchy. All other regions = 95–100%	K2P interspecific sequence divergence: <i>trnH-psbA</i> (0.0216) <i>matK</i> (0.0125) <i>ycf5</i> (0.01) <i>rbcL</i> (0.0079) <i>rpoB</i> (0.0061) <i>ndhJ</i> (0.0046) <i>accD</i> (0.0038) <i>rpoC1</i> (0.0019)	<i>matK</i> (or <i>matK</i> and <i>trnH-psbA</i>)¶
Fazekas <i>et al.</i> 2008	<i>cox1</i> , 23S rDNA, <i>rpoB</i> , <i>rpoC1</i> , <i>rbcL</i> , <i>matK</i> , <i>trnH-psbA</i> , <i>atpF-atpH</i> , <i>psbK-psbI</i>	251 individuals/92 species from 32 genera of land plants	23S rDNA = 100% <i>rbcL</i> = 100% (2 primer pairs used) <i>trnH-psbA</i> = 99% <i>rpoC1</i> = 95% (3 primer pairs used) <i>rpoB</i> = 92% (5 primer pairs used) <i>matK</i> = 88% (10 primer pairs used) <i>psbK-psbI</i> = 85% <i>cox1</i> = 72% <i>atpF-atpH</i> = 65%	No. of parsimony- informative characters: <i>matK</i> (386) <i>trnH-psbA</i> (350) <i>atpF-atpH</i> (306) <i>psbK-psbI</i> (263) <i>rbcL</i> (242) <i>rpoB</i> (179) <i>cox1</i> (146) <i>rpoC1</i> (134) 23S rDNA (19)	Broadly equivalent performance from various combinations of loci; suggested selecting 3–4 regions from: <i>rbcL</i> , <i>rpoB</i> , <i>matK</i> , <i>trnH- psbA</i> , <i>atpF-atpH</i>
Chase <i>et al.</i> 2007	In preparation, empirical data currently unpublished				<i>rpoC1</i> , <i>rpoB</i> and <i>matK</i> or <i>rpoC1</i> , <i>matK</i> and <i>trnH-psbA</i>
Kim <i>et al.</i> cited in Pennisi 2007	In preparation, empirical data currently unpublished				<i>matK</i> , <i>atpF-atpH</i> and <i>psbK-psbI</i> or <i>matK</i> , <i>atpF-atpH</i> and <i>trnH-psbA</i>

Il concetto di specie e DNA barcoding

La classificazione

Utilizzare una nomenclatura standard è necessario per consentire la comunicazione sia in generale che in ambiti tecnici specifici. Nell'ambito delle scienze naturali all'inizio del settecento i naturalisti credevano che Dio avesse creato le specie secondo un piano ordinato e cominciarono quindi a catalogare le opere del Creatore credendo di scoprire la vera classificazione naturale.

Il modello di nomenclatura utilizzato allora (e usato ancora oggi) fu messo a punto da C. Linneo. Egli introdusse la nomenclatura binomia secondo la quale ogni specie è indicata con una coppia di nomi: il primo è il nome del genere mentre il secondo è l'epiteto della specie. Le specie erano successivamente raggruppate secondo una organizzazione gerarchica in generi, famiglie ect., sulla base di somiglianze e caratteristiche comuni riconducibili in ogni caso al disegno divino.

Verso la metà dell'ottocento Darwin, sulla base delle osservazioni e studi fatti presso le isole Galapagos, formulò l'ipotesi secondo cui le somiglianze e i caratteri in comune tra le specie dipendessero non da una assegnazione divina, bensì dal fatto che esse condividevano un antenato comune da cui queste caratteristiche erano state ereditate. Tuttavia nel corso del tempo si erano accumulate anche molte differenze tanto da rendere le specie diverse ai nostri occhi. Queste differenze sono generate secondo Darwin da una forza detta “selezione naturale” che agisce in modo tale che le caratteristiche più vantaggiose per la sopravvivenza siano preferite e quindi trasmesse alla discendenza.

Quindi, con molto coraggio, Darwin arrivò ad ipotizzare l'esistenza di un progenitore comune a tutte le forme viventi e la possibilità di un grande “albero della vita” di cui fanno parte tutte le specie e come scrisse Darwin stesso “ ...le foglie e i ramoscelli germoglianti possono rappresentare le specie esistenti; i rametti che si sono sviluppati negli anni precedenti possono rappresentare la lunga successione delle specie estinte....”

Dunque in questo caso la struttura gerarchica della classificazione sarebbe generata da un processo evolutivo avvenuto nel corso del tempo anziché dal disegno divino.

Nonostante le polemiche, la teoria di Darwin fu accettata dalla maggior parte dei biologi e oggi possiamo affermare che l'evoluzione è un fatto. Tuttavia la spiegazione di come avvengano le modificazioni e come gli antenati abbiano originato i diversi discendenti è ancora “teorica”.

Infatti, sebbene l'ipotesi di Darwin sulla selezione naturale sia corretta, oggi sappiamo che le cause dell'evoluzione sono molto più complesse e comprendono concetti come ricombinazione, flusso genico, isolamento, deriva genetica, etc..

Le specie

La concezione di specie non è uniforme tanto che, in tempi e contesti diversi, sono state proposte molte definizioni di “specie”, e le principali sono riassunte in tabella 1.

Linneo e gli altri tassonomisti del passato possedevano quello che Mayer definisce nozione *tipologica* o essenzialista di specie, seconda la quale gli individui appartengono ad una data specie se si conformano in modo sufficiente ad un “tipo” che possiede caratteri morfologici ritenuti come proprietà “essenziali” fissate.

Questa nozione tipologica di specie fu messa in crisi dallo studio della variazione entro e tra le popolazioni, e attualmente i concetti più utilizzati sono le varianti del *concetto filogenetico* e il *concetto biologico* di specie. Fondamentalmente, nel concetto filogenetico la specie viene considerata come il prodotto di una storia evolutiva di divergenza che ha infine generato la specie, mentre nel concetto biologico vengono sottolineati i processi biologici adottati dalle specie stesse.

Tuttavia, alcuni autori sottolineano come, molto probabilmente, non esista una definizione idonea in tutti i contesti, in particolare in una review sulle specie (Coyne et al 2004) si ribadisce come una specie possa essere diversamente definita in funzione degli obiettivi prefissati. Infatti è molto improbabile che un singolo concetto di specie possa adattarsi a questa lista di scopi diversi:

- 1) consenta di classificare gli organismi in modo sistematico
- 2) corrisponda a gruppi discreti di individui simili
- 3) aiuti a capire come gruppi discreti di organismi compaiano in natura
- 4) rappresenti i prodotti della storia evolutiva
- 5) si applichi alla varietà più ampia possibile di organismi

Ciononostante è necessario stabilire dei criteri oggettivi che consentano di identificare le specie per poterle distinguere nella pratica. I caratteri morfologici e altri caratteri fenotipici sono le caratteristiche degli organismi solitamente utilizzati per identificare le specie. Questi, opportunamente interpretati, servono come marcatori (*sintomi*) dell'isolamento riproduttivo o al contrario dell'esistenza di una interfecondità riproduttiva, ma raramente viene messa direttamente alla prova la propensione a incrociarsi o la capacità di produrre prole fertile.

In questo contesto, la cui complessità è stata solo accennata, si inserisce la tecnica del barcoding. In teoria essa, come definito dallo stesso Hebert, non ha lo scopo di definire o scoprire specie, bensì semplicemente quello di identificarle e quindi differenziarle le une dalle altre.

In effetti il concetto di specie non viene mai citato nelle varie descrizioni del metodo, tuttavia, per una tecnica che abbia tale scopo, la definizione di ciò che si vuole discriminare costituisce un punto fondamentale, soprattutto nel momento in cui si voglia quantificare il successo o il fallimento del metodo stesso.

Solitamente negli studi applicati ai vari taxon, il successo della tecnica viene attribuito quando è possibile discriminare le entità molecolari ottenute collocandole secondo la chiave di lettura della tassonomia classica. Quindi, implicitamente, nel momento della quantificazione del successo si definisce quale concetto di specie sottenda il progetto Barcode. Tuttavia, la tassonomia classica può essere considerata un compromesso tra complessità della natura e dei processi biologici, e la necessità dell'uomo di modellizzarla e semplificarla per comprenderla. Il risultato è quindi una visione imperfetta e statica delle specie, come viene talvolta evidenziato dai risultati molecolari di loci barcode.

In maniera quasi provocatoria potremmo dire che il metodo del DNA barcode, pur affermando di essere un nuovo modo per studiare, osservare e conoscere la biodiversità del nostro pianeta, lo fa poggiando le sue fondamenta su concetti e criteri vecchi e notoriamente imperfetti.

<i>Concetto biologico di specie</i>	Le specie sono gruppi di popolazioni naturali effettivamente o potenzialmente interfeconde che sono riproduttivamente isolati da gruppi simili (Mayr 1942)
<i>Concetto evolutivo di specie</i>	Una specie è una singola linea di discendenza (una sequenza antenato-discendente) di popolazioni o individui che mantiene una identità separata da altre linee simili e che possiede una propria tendenza evolutiva e un suo destino storico (Wiley 1978)
<i>Concetto filogenetico di specie</i>	(1) Una specie è un gruppo non riducibile (basale) di organismi evidentemente diverso da altri gruppi simili, entro il quale vi sono relazioni genealogiche di ascendenza e discendenza (Cracraft 1989) (2) Una specie è il più piccolo gruppo monofiletico di comune ascendenza (de Queiroz e Donoghue 1990)
<i>Concetto genealogico di specie</i>	Le specie sono gruppi “esclusivi” di organismi, cioè gruppi i cui membri sono tutti più strettamente imparentati tra loro che con ogni altro organismo al di fuori del gruppo (Baum e Shaw 1995)
<i>Concetto di specie per “recognition”</i>	Una specie è la popolazione più inclusiva di individui di organismi biparentali che condividono un comune sistema di fecondazione (Paterson 1985)
<i>Concetto coesivo di specie</i>	Una specie è la più inclusiva popolazione di individui che hanno il potenziale per una coesione fenotipica attraverso meccanismi intrinseci di coesione (Templeton 1989)

Tabella 1 Alcuni concetti di specie. (Da Futuyma Douglas J. (2005): Evolution)

L'analisi dei dati

L'analisi dei dati nel metodo barcode ha il solo scopo di verificare se la sequenza ottenuta con i marcatori standard definiti dal protocollo corrisponda ad un *oggetto tassonomico* definito dalla classica tassonomia e nomenclatura Linneana. Questo presuppone l'esistenza di un database completo di sequenze vouchers relative a tutte le specie, all'interno del quale effettuare la ricerca del "match" con la nostra sequenza ignota. Attualmente non è chiaro quale criterio sarà adottato per effettuare questa ricerca; per le sequenze presenti su GenBank, sicuramente questo processo già ora avviene attraverso gli algoritmi di ricerca di similarità come Blastn o MegaBlast, tuttavia l'utilizzo del BIT score (che tiene conto della percentuale di identità e di E-value) associato ad ogni sequenza, non è una misura rigorosa della similarità genetica poiché dipendente dalla grandezza del database in cui si effettua la ricerca.

Quindi, per rispondere a questa esigenza specifica è nato Barcode of Life Data Systems, BOLD ([http:// www.barcodinglife.org/](http://www.barcodinglife.org/)) che costituisce il maggior tool del progetto Barcode.

Dapprima esclusivo contenitore per le sequenze di COI, è stato successivamente ampliato per ospitare sia le sequenze ufficiali ITS dei funghi che per le piante le sequenze rbcL e matK; tuttavia gli algoritmi di ricerca utilizzati al suo interno sono ancora oggetto di sperimentazione e quindi in costante aggiornamento.

Hebert *et al.* (2004b) propose di validare il successo della tecnica attraverso la valutazione delle distanze genetiche inter e intra specie e la costruzione di un albero Neighbour Joining, tuttavia, allo stato attuale, soprattutto in campo vegetale, non esiste un metodo di analisi dei dati condiviso dalla comunità di coloro che lavorano ai progetti barcode, e spesso gli autori sperimentano diversi metodi analitici sullo stesso dataset ottenendo, nella maggior parte dei casi, risultati identici. E' possibile quindi affermare che non esiste un "miglior metodo" che consenta "*a priori*" di aumentare le potenzialità di identificazione del locus utilizzato, ma che invece, a seconda delle caratteristiche dei dati da analizzare, sia possibile e anzi necessario utilizzare un metodo anziché un altro.

In genere gli autori analizzano il proprio dataset di sequenze identificando in maniera indipendente i campioni che andranno a costituire il database dei vouchers a cui assoceranno la relativa *sequenza barcode*. Questo passaggio ha lo scopo di contribuire al completamento della banca dati centrale che raccoglierà le informazioni relative a tutte le specie e che è, chiaramente, la condizione *necessaria* per il funzionamento del DNA barcoding. Al contempo i dati ottenuti vengono analizzati per valutare se il pattern di variabilità osservato è tale da poter

predire il successo del metodo, intendendo per successo la separazione delle sequenze in gruppi inter e intra specie che corrispondano alla tassonomia corrente.

Alcuni protocolli di analisi sono molto più utilizzati di altri forse anche a causa della loro minore o maggiore complessità.

Un primo tipo di approccio è quello basato sui *caratteri* (*character-based approach*); in questo tipo di analisi le entità molecolari generate vengono identificate e raggruppate in funzione della presenza di caratteri (o combinazioni di caratteri) condivisi, che, trattandosi di una stringa di DNA, saranno dei nucleotidi. Quindi, il concetto fondamentale è che i membri di un determinato gruppo tassonomico condividano degli attributi caratteristici (CAs), per esempio un certo polimorfismo, che sono assenti in altri gruppi (Rach *et al.*, 2008).

Uno gruppo di metodi senza dubbio più utilizzati sono quelli basati su un approccio *filogenetico* (*tree-based approach*), che generano come risultato un albero. Purtroppo questo tipo di rappresentazione non aiuta a placare le critiche nei confronti della tecnica (come accennato in precedenza), aumentando la confusione tra *DNA barcoding* e *DNA taxonomy*. Infatti, sebbene i due approcci siano palesemente diversi, spesso in letteratura i due termini perdono il loro reale significato a causa dell'uso improprio o ambiguo che ne viene fatto. Brevemente, il *DNA barcoding* vuole essere un tool esclusivamente identificativo, ossia separare le specie a partire da un enorme database contenente la struttura tassonomica classica associata alle sequenze barcode standard; differentemente, nel caso della *DNA taxonomy*, esiste la volontà di circoscrivere e delineare le sequenze in gruppi che, non necessariamente, devono corrispondere alla struttura gerarchica propria della tassonomia classica, tanto che, in alcuni casi ci sarà una corrispondenza con la nomenclatura Linneana, in altri questa corrispondenza non esisterà e saranno quindi attribuiti altri identificativi. Inoltre l'acquisizione dei dati può avvenire attraverso uno o più marcatori provenienti da diverse regioni genomiche non legate a loci standard come nel caso del barcode.

Quindi, nel *tree-based approach* vengono generati degli alberi (spesso denominati dendrogrammi), attraverso l'utilizzo di vari criteri che vanno dal Neighbour Joining ai metodi Bayesiani, e la valutazione del successo viene spesso effettuata attraverso la monofilia di individui co-specifici supportata da alti valori di bootstrap (Fazekas *et al* 2008).

Un altro metodo utilizzato per la valutazione del successo prevede l'utilizzo degli algoritmi per la ricerca di similarità BLAST, BLAT, e megaBLAST. Questi metodi consentono di modificare alcuni parametri relativi al grado di similarità desiderato, e se il pattern di variabilità mostra opportuni valori inter e intra specie, è possibile individuare una soglia che consenta non

solo di separare le specie, ma anche di raggruppare insieme gli individui della stessa specie ma di diverse provenienze geografiche.

Infine ricordiamo i metodi basati sulle distanze (*p-distance* o *Kimura-2P*), che in base a quanto scritto nella proposta ufficiale del PlantWG, corrispondono ai metodi che sono stati utilizzati come criterio per la valutazione dei sette loci candidati. In effetti, forse proprio attraverso questo approccio, si evidenzia quali caratteristiche di variabilità debba possedere il locus barcode, cioè, come già accennato, una bassa variabilità nucleotidica entro le specie contrapposta ad una alta variabilità tra le specie (secondo Hebert circa 10 volte maggiore). Più nel dettaglio diciamo che il successo viene attribuito se la più piccola distanza interspecie è più grande della massima distanza intraspecie. In tal modo viene valutata la presenza del cosiddetto “*barcoding-gap*”, anche se, attraverso un lavoro di comparazione di moltissimi dati già pubblicati relativi ad animali e piante, è stato evidenziato che nel regno vegetale sia molto più frequente una sovrapposizione delle distanze inter e intra specie, riducendosi così il potenziale discriminatorio dei loci (Figura 7 (Meyer *et al* 2005)). Inoltre, in passato venivano utilizzati i valori medi delle distanze inter e intraspecie, ma alcuni autori hanno evidenziato come questo determini una sovrastima del *barcoding gap* (Meier *et al* 2008).

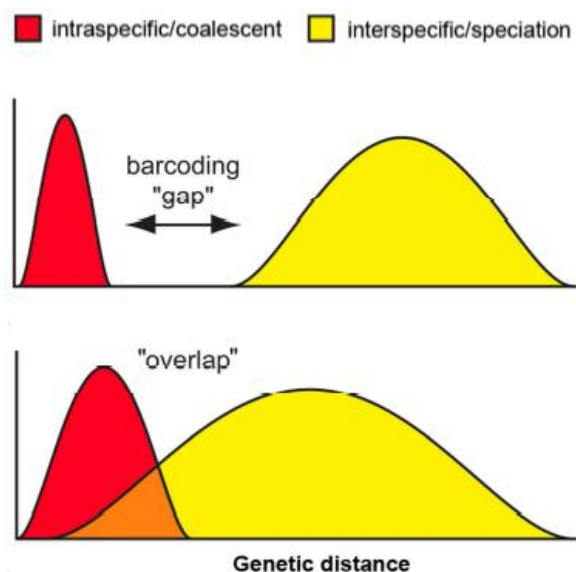


Figura 7: Rappresentazione grafica del *barcoding gap*. (Meyer and Paulay, 2005).

2 IL GENERE QUERCUS

INTRODUZIONE

Le querce sono abbondanti e spesso dominanti nei paesaggi Mediterranei dell'emisfero settentrionale e per questo motivo sono rappresentative della cultura, dell'economia e del paesaggio dell'uomo moderno. Inoltre, esse da tempo sono uno dei gruppi vegetali forestali che ha più riscosso l'interesse degli studiosi circa il concetto di specie, dovuto alla forte attitudine all'incrocio che queste esibiscono, stimolando le ricerche sulla biosistemica, l'ibridazione interspecifica, l'introggressione e la filogeografia.

Le problematiche del genere *Quercus*:

Il genere *Quercus* presenta grosse difficoltà tassonomiche poiché, a causa dell'elevato polimorfismo, ne sono stati descritti numerosissimi taxa che spesso sono risultati di incerta collocazione. L'incertezza deriva principalmente dalla difficoltà (o dal disaccordo) dei vari studiosi nel definire una chiave sistematica incontrovertibile e ripetibile, infatti, i caratteri differenziali nelle querce vanno esaminati in maniera alquanto differente rispetto a quanto è usuale negli altri gruppi di vegetali. Occorre utilizzare criteri di identificazione che impieghino caratteri qualitativi (e non quantitativi), che non siano influenzati dall'ambiente, dando origine a fenomeni di evoluzione parallela o convergente che disturbino il segnale filogenetico, o dall'età degli individui (Tucker, 1974; Jones, 1986). Viene pertanto a mancare una chiara discriminazione tra moltissimi taxa, a fronte di diversi sistemi nomenclaturali, con contrastanti suddivisione infra-generica, inter- e sub-specifica, molto spesso compilati a livelli geografici troppo ristretti. A parziale giustificazione di questa situazione bisogna tuttavia ricordare la forte (la cui portata è spesso sconosciuta) attitudine delle querce ad ibridarsi inter-specificamente (Burger, 1975) e la forte plasticità morfologica dell'intero genere nei confronti delle mutevoli condizioni ambientali (Van Valen, 1976). La determinazione delle specie è infatti resa difficile dalla presenza di moltissimi ibridi, che riducono la delimitazione delle singole specie l'una rispetto all'altra; la corrispondenza tra caratteri morfologici ed ecologici risulta invece complessa soprattutto laddove più specie affini vivono in contatto, presentando serie complete di tutti gli intermedi lungo i clini ecologici.

Se nel 1911 un botanico siciliano scriveva: “...mi è lecito affermare che il genere *Quercus* rappresenta la perfetta negazione del concetto di specie...si tratta di un immenso caos...” (Borzì, 1911), la questione non è migliorata con le nuove conoscenze: basti pensare che negli anni ‘30 la Camus censiva circa 800 taxa (Camus, 1936-38), in massima parte attribuibili alla variabilità esistente all’interno del complesso che comprende *Quercus pubescens* Willd. e *Quercus petraea* (Matt.) Liebl. e ancora negli anni ‘80 Krüssmann (1984-86) riportava circa 450 specie. Nixon (1993) riferisce che esistono, a seconda dei criteri tassonomici adottati, da 394 a 448 specie distribuite nell’emisfero boreale: esse occupano soprattutto le regioni temperate di Nord America, Europa e Asia e si spingono fino ad alcune aree tropicali e subtropicali del Nord Africa, America centro-meridionale e Asia.

Anche se la tendenza più recente è quella di ridurre il numero delle specie riunendo insieme quelle con caratteri molto simili, la più recente checklist del genere *Quercus* classifica 531 specie (Govaerts & Frodin, 1998), a confronto con le circa 300 specie descritte nella prima lista completa definita da De Candolle (1862).

Le classificazioni Sub Generiche proposte

Un primo esame della classificazione delle specie del genere *Quercus* può essere attribuito ad Ørsted (1867). Egli ha suddiviso il genere in quattro sottogeneri tre dei quali composti da specie del Vecchio Mondo: *Lepidobalanus* con stilo corto e capitato e con squame della cupola appressate (come *Q. ilex*, *Q. petraea*, *Q. pubescens*, *Q. robur*); *Heterobalanus* con stili lunghi ed appuntiti e squame come in *Lepidobalanus* (comprendente solo *Q. semecarpifolia*); *Cerris* con stili lunghi, stretti e appuntiti e squame della cupola divelte (*Q. cerris*, *Q. suber*, *Q. trojana*).

In seguito la Camus (1936-38) ha preferito suddividere il genere *Quercus* in due sottogeneri: *Cyclobalanopsis* Schneid., che comprende querce con cupole formate da serie di squame fuse assieme, e *Euquercus* Hickel & Camus, con le squame libere. Nel primo sottogenere sono collocate specie esclusive delle Montagne dell'Asia orientale temperata o tropicale; il secondo sottogenere è suddiviso ancora nelle sezioni: *Erythrobalanus*, *Protobalanus* e *Macrobalanus* (del continente americano); *Lepidobalanus* (dell'emisfero boreale) e, infine, *Cerris* e *Mesobalanus* dell'Eurasia (tabella 2).

SECT. CERRIS SPACH	SECT. LEPIDODALANUS ENDL.
SUBSECT. COCCIFERAE	SUBSECT. ILEX
Q. alnifolia Poech.	Q. ilex L.
Q. coccifera L.	
Q. calliprinos Webb	SUBSECT. GALLIFERAE
Q. rigida Willd.	Q. mirbeckii Dur.
Q. pseudorigida Kotschy	Q. alpestris Boiss.
Q. aucheri Jaub. et Spach.	Q. fruticosa Brot.
Q. bulgurluensis A. Camus	Q. faginea Lam.
	Q. infectoria Olivier
	Q. boissieri Reuter
SUBSECT. SUBER	Q. araxina Gross
Q. suber L.	Q. woronowii Maleev
Q. occidentalis G. Gay	
	SUBSECT. HARTWISSIANAE
SUBSECT. MACROLEPIDES	Q. hartwissiana Steven
Q. libani Olivier	
Q. regia Lindley	SUBSECT. SESSILIFLORAE
Q. vesca Kotschy	Q. sessilis Ehrh.
Q. persica Jaubert	Q. mas Thore
Q. ehrenbergii Kotschy	Q. dshorochensis C. Koch
Q. aegilops L.	Q. pinnatiloba C. Koch
Q. macrolepis Kotschy	Q. komarovii A. Camus
Q. vallonea Kotschy	Q. iberica Steven
Q. ithaburensis Decaisne	Q. hypochrysa Steven
Q. brantii Lindley	Q. mannifera Lindley
Q. oophora Kotschy	Q. atropatena Schwarz
Q. pyrami Kotschy	Q. cedrorum Kotschy
Q. ungeri Kotschy	Q. lanuginosa Lam.
Q. look Kotschy	Q. palensis Palassou
Q. castaneifolia Meyer	Q. brachyphylla Kotschy
Q. aitchisoniana A. Camus	Q. dalechampii Ten.
Q. afares Pomel	Q. crispata Steven
Q. trojana Jaub. & Spach.	Q. kotschiana Schwarz
	Q. kozlowsky Woronow
SUBSECT. EUCERRIS	Q. esculiformis Schwarz
Q. cerris L.	
	SUBSECT. PEDUNCULATAE
SECT. MESOBALANUS A. CAMUS	Q. robur L.
	Q. estremadurensis A. Camus
SUBSECT. PONTICAE	Q. fastigiata Lam.
Q. pontica C. Koch	Q. brutia Ten.
	Q. imeretina Stev. & Woronow
SUBSECT. MACRANTHERAE	Q. pedunculiflora K. Koch.
Q. macranthera Fisch. & Meyer	Q. rhodopea Vel.
Q. syspirensis C. Koch	Q. longipes Steven
Q. frainetto Ten.	Q. erucaefolia Steven
Q. toza Gill. & Bosc	Q. haas Kotschy
	Q. sicula Borzi

Tabella 2: -Le querce europee e mediterranee secondo Camus (1936-1939)
(Bussotti e Grossoni, 1997)

Krüssman (1984-1986) propone invece una classificazione basata su tre sottogeneri: *Cyclobalanus* (con lo stesso valore tassonomico e corologico proposto dalla Camus), *Erythrobalanus* Oerst, (querce esclusivamente americane con stili lineari e allungati, ghiande maturanti in due anni con tegumento internamente tomentoso) e *Lepidobalanus* (querce in prevalenza euroasiatiche, alcune nordafricane, con stili subulati e acuti, ghiande maturanti in uno o due anni con tegumento mai internamente tomentoso). Quest'ultimo sottogenere è ripartito nelle seguenti sezioni: *Robur*, *Gallifera*, *Cerris*, *Suber*, *Ilex* (Tabella 3).

Sect. Robur	Sect. Gallifera	Sect. Cerris	Sect. Suber	Sect. Ilex
Q. canariensis Willd.	Q. faginea Ten.	Q. aegilops Lam.	Q. alnifolia Poech	Q. x audleyensis
Q. congesta Borzi	Q. fruticosa Brot.	Q. afares Pomel	Q. calliprinos Webb	Henry
Q. dalechampii Ten.	Q. infectoria Olivier	Q. castaneifolia	Q. coccifera L.	Q. gramuntia L.
Q. frainetto Ten.		C.A. Mey.	Q. x hispanica Lam.	Q. ilex L.
Q. haas Kotschy		Q. cerris L.	Q. suber L.	
Q. hartwissiana Steven		Q. ehrenbergii		
Q. x hickelii		Kotschy		
A. Camus		Q. x libanerris		
Q. iberica M. Bieb.		Boom.		
Q. macranthera		Q. libani Olivier		
Fisch. & C.A. Mey.		Q. macrolepis		
Q. mas Thore		Kotschy		
Q. petraea Liebl.		Q. pyrami		
Q. pedunculiflora		Kotschy		
C. Koch.		Q. trojana Webb		
Q. polycarpa Schur		Q. vallonea Kotschy		
Q. pontica C. Koch.				
Q. pubescens Willd.				
Q. pyrenaica Willd.				
Q. robur L.				
Q. x rosacea Bechst.				
Q. x turneri Willd.				
Q. virgiliana Ten.				

Tabella 3: -Le querce europee e mediterranee secondo Krüssmann (1986)- (Bussotti e Grossoni, 1997)

Nixon (1993) si rifà alla proposta della Camus (1936-38) dandone però una interpretazione strettamente filogenetica: anch'egli suddivide il genere in due sottogeneri, *Cyclobalanopsis* e *Quercus* (= *Euquercus*), ma si differenzia dalla Camus per il diverso modo di aggregare le specie in sezioni. Le sezioni del subg. *Quercus* sono: *Lobatae* (gruppo delle querce rosse Nordamericane), *Protobalanus* (querce del settore occidentale del continente americano, dall'Oregon al Messico, ritenute ancestrali) e *Quercus* (specie del Nordamerica e dell'Eurasia che racchiude anche i “gruppi *Cerris* e *Ilex*”.

Esclusivamente per le querce europee, particolarmente seguita è la classificazione di Schwarz (Tab.4) riportata in *Flora Europea* (1993). Per questo autore esse appartengono a 3 sottogeneri ai quali sono eventualmente da affiancare alcune querce del subgen. *Erythrobalanus* che sono più o meno diffusamente coltivate anche in Europa:

- Cerris* (Spach) Oersted: foglie persistenti, semipersistenti o decidue, ghianda maturante di norma nel secondo anno con endocarpo tomentoso (proprio quest'ultimo sembra essere il carattere morfologico distintivo di queste querce dalle “*querce bianche*” (Manos et al. 1999)).

- *Quercus* Örsted: (Subgen. *Quercus sensu* Schwartz, Sect. *Quercus sensu* Nixon, Sect. *Mesobalanus* e *Lepidobalanus sensu* Camus, Sect. *Robur* e *Gallifera sensu* Krüssmann: “*Querce bianche*”), gruppo polifiletico di specie distribuite in Nordamerica ed Eurasia, a foglie decidue o semipersistenti, con caratteristica posizione basale degli ovuli abortivi sulla superficie del frutto, quest' ultimo maturante nell'anno e con endocarpo glabro; secondo Nixon, il gruppo racchiude anche i “gruppi *Cerris* e *Ilex*”, secondo Manos et al. (1999), i due gruppi vanno separati dalle querce bianche e dovrebbero assumere rango di Sezioni; secondo Camus, il gruppo raccoglie anche *Q. ilex*;

- Sclerophyllodrys*: O. Schwarz: gruppo monofiletico di specie esclusivamente euroasiatiche con foglie persistenti a lamina spessa, ghianda maturante in uno o due anni, endocarpo tomentoso. Il sottogenere *Sclerophyllodris corrisponde a: Ilex group sensu* Nixon, Sect. *Ilex* e *Suber sensu* Krüssmann (includendo, perciò, *Q.suber*), Sect. *Cerris* (Subsect. *Cocciferae*) e Sect. *Lepidobalanus* (Subsect. *Ilex*) *sensu* Camus (raggruppando *Q.ilex* con le querce bianche e le altre sclerofille con il gruppo *Cerris*).

- Erythrobalanus* (Spach) Örsted (America centro-settentrionale): ghianda maturante nel secondo anno e endocarpo tomentoso. Comprende la cosiddette “*querce rosse*” e “*querce nere*” americane.

Subgen. Quercus	Subgen. Cerris	Subgen. Sclerophyllodrys
Q. mas Thore	Q. suber L.	Q. coccifera L.
Q. polycarpa Schur	Q. trojana Webb	Q. ilex L.
Q. petraea Liebl.	Q. macrolepis Kotschy	
Q. dalechampii Ten.	Q. cerris L.	
Q. hartwissiana Steven		
Q. robur L.		
Q. pedunculiflora C. Koch		
Q. frainetto Ten.		
Q. pyrenaica Willd.		
Q. pubescens Willd.		
Q. canariensis Willd.		
Q. faginea Lam.		
Q. lusitanica Lam.		

Tabella 4: -Le querce europee secondo Schwarz (1993)- (Bussotti e Grossoni, 1997)

Il sistema di classificazione più recente è attribuibile a Yu.L. Menitsky (2005). Egli riprende quello del 1938 di V.P. Maleev modificato da S.Ya. Sokolov nel 1951 in cui le querce euro-asiatiche sono divise nei sottogeneri *Lepidobalanus*, *Protobalanus* (con le sezioni *Ilex* e *Coccifera*) e *Cyclobalanopsis* senza suddivisioni (Tab.6). Il sottogenere *Lepidobalanus* include: la sezione *Cerris* che contiene, come nel sistema di Schwarz, non solo le querce decidue ma anche quelle sempreverdi del gruppo *Q. suber*, la sezione *Cerridopsis* (con specie del gruppo *Q. macranthera*) isolata più di recente in una sezione *Dentata* (con *Q. dentata*) e la sezione *Eulepidobalanus* con le sottosezioni *Robur*, *Ponticae*, *Galliferae* e *Diversipilosae* (con corteccia liscia).

Menitsky (2005) ricalca il sistema di classificazione precedentemente descritto ad eccezione del trasferimento delle specie sempreverdi (affinità con il gruppo della *Q. suber* e *Q. semecarpifolia*), dalla sezione *Cerris* delle querce decidue ad un altro sottogenere con le querce a foglia coriacea, sulla base della somiglianza del tipo di struttura della lamina fogliare e di peculiarità anatomiche del legno, abbassando il rango delle rimanenti sottodivisioni del sottogenere a sottosezioni basate sul riconoscimento delle origini comuni e della vicinanza

filogenetica dei gruppi affini *Q. robur*, *Q. macranthera*, *Q. pontica*, *Q. dentata* aventi un antenato comune Est-Asiatico più prossimo a *Q. griffithi*. Nel sottogenere *Protobalanus*, chiamato in seguito *Heterobalanus* come per Örsted, ha incluso le querce suberoidi (*Q. semecarpifolia*, *Q. suber* e altre).

Subgen ∴	Quercus L. Gerontogaeae		Heterobalanus Oerst.	
Sect.:	Quercus	Cerris Dumort	Heterobalanus (Oerst.) Menits	Ilex (Loud) Guerke
Subject ∴	- <i>Pontica</i> (Stef.) A. Camus	- <i>Cerris</i> (Dumort.) Guerke	- <i>Heterobalanus</i> (Oerst.) O. Schwarz	- <i>Ilex</i> (Loud)
	- <i>Quercus</i>	- <i>Aegilops</i> (Reichenb.) Menits	- <i>Subers</i> (Spach))Maleev	- <i>Cocciferae</i> (Spach) Guerke
	- <i>Macrantherae</i> (Stef.) A. Camus			
	- <i>Galliferae</i> (Spach) Guerke			

Tab.6: –Schema di classificazione per le querce europee proposto da Menitsky (2005)–

La situazione tassonomica delle querce italiane

Confrontando nei principali testi la distribuzione e la classificazione del genere *Quercus* nell'Europa e nella regione mediterranea, viene messa in evidenza una notevole incertezza tassonomica: quelle che per alcuni autori sono specie valide, per altri sono solo sottospecie e varietà.

Ad esempio, dal censimento di Tutin per *Flora Europaea* 1964-1980 (Tab.7) risultano ben 27 specie: tra queste le Nord Americane, ormai naturalizzate, *Q. rubra* L.e *Q. palustris* Muenchh. appartenenti al Subgen. *Erythrobalanus* (Spach) Örsted; *Q. coccifera* L., *Q. ilex* L., *Q. rotundifolia* Lam. del Subgen. *Sclerophyllodrys* O. Schwarz.; *Q. suber* L., *Q. trojana* Webb in

Loudon, *Q. macrolepis* Kotschy, *Q. cerris* L. del Subgen. *Cerris* (Spach) Örsted; *Q. mas* Thore, *Q. polycarpa* Schur., *Q. petraea* (Mattuschka) Liebl., *Q. dalechampii* Ten., *Q. hartwissiana* Steven, *Q. robur* L., *Q. pedunculiflora* C. Koch, *Q. sicula* Borzì in Lojac, *Q. frainetto* Ten., *Q. pyrenaica* Willd, *Q. congesta* C. Presl in J. & C. Presl. , *Q. brachyphylla* Kotschy, *Q. virgiliana* (Ten.) Ten., *Q. pubescens* Willd., *Q. canariensis* Willd., *Q. faginea* Lam., *Q. fruticosa* Brot., *Q. infectoria* Olivier del subgen. *Quercus*.

Subgen. <i>Quercus</i>	Subgen. <i>Cerris</i>	Subgen. <i>Sclerophylloides</i>
<i>Q. mas</i> Thore	<i>Q. suber</i> L.	<i>Q. coccifera</i> L.
<i>Q. polycarpa</i> Schur	<i>Q. trojana</i> Webb	<i>Q. ilex</i> L.
<i>Q. petraea</i> (Matt.) Liebl.	<i>Q. macrolepis</i> Kotschy	<i>Q. rotundifolia</i> Lam.
<i>Q. dalechampii</i> Ten.	<i>Q. cerris</i> L.	
<i>Q. hartwissiana</i> Steven		
<i>Q. robur</i> L.		
<i>Q. pedunculiflora</i> C. Koch		
<i>Q. sicula</i> Borzì		
<i>Q. frainetto</i> Ten.		
<i>Q. pyrenaica</i> Willd.		
<i>Q. congesta</i> C. Presl		
<i>Q. brachyphylla</i> Kotschy		
<i>Q. virgiliana</i> Ten.		
<i>Q. pubescens</i> Willd.		
<i>Q. canariensis</i> Willd.		
<i>Q. faginea</i> Lam.		
<i>Q. fruticosa</i> Brot.		
<i>Q. infectoria</i> Olivier		
<i>Q. lusitanica</i> Lam.		

Tabella 7: -Classificazione delle querce europee e mediterranee secondo Schwarz riportate in Flora Europea di Tutin 1964-1980- (Bussotti e Grossoni, 1997)

Nella seconda edizione (*Flora Europaea*1993) sono riportate 20 diverse specie quercine, differentemente dalla Camus che, pochi anni prima, ne aveva censite una settantina. Secondo *Med-Checklist* (Greuter et al., 1986) esse sono 17 (oltre a 5 dubbie) mentre nella totalità del bacino mediterraneo le entità salgono a 25 (oltre a 6 per le quali vengono posti seri dubbi sulla validità del rango di specie).

Limitando questo inventario al territorio italiano, per *Flora Europaea* (1993) le specie quercine sono 12: *Q. coccifera*, *Q. ilex*, *Q. cerris*, *Q. macrolepis*, *Q. suber*, *Q. trojana*, *Q. congesta*, *Q. dalechampii*, *Q. petraea*, *Q. pubescens*, *Q. pyrenaica*, *Q. robur* e considera specie dubbia *Q. crenata*.

Secondo Pignatti (1982) la flora quercina è costituita da 15 specie indigene:

Q. robur, *Q. petraea*, *Q. pubescens*, *Q. frainetto*, *Q. cerris*, *Q. trojana*, *Q. macrolepis*, *Q. suber*, *Q. ilex*, *Q. coccifera*, *Q. calliprinos*, *Q. dalechampii*, *Q. virgiliana* e *Q. congesta*, *Q. pyrenaica*.

R. Gellini in *Botanica Forestale* (1997), aggiunge a questo elenco anche *Q. crenata* e la Nordamericana *Q. rubra*.

Di seguito sono riportate delle brevi descrizioni delle specie quercine maggiormente riconosciute dai diversi autori e le relative considerazioni tassonomiche (Tab.8).

1. *Q. aegilops* L. (nome ambiguo, aff. *Q. ithaburensis* Decaisne).
2. *Q. aegilops* Lam. (sin. *Q. macrolepis*, aff. *Q. ithaburensis* Decaisne)
3. *Q. afares* Pomel. (aff. *Q. castaneifolia*)
4. ***Q. alnifolia* Poech.**
5. *Q. aucheri* Jaub. & Spach. (aff. *Q. coccifera*)
6. *Q. boissieri* Reut. (aff. *Q. infectoria*)
7. *Q. brachyphylla* Kotschy (aff. *Q. pubescens*)
8. *Q. brantii* Lind. (aff. *Q. ithaburensis* Decaisne)
9. *Q. calliprinos* Webb (aff. *Q. coccifera*)
10. ***Q. canariensis* Willd.**
11. ***Q. castaneifolia* C.A. Mey**
12. *Q. cerrioides* Willk & Costa (aff. *Q. pubescens*; prob. ibrido *Q. pubescens* x *Q. faginea*)
13. ***Q. cerris* L.**
14. ***Q. coccifera* L.**
15. *Q. conferta* Kit. (sin. *Q. frainetto*)
16. *Q. congesta* C. Presl (aff. *Q. pubescens*)
17. *Q. crenata* Lam. (aff. *Q. cerris*; prob. ibrido *Q. cerris* x *Q. suber*)
18. *Q. dalechampii* Ten. (aff. *Q. pubescens* - *Q. petraea*)
19. *Q. dshorochensis* C. Koch (sin. *Q. polycarpa*; e *Q. iberica* secondo Greuter *et al.*; aff. *Q. petraea*)
20. *Q. dshorochensis* Wenz (sin. *Q. iberica* secondo Krüssmann; aff. *Q. petraea*)
21. *Q. ehrenbergii* Kotschy (sin. *Q. macrolepis*; Aff. *Q. ithaburensis* Decaisne)
22. ***Q. faginea* Lam.**
23. *Q. fontanesii* Guss. (aff. *Q. cerris* secondo Pignatti, 1982; sin. *Q. crenata* secondo Greuter, 1986)
24. ***Q. frainetto* Ten.**
25. *Q. fruticosa* Brot. (sin. *Q. lusitanica* Lam.; aff. *Q. faginea*)
26. *Q. gramuntia* L. (aff. *Q. ilex*)
27. *Q. haas* Kotschy (aff. *Q. robur*)
28. *Q. hartwissiana* Steven (aff. *Q. robur* secondo Camus; aff. *Q. pubescens* secondo Pardé)
29. *Q. humilis* Lam. (sin. *Q. lusitanica* Lam.; aff. *Q. faginea*)
30. *Q. humilis* Miller (sin. *Q. pubescens*)
31. *Q. iberica* M. Bieb. (sin. *Q. polycarpa* secondo Greuter *et al.*; aff. *Q. petraea*)
32. ***Q. ilex* L.**
33. *Q. infectoria* Olivier (aff. *Q. faginea* secondo Krüssmann, 1986; aff. *Q. pubescens* secondo Schwarz, 1993)
34. ***Q. ithaburensis* Decaisne**
35. *Q. lanuginosa* Thuill. (sin. *Q. pubescens*)
36. ***Q. libani* Oliv.**
37. *Q. lusitanica* Lam. (aff. *Q. faginea*)
38. *Q. lusitanica* Webb (sin. *Q. faginea* e *Q. canariensis*)
39. *Q. macedonica* A. DC. (sin. *Q. trojana*)
40. ***Q. macranthera* Fisch. & Mey**
41. *Q. macrolepis* Kotschy (sin. *Q. aegilops* Lam., aff. *Q. ithaburensis* Decaisne)
42. *Q. mas* Thore (aff. *Q. petraea*)
43. *Q. mirbeckii* Dur. (sin. *Q. canariensis*)
44. *Q. occidentalis* Gay (aff. *Q. suber*)
45. *Q. pedunculata* Ehrh. (sin. *Q. robur*)
46. *Q. pedunculiflora* C. Koch (aff. *Q. robur*)
47. ***Q. petraea* (Matt.) Liebl.**
48. *Q. polycarpa* Schur (aff. *Q. pubescens* secondo Camus; aff. *Q. petraea* secondo Krüssmann)
49. ***Q. pontica* C. Koch**
50. ***Q. pubescens* Willd.**
51. *Q. pyrami* Kotschy (aff. *Q. ithaburensis* Decaisne)
52. ***Q. pyrenaica* Willd.**
53. ***Q. robur* L.**
54. *Q. rotundifolia* Lam. (aff. *Q. ilex*)
55. *Q. sessiliflora* Salisb. (sin. *Q. petraea*)
56. *Q. sicala* Borzi (aff. *Q. pubescens*)
57. ***Q. suber* L.**
58. *Q. toza* Bosc (sin. *Q. pyrenaica*)
59. ***Q. trojana* Webb**
60. *Q. virgiliana* Ten. (aff. *Q. pubescens*)
61. *Q. valentina* Cav. (sin. *Q. lusitanica* Lam.; aff. *Q. faginea*)
62. *Q. vallonea* Kotschy (aff. *Q. ithaburensis* Decaisne)
63. *Q. vulcanica* Kotschy (aff. *Q. petraea*)

Tab.8: -Elenco dei sinonimi specifici più comunemente usati (aff.=specie affine a; sin.=sinonimo di). In grassetto sono indicate le specie principali- (Bussotti e Grossoni, 1997)

Subgenere *Quercus*¹

Specie quali *Q. robur* L., *Q. petraea* (Matt.) Liebl, *Q. pubescens* Willd, *Q. frainetto* Ten. , risultano essere, per le classificazioni di seguito citate, tutte appartenenti ad un medesimo sottogenere e/o ad una stessa sezione.

Per Schwarz sono incluse nel Subgen. *Quercus* corrispondente all'insieme delle due sezioni *Robur* e *Gallifera* di Krüsmann (secondo il quale esse rientrano soltanto nella sezione *Robur*).

Dal punto di vista di Menitsky (2005) sono da includere nel sottogenere e nella sezione *Quercus*. In particolare: *Q. robur* è inserita nella sottosezione *Quercus* (membri che non hanno apici papilliformi sui denti delle foglie); *Q. frainetto* nella sottosezione *Macrantherae* (con densa pubescenza nella pagina inferiore della foglia, stipole lunghe e persistenti, e squame della cupola non strettamente appressate); *Q. petraea* e *Q. pubescens* nella sottosezione *Galliferae* (base della foglia cuneata, microfillia, sviluppata pubescenza, e talvolta schizofillia)

-*Q.robur* L.

Albero di grandi dimensioni a foglie caduche la cui area di distribuzione è molto vasta e comprendente tutta l'Europa.

In Italia è presente in tutte le regioni (con l'eccezione della Sardegna), soprattutto nelle pianure alluvionali o lungo valli e vallette umide; la maggior parte di questi ambienti è stata modificata dall'uomo e attualmente la farnia è in forte contrazione e si ritrova in maniera puntiforme lungo la penisola (Gellini e Grossoni, 1997).

Secondo Krüssman (1986), Schwarz (1964;1963) e Pignatti (1982) questo taxon comprende due sottospecie:

- subsp. *robur*, che copre l'intera area di distribuzione di questa specie.
- subsp. *brutia*, dell'Italia meridionale e costa adriatica dei Balcani.

-*Q.petraea* (Matt.) Liebl.

Grande albero a foglie caduche distribuito in tutta Europa e nell'Asia Minore. In Italia è specie di collina e di bassa montagna. Nell'Italia centrale, sono stati segnalati popolamenti puri o misti di rovere: nei boschi di latifoglie della Toscana settentrionale (Lunigiana, Garfagnana, Alpi Apuane), nel bosco di Tatti nel Volterrano, a Sargiano nei pressi di Arezzo, nel Senese, sul Monte Corona presso Umbertide (Umbria), sulle colline del Trasimeno, sul Monte Gemmo

¹

Il sistema di classificazione in sottogeneri, per le specie trattate da questo punto in poi, ricalca quello di Schwarz per Flora Europea 1993

(Camerino), sui Monti della Tolfa e sui Monti Cimini nel Viterbese, nel Teramano; in Sicilia sulle Madonie vi è un popolamento molto interessante non solo per la sua estensione ma anche per la sua dislocazione periferica (Gellini e Grossoni, 1997).

Schwarz in Flora Europaea non suddivide la Rovere in sottospecie; tuttavia a *Q. petraea* vengono attribuite almeno 3 sottospecie (Greuter et al. 1986): subsp. *iberica* (Europa orientale e Asia Minore), subsp. *petraea* (in tutto l'areale), subsp. *pinnatiloba* (diffusa nell'Asia Minore). Castroviejo *et al.* (1990) considera, oltre alla subsp. *petraea*, anche la subsp. *huguetiana* Franco & G. López.

-*Q. pubescens* Willd.

Albero deciduo di medie dimensioni. È distribuito in tutta l'Europa meridionale e mediterranea e nell'Asia Minore. In Italia è presente in tutte le regioni, principalmente nella sottozona calda del *Castanetum* ma anche nella sottozona fredda del *Lauretum*, in ambienti molto diversificati grazie alla sua frugalità ed al suo polimorfismo ecologico (Gellini e Grossoni, 1997).

Q. pubescens è una specie estremamente polimorfa, ed è stata suddivisa in numerose sottospecie e varietà. La grande variabilità morfologica della roverella e la sua sovrapposizione morfologica ed ecologica con la rovere hanno creato notevoli difficoltà anche da un punto di vista tassonomico portando da una parte ad una estrema semplificazione (Fiori nella sua Flora d'Italia 1923-25) riportava solo

Q. robur in cui sono comprese farnia, rovere e roverella) e dall'altra alla frammentazione del taxon in numerose piccole specie a distribuzione molto localizzata (ad esempio: *Q. brachyphylla* Kotschy dell'Egeo, *Q. infectoria* dell'Egeo e dell'Anatolia, *Q. sicula* Borzi e *Q. virgiliana* Ten. Le numerosissime entità, in passato descritte, non vengono più riconosciute.

Schwarz (1964-1963) distingue tre sottospecie: subsp. *anatolica* (nella parte orientale della penisola balcanica e della Crimea), subsp. *palensis* (dei Pirenei del Nord-Est e della Spagna), subsp. *pubescens* (in tutto l'areale mediterraneo tranne Spagna e Pirenei).

-*Q. frainetto* Ten.

Grande albero caducifoglio ad areale europeo sud-orientale, diffuso in Anatolia settentrionale, Bosnia, Bulgaria, Albania, Macedonia e Grecia. Presente nell'Italia centro-orientale ma più frequente al Meridione. La denominazione frainetto deriva da un errore tipografico: il nome originario era infatti *Q. farnetto* Ten.

Il Subgenere Cerris

Seguendo la classificazione di Schwarz (1964,1993) *Q. cerris*, *Q. trojana*, *Q. macrolepis*, *Q. suber* sono riunite nello stesso subgen. *Cerris*. Anche Krüssmann avvicina le quattro specie ma considera come specie differenti *Q. macrolepis* e

Q. aegilops. Menitsky (2005) colloca *Q. cerris* nel sottogenere *Quercus*, nella sezione e nella sottosezione *Cerris*, mentre *Q. aegilops* e *Q. trojana* (con foglie completamente glabre) nell'altra sottosezione: *Aegilops*. Questa si differenzia per avere foglie molto dentate e apici più lunghi sulla sommità dei denti. *Q. suber* ha una collocazione particolare. Come nella vecchia classificazione di Fiori (1923-1929), dove è lontana dal cerro e più vicino al leccio (*Sez. suber* con: *Q. suber*, *Q. ilex*, *Q. coccifera*), Menitsky (2005) non la include nel sottogenere *Quercus* ma nel sottogenere e nella sezione *Heterobalanus* e nella sottosezione *Suber*.

-*Q.cerris* L.

Albero deciduo diffuso nell'Europa centro-meridionale orientale. In Italia è frequente sulla dorsale appenninica, Maremma toscana, Molise, Sannio, Irpinia, Basilicata. Raro in Sicilia e in Pianura Padana, dove forse è stato introdotto. Assente in Sardegna.

È una delle specie quercine tassonomicamente meno controverse di cui, tuttavia, sono state descritte alcune varietà tra cui:

-var. *tournefortii* O. Swarz: presente in Asia minore e Sicilia, con foglie giallo-verdastre, tomentose inferiormente, di forma variabile anche sulla stessa pianta ma sempre pinnatifide, spesso con la lamina praticamente limitata alla parte apicale e a quella basale dal momento che la porzione centrale è costituita da due seni ampi e profondi fino alla nervatura principale.

-var. *pseudocerris* Boiss: tipica dell'Europa orientale e dell'Asia occidentale. Piuttosto polimorfa con foglie profondamente incise e lobi interi o leggermente sub lobati.

-var. *austriaca* Loud.: ha foglie ovate, con margine dentato ma con lobi molto ridotti o anche assenti; malgrado la specificazione del nome si ritrova in tutto l'areale.

-var. *haliphloeos* Lam. & DC.: è presente in Francia e Italia, ha foglie molto grandi e molto allargate nella metà apicale.

In Italia, non solo non è infrequente che cerri riferibili alla var. *haliphloeos* si ritrovino nei medesimi soprassuoli, ma spesso anche lo stesso albero può portare foglie ascrivibili ai due taxa precedenti, (specialmente la seconda cacciata ha sempre foglie del tipo “*austriaca*”; Schirone et al., 1988)

-*Q. trojana* Webb.

Piccolo albero a foglie semipersistenti o tardivamente caduche, diffuso nella parte occidentale della penisola balcanica ed in Puglia (Italia). I popolamenti italiani sono rimasti ignorati fino al 1885 quando la specie fu descritta col nome di *Q. fragnus* da Longo; successivamente Borzì la identificò come *Q. macedonica* A. DC. E quindi Martelli come *Q. trojana*. Numerose sono le specie come *Q. trojana* e la successiva *Q. macrolepis*, che hanno corologia anfiadriatica. Questo è potuto accadere in quanto più volte il Mare Adriatico si è ritirato creando ponti fra le due sponde; in particolare nel Miocene superiore (Pontico) la Puglia è stata ampiamente legata alla regione illirica. Le ultime congiunzioni adriatiche sono avvenute nel Rissiano e nel Würmiano quando il mare si prosciugò fin poco sotto il Conero (Montelucci, 1972).

-*Q. ithaburensis* Decaisne

Albero di piccole dimensioni con distribuzione mediterraneo-orientale che va dalla Puglia ai Balcani, alla Grecia ed all'Asia Minore. Questa quercia è stata oggetto di studi da parte di vari botanici che hanno operato autonomamente con conseguenti denominazioni e attribuzioni diversificate, tutto ciò ha ovviamente generato un quadro tassonomico piuttosto confuso. Nel passato veniva riconosciuta come *Q. aegilops* L., questo nome oggi è considerato confuso e ne è stato proposto l'abbandono da Greuter (1986). Krüssmann nel 1986 le considerava specie separate.

-*Q. crenata* Lam.

Presunto ibrido naturale, probabilmente di origine antica e stabilizzato (Fiori, 1930; Camus, 1936-1939; Schwarz 1964, 1993; Pignatti 1982) dato da *Q. cerris* x *Q. suber*. È un albero a foglie semipersistenti ed è diffuso, più o meno sporadicamente, in gran parte d'Italia e lungo la costa dalmata.

-*Q. suber* L.

Albero di 10-20 m a foglie sempreverdi. È presente nel settore occidentale del bacino del mediterraneo, dalle coste occidentali dell'Italia fino all'Oceano Atlantico. È diffuso soprattutto in Portogallo, Spagna e in Sardegna. *Quercus suber* è un taxon che non viene ulteriormente suddiviso; tuttavia viene spesso riconosciuta una varietà *occidentalis*: *Q. s. var. occidentalis*

(Gay) Arcang. (riportata anche come sottospecie) che comprende individui più rustici, con maturazione biennale del frutto, foglie semipersistenti e legno semiporoso.

Secondo Menitsky (2005), la differenziazione dei gruppi extra-tropicali del genere *Quercus*, come mostrano le tendenze della sua evoluzione morfologica, fu causata dall'adattamento di alcune forme ancestrali del genere *Lithocarpus* del clima equatoriale alle condizioni del clima monsonico con periodi invernali freddi e secchi e siccità durante il periodo di fioritura che comportò il passaggio all'anemofilia.

Nei due sottogeneri: *Quercus* e *Heterobalanus*, l'adattamento ai climi secchi proseguì in due modi differenti: nel primo con la conservazione della struttura mesofitica della foglia attraverso lo sviluppo della caduta regolare delle foglie in autunno, e nel secondo sviluppando l'insieme di caratteristiche di sclerofillia e progressiva microfillia della foglia, dando la possibilità agli alberi di superare il periodo di aridità fisiologica attraverso la riduzione della traspirazione nonostante la conservazione delle foglie verdi in inverno. Nel sottogenere *Heterobalanus* sono incluse due sezioni: la sezione *Heterobalanus* e la sezione *Ilex*. la prima è a sua volta suddivisa nelle due sezioni *Heterobalanus* e *Suber* quest'ultima particolarmente interessante perchè ospitante solamente la *Quercus suber*. Secondo Menitsky (2005) la sughera dà l'impressione di essere un ibrido intersezionale tra alcune specie iniziali della sottosezione *Heterobalanus* e specie del gruppo della *Q. aegilops* della sezione *Cerris*. Se questa ibridazione è avvenuta, deve tuttavia essere avvenuta in tempi molto antichi, durante il periodo della dispersione di specie della sezione *Heterobalanus* sull'attuale territorio Mediterraneo, poichè, ad oggi, non sono presenti querce riferibili al gruppo della *Q. aegilops* nell'areale della *Q. suber* (massiccio dell'atlante, Penisola Pirenaica, Sud della Francia).

Il Subgenere Sclerophyllodrys

Schwarz indica, come appartenenti al genere *Sclerophyllodrys*, sia la *Q. coccifera* che *Q. ilex*. Quest'ultima è inserita nel subgen. *Heterobalanus*, sezione e sottosezione *Ilex* (Menitsky 2005) e viene descritta come specie con foglie piccole, dure, ovate o oblunghie, dentate, usualmente con una pubescenza biancastra sulla pagina inferiore distribuita nei boschi umidi subtropicali delle regioni dell'Emisfero Settentrionale.

-*Q.ilex* L.

Albero sempreverde che può raggiungere i 25 m di altezza. È presente in tutta la regione mediterranea ma lungo le coste balcaniche si presenta in stazioni disgiunte che diventano sporadiche e miste nel versante orientale. In Italia è diffuso principalmente nelle isole e nelle regioni costiere tirreniche e ioniche; sul versante adriatico i popolamenti diventano discontinui e sporadici anche se si possono ancora trovare formazioni estese in Puglia, Abruzzo e Marche; piccoli nuclei isolati e relittuali sono in Val Padana lungo le coste dei grandi laghi insubrici, sui colli Euganei, in Friuli e nel Ferrarese (Bosco della Mesola). Lungo la penisola risale spesso le valli interne disponendosi sui versanti calcarei soleggiati (Gellini e Grossoni, 1997).

Il leccio è in grado di ibridarsi con la sughera dando luogo alla *Q. morisii* Borzì: questo albero sempreverde è diffuso nella regione del Mediterraneo occidentale e nel caso italiano in Toscana, Lazio, Puglia e Sardegna. Si differenzia dalla sughera per la corteccia poco sviluppata e per le foglie con un colore verde intenso.

Anche se alcuni botanici e *Flora Iberica* (Castroviejo *et al.*, 1990) continuano a separare *Q. ilex* da *Q. rotundifolia*, attualmente si tende a considerare il leccio con le sue due sottospecie:

- subsp. *ilex*, distribuita nel Mediterraneo centrale e orientale;

- subsp. *rotundifolia*, distribuita nel Portogallo e nella maggior parte della Spagna.

-*Q. coccifera* L.

Inserita nel subgen. *Heterobalanus*, sez. *Ilex*, sottosez. *Cocciferae*.

È un arbusto o alberello a foglie sempreverdi diffuso nelle zone più aride del bacino del Mediterraneo. *Q. coccifera* è costituita da due entità tassonomiche (subsp. *coccifera* e subsp. *calliprinos*) che a loro volta hanno distribuzione nel settore centro occidentale (Libia, Tunisia, Algeria, Marocco, Portogallo, Spagna) e nel settore centro-orientale (Sicilia, Basilicata, Puglia, coste adriatiche, Crimea, Anatolia). Come per molte altre specie quercine, rimane aperta la questione sull'inquadramento intraspecifico. Mentre *Flora Europaea* (1993) e Greuter *et al.* (1986) riconoscono la sola *Q. coccifera*, senza ulteriori suddivisioni, Pignatti, in *Flora d'Italia* (1982) mantiene la divisione della quercia spinosa in due specie distinte: *Q. coccifera* L. e *Q. calliprinos* Webb. Quest'ultima viene descritta come un albero sempreverde diffuso nelle regioni del mediterraneo orientale. Costituisce l'elemento fondamentale delle formazioni di macchia o di foresta sempreverde climax del Mediterraneo orientale dove la sua presenza si è ormai rarefatta per cause antropiche.

Le querce di dubbia collocazione

Tra le querce italiane di dubbia collocazione tassonomica rientrano: *Q. dalechampii*, *Q. virgiliana*, *Q. congesta* e *Q. pyrenaica*; esse appartengono, secondo Schwarz (1964), al

sottogenere *Quercus*. Nella classificazione del 1993 non sono presenti *Q. virgiliana*, *Q. congesta* né l'endemica *Q. sicula* Borzi in Lojac; perchè ricomprese nell'ambito di *Quercus pubescens*. Al contrario, Krussmann, riprendendo il primo sistema di Schwarz (1964), mantiene separate le varie specie nell'ambito della stessa sezione *Robur*.

-*Q. dalechampii* Ten.

Albero di piccole dimensioni, deciduo e diffuso nell'Europa Sud-Orientale, Dalla Slovacchia all'Italia occidentale. Sia *Flora d'Italia* (1982) sia *Flora Europaea* considerano questo taxon correlato alla serie *petraea-pubescens* e più strettamente connesso a *Q. petraea*. Rispetto a *Q. petraea* si differenzia principalmente per avere la foglia glabra e maggiormente incisa e per la ghianda che appare più rugosa a causa delle squame non strettamente appressate.

-*Q. virgiliana* Ten.

Albero deciduo di grandi dimensioni, con ghiande dolci, diffuso nell'Europa meridionale: dalla Corsica fino al Mar Nero. La si ritrova in ambienti prevalentemente aridi con suoli subacidi insieme alla roverella o come sua vicariante.

-*Q. congesta* C. Presl.

Arbusto o piccolo alberello deciduo, diffuso in Sardegna, Sicilia e nel Sud della Francia. Ha una morfologia strettamente correlata con quella della roverella da cui se ne distingue per i rami penduli e per i lobi sovrapponibili.

-*Q. pyrenaica* Willd.

Detta anche *Q. toza*, è un albero deciduo di piccole dimensioni, diffuso nell'Europa meridionale atlantica e comune sulle colline sud-orientali della Francia, nella Loira, nella Bretagna e sui Pirenei. In Italia è segnalata come rarissima nella Val di Susa in Piemonte.

-*Q. rubra* L.

Quercia del versante orientale del Nord America appartenente al subg. *Erythrobalanus*. È la più importante e coltivata delle querce rosse americane, è stata introdotta e diffusa in Europa nel 1724 per il suo valore decorativo. In Italia sono ben noti gli esempi dei Parchi della Villa Reale di Monza e della Mandria vicino a Torino.

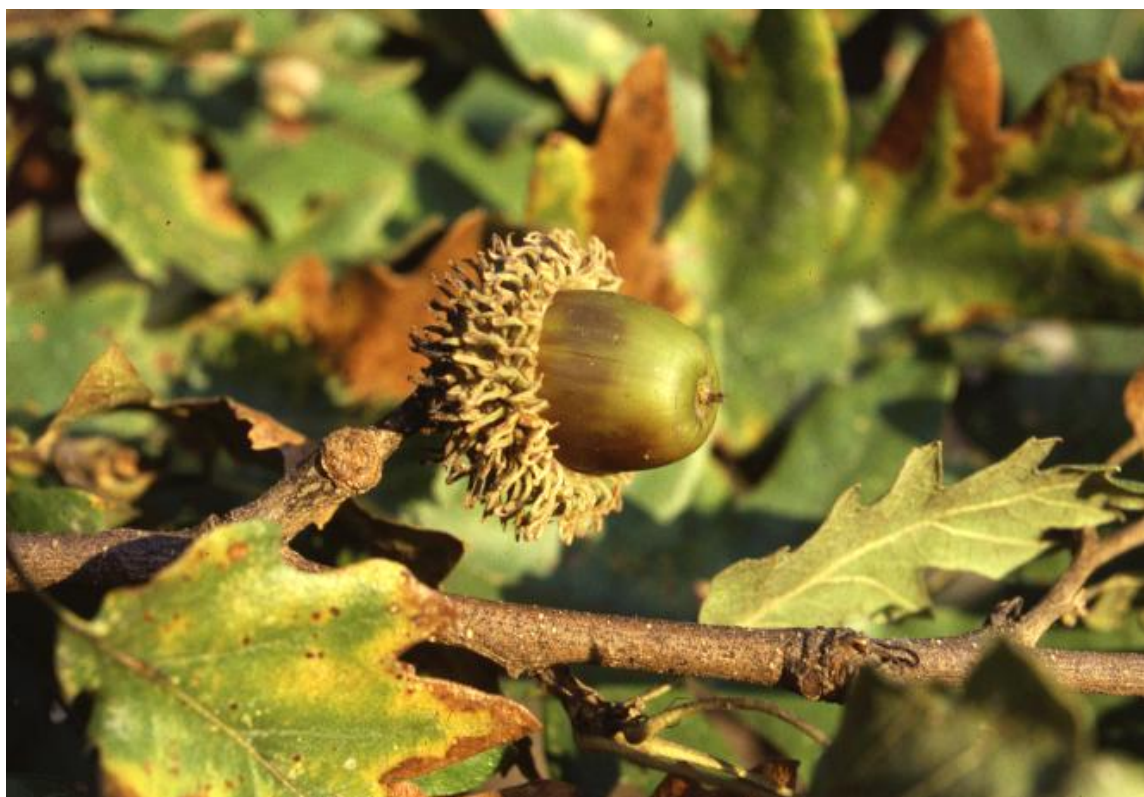


Foto 1: Q.Cerris



Foto 2: Q.crenata



Foto 3: Q.pubescens

3 IL DISEGNO SPERIMENTALE DELLA TESI

Strategia di campionamento

Fase 1: Floristic dataset

Come illustrato nella parte introduttiva, sono possibili diversi approcci di campionamento ognuno dei quali è in grado di fornire informazioni diverse.

Poiché non esistevano precedenti studi su specie forestali dell'area Mediterranea o italiana, l'esplorazione della biodiversità è cominciata con un approccio di campionamento floristico della dendroflora Italiana.

E' stata così fatta una lista delle specie più diffuse e rappresentative del nostro territorio e queste sono state studiate utilizzando un solo individuo per specie.

Le informazioni derivanti da questo tipo di campionamento riguardano principalmente la parte di tecnica molecolare ossia l'universalità dei primers ed eventuali variazioni sui protocolli di amplificazione e sequenziamento. Tuttavia consente di valutare anche problematiche relative ai marcatori (specialmente negli omopolimeri dei loci non codificanti) o, come nel nostro caso, evidenziare generi con percentuali di successo o insuccesso nell'identificazione di specie molto particolari.

Fase 2: Taxon-based dataset

La prima esplorazione della dendroflora italiana ha evidenziato il genere *Quercus* come degno di nota per le sue pessime percentuali di successo, e poiché tale genere rappresenta un punto chiave del nostro patrimonio forestale, nella seconda parte del periodo di dottorato proprio tale genere è stato esplorato con un approccio *taxon-based*. Quindi le dodici specie del genere *Quercus* presenti in Italia sono state campionate con diversi individui provenienti da diverse regioni, consentendo di ottenere notevoli informazioni non solo come risposta all'approccio barcode in senso stretto, ma anche sui rapporti evolutivi che probabilmente legano queste entità.

Marcatori barcode utilizzati

La scelta dei marcatori da utilizzare per i nostri scopi è stata fatta dopo la Conferenza Internazionale di Taipei, quando le informazioni disponibili relative ai marcatori erano ancora molto generiche e, di conseguenza, non è stato facile operare la scelta.

Tuttavia le regioni da noi selezionate sono risultate essere quelle che ora, a distanza di tre anni, il CBOL ha indicato come barcode ufficiali (*rbcL* e *matK*) e *trnH-psbA* il terzo candidato potenziale, rendendo il nostro studio perfettamente inserito nel contesto attuale del dibattito.

Per la prima parte del lavoro (*approccio floristico*) sono stati scelti quattro marcatori:

rbcL e ***rpoC1*** (regioni codificanti conservate) ***matK*** (regione codificante a rapida evoluzione) e ***trnH-psbA*** (regione intergenica a rapida evoluzione).

Sulla base dei risultati ottenuti da questa prima esplorazione il locus *rpoC1* è stato escluso, quindi, per lo studio *taxon-based* sono stati utilizzati ***rbcL***, ***matK*** e ***trnH-psbA***.

Inoltre è stata sperimentata una terza regione, da noi disegnata, localizzata nella regione a monte della regione classica *matK* e indicata come ***matK2***.

4 MATERIALI E METODI

Il Materiale vegetale

Campionamento floristico

Il campione è stato costruito scegliendo specie rappresentative dei diversi ambienti della nostra penisola.

Partendo dalla zona alpina (*Larix decidua*), passando poi per quella montana e submontana (*Fagus sylvatica*, *Pinus nigra*, *Acer spp.*, *Corylus avellana*, *Fraxinus spp.*, *Quercus spp.*, *Castanea sativa*, *Prunus spinosa*, *Craetagus monogyna*), quella umida (*Salix alba*, *S. triandra*, *Populus alba*, *P. nigra*, *Fraxinus angustifolia*, *Alnus glutinosa*, *Ulmus minor*), fino alla macchia Mediterranea e foreste sempreverdi (*Quercus spp.*, *Olea europaea*, *Phyllirea spp.*, *Pinus pinea*, *P. halepensis*, *Cupressus sempervirens*, *Juniperus phoenicea*, *Ligustrum vulgare*, *Fontanesia phyllireoides*), e includendo alcune piante rampicanti e arbusti (*Rosa canina*, *Rubus spp.*, *Humulus lupulus*, *Hedera helix*).

Lo studio ha incluso quindi 50 specie Italiane e due specie Nord Americane ormai naturalizzate (*Acer negundo*, *Quercus rubra*) appartenenti a 24 diversi generi e 11 famiglie.

Tutti i campioni sono stati inoltre conservati sia come campione di erbario sia come materiale liofilizzato nella Banca del DNA Forestale dell'Università della Tuscia. La tabella 9 riporta la lista delle specie, il corrispondente Voucher della Banca del DNA e infine la sigla attribuita da GenBank per ogni sequenza che è stata depositata e quindi resa disponibile alla comunità scientifica, per un totale di 156 sequenze.

Campionamento taxon-based

Questo campionamento è stato concentrato all'interno del genere *Quercus* su campioni italiani per un totale di 30 individui: 3 individui per nove specie e un solo individuo per 3 specie endemiche rare come mostrato nella tabella 10 dove sono indicati anche il dettaglio dell'origine geografica, i marcatori utilizzati e gli accession numbers, attribuiti da GenBank, per un totale di 120 sequenze.

Family	Genus	Species	Voucher	GenBank acc. numbers		
				trnH-psbA	rbcL	rpoC1
Aceraceae	<i>Acer</i>	<i>campestre</i> L.	001.01/IT.RI.016	FN687508	FN689361	FN689644
		<i>lobelii</i> Ten.	A.lob.01/IT.CH.078	FN687507	FN689360	FN689643
		<i>monspessulanum</i> L.	004.01/IT.RI.016	FN689729	FN689377	FN689659
		<i>negundo</i> L.	A.neg.01/IT.VT.017	FN687502	FN689355	FN689638
		<i>obtusatum</i> W. et K.	A.ob.01/IT.RI.016	FN687506	FN689359	FN689642
		<i>platanoide</i> L.	004.01/IT.RI.016	FN687504	FN689356	FN689639
		<i>pseudoplatanus</i> L.	005.01/IT.VT.048	FN687503	FN689357	FN689640
		<i>xperonai</i> Schwer.	A.per.01/IT.FI	FN687505	FN689358	FN689641
Fagaceae	<i>Fagus</i>	<i>sylvatica</i> L.	020.01/IT.VT.000	FN687511	FN689362	FN689626
	<i>Castanea</i>	<i>sativa</i> Miller	015.01/IT.VT.000	FN687510	FN689363	FN689628
	<i>Quercus</i>	<i>cerris</i> L.	041.01/IT.VT.017	FN687489	FN675724	FN689627
		<i>coccifera</i> L.	041.01/IT.BA.00	FN687472	FN675713	FN689614
		<i>crenata</i> Lam.	Q.cren.01/IT.VT	FN687486	FN675732	FN689625
		<i>frainetto</i> Ten.	044.01/IT.RM.010	FN687498	FN675741	FN689624
		<i>ilex</i> L.	046.01/VT.011	FN687476	FN675717	FN689623
		<i>macrolepis</i> Kotschy	047.01/IT.BR	FN687483	FN675729	FN689622
		<i>xmorisii</i> Borzi	Q.mor.01/IT.BR	FN687475	FN675716	FN689621
		<i>petraea</i> (Matt.) Liebl.	048.01/IT.RM.010	FN687495	FN675738	FN689620
		<i>pubescens</i> Willd.	049.01/IT.VT.017	FN687493	FN675736	FN689619
		<i>robur</i> L.	051.01/IT.RM.010	FN687490	FN675733	FN689618
		<i>rubra</i> L.	Q.rub.01/IT.VT	FN687509	FN689354	FN689617
		<i>suber</i> L.	054.01/IT.VT.016	FN687481	FN675727	FN689616
		<i>trojana</i> Webb.	055.01/IT.BR	FN687479	FN675725	FN689615
	<i>Fontanesia</i>	<i>phylliraeoides</i> Labill.	F.ph.01/IT.PA	GU120318	FJ862058	FN689629
	<i>Fraxinus</i>	<i>angustifolia</i> L.	021.01/IT.LT.007	GU120315	FJ862055	FN689637
		<i>excelsior</i> L.	022.01/IT.RM.010	GU120316	FJ862056	FN689636
		<i>ornus</i> L.	023.01/IT.VT.011	GU120317	FJ862057	FN689635
	<i>Ligustrum</i>	<i>vulgare</i> L.	L.vul/IT.GR.015	GU120319	FJ862059	FN689634
	<i>Olea</i>	<i>europaea</i> L.	028.01/IT.GR.015	GU120320	FJ862060	FN689633
	<i>Phillyrea</i>	<i>angustifolia</i> L.	083.01/IT.GR.018	GU120321	FJ862062	FN689632
		<i>latifolia</i> L.	082.01/IT.GR.015	GU120322	FJ862063	FN689631
		<i>x media</i> L.	Ph.me.01/IT.CA.090	GU120323	FJ862064	FN689630
Araliaceae	<i>Hedera</i>	<i>helix</i> L.	H.hel.01/IT.TR.003	FN687521	FN689376	FN689658
Betulaceae	<i>Alnus</i>	<i>glutinosa</i> L.	007.01/IT.TR.003	FN687523	FN689372	FN689654
	<i>Corylus</i>	<i>avellana</i> L.	016.01/IT.TR.003	FN687522	FN689371	FN689653
Cannabaceae	<i>Humulus</i>	<i>lupulus</i> L.	Hu.lup.01/IT.TR.003	FN687524	FN689369	FN689651
Rosaceae	<i>Prunus</i>	<i>spinosa</i> L.	Pr.sp.01/IT.TR.003	FN687520	FN689384	FN689648
		<i>monogyna</i> Jacq.	090.01/IT.TR.003	FN687519	FN689370	FN689652
	<i>Rosa</i>	<i>canina</i> L.	Ro.ca.01/IT.TR.003	FN687518	FN689381	FN689661
	<i>Rubus</i>	<i>caesius</i> L.	Ru.ca.01/IT.TR.003	FN687517	FN689382	FN689664
		<i>ulmifolius</i> Schott.	Ru.ul.01/IT.TR.003	FN687516	FN689383	FN689665
Salicaceae	<i>Populus</i>	<i>alba</i> L.	031.01/IT.TR.003	FN687513	FN689368	FN689650
		<i>nigra</i> L.	034.01/IT.TR.003	FN687512	FN689367	FN689649
	<i>Salix</i>	<i>alba</i> L.	057.01/IT.TR.003	FN687514	FN689366	FN689647
		<i>triandra</i> L.	S.tr.01/IT.TR.003	FN687515	FN689365	FN689646
Ulmaceae	<i>Ulmus</i>	<i>minor</i> Miller	072.01/IT.TR.003	FN687525	FN689364	FN689645
Cupressaceae	<i>Cupressus</i>	<i>sempervirens</i> L.	110.01/IT.TR.005	FN689387	FN689378	FN689660
	<i>Juniperus</i>	<i>phoenicea</i> L.	113.01/IT.LT.025	FN689390	FN689373	FN689655
Pinaceae	<i>Larix</i>	<i>decidua</i> Miller	116.01/IT.UD.071	FN689386	FN689379	FN689662
	<i>Pinus</i>	<i>halepensis</i> Miller	125.01/IT.TA.003	FN689388	FN689375	FN689657
		<i>pine</i> L.	131.01/IT.GR	FN689389	FN689374	FN689656
		<i>nigra</i> Arnold	129.01.01/IT.UD.051	FN689385	FN689380	FN689663

Tabella 9 Campionamento floristico

Sono indicati i nomi delle specie, il Voucher, e gli Accession numbers GenBank relativi ai tre marcatori Barcode

Species	Subgenus	Origin	GenBank Accession numbers			
			trnH-psbA	rbclA	matK	matK2
<i>Q. ilex</i>	Sclerophyllodris	Latium	FN687476	FN675717	FN675317	FN674513
		Sardinia	FN687477	FN675718	FN675314	FN674514
		Apulia	FN687478	FN675719	FN675315	FN674515
<i>Q. coccifera</i>		Apulia	FN687472	FN675713	FN675318	FN674487
		Sardinia	FN687473	FN675714	FN675312	FN674511
		Sardinia	FN687474	FN675715	FN675313	FN674512
<i>Q. morisii</i> *		Apulia	FN687475	FN675716	FN675316	FN674516
<i>Q. cerris</i>	Cerris	Basilicata	FN687488	FN675722	FN675335	FN674500
		Liguria	FN687487	FN675723	FN675340	FN674501
		Apulia	FN687489	FN675724	FN675341	FN674502
<i>Q. suber</i>		Latium	FN687481	FN675727	FN675332	FN674507
		Liguria	FN687480	FN675726	FN675336	FN674509
		Apulia	FN687482	FN675728	FN675337	FN674508
<i>Q. crenata</i>		Trentino	FN687485	FN675730	FN675334	FN674503
		Liguria	FN687484	FN675731	FN675338	FN674505
		Latium	FN687486	FN675732	FN675339	FN674504
<i>Q. trojana</i> **		Apulia	FN687479	FN675725	FN675331	FN674510
<i>Q. macrolepis</i> **		Apulia	FN687483	FN675729	FN675333	FN674506
<i>Q. petraea</i>	Quercus	Latium	FN687495	FN675738	FN675321	FN674496
		Umbria	FN687496	FN675739	FN675327	FN674497
		Piedmont	FN687497	FN675740	FN675328	FN674498
<i>Q. robur</i>		Latium	FN687490	FN675733	FN675319	FN674499
		Emilia	FN687491	FN675734	FN675323	FN674494
		Tuscany	FN687492	FN675735	FN675324	FN674495
<i>Q. frainetto</i>		Latium	FN687498	FN675741	FN675322	FN674489
		Apulia	FN687499	FN675721	FN675329	FN674490
		Basilicata	FN687500	FN675742	FN675330	FN674491
<i>Q. pubescens</i>		Latium	FN687493	FN675736	FN675320	FN674488
		Campania	FN687494	FN675737	FN675325	FN674492
		Sardinia	FN687501	FN675720	FN675326	FN674493

Tabella 10 Campionamento taxon-based

Sono riportate le specie, l'origine geografica e Accession numbers GenBank dei loci Barcode

* raro ibrido F1

** specie la cui presenza è limitata ad un'area molto limitata

I marcatori molecolari

matK

Il gene *trnK* (*tRNA^{Lys}(UUU)*) contiene un introne del gruppo II (*trnK11*) al cui interno si trova l'open reading frame (ORF) *matK* (Figura 8). Gli introni di tipo II infatti possono contenere ORFs che codificano per peptidi coinvolti nella propagazione degli introni e nello splicing come maturasi e trascrittasi inversa.

Come mostrato in figura 9 le ORF che si trovano negli introni di tipo II possiedono una struttura canonica caratterizzata da tre domini: un dominio RT (Retro-Transcriptasi) che contiene diversi subdomini (0-7); un dominio X associato alla funzione di maturasi; un dominio opzionale En associato alla funzione di endonucleasi; può essere presente un dominio D caratterizzato funzionalmente ma non conservato (San Filippo and Lambowitz 2002). Tuttavia *matK* ORF possiede un dominio RT molto degenerato, mentre mantiene ben conservato il dominio X associato, come già detto, alla funzione di maturasi che quindi è quella che viene solitamente associata a questa proteina, anche se la completa funzione biochimica deve essere ancora dimostrata (G. Hausner et al 2006).

Sia l'introne che la sua ORF codificante sono stati oggetto di numerosi studi nel campo dell'evoluzione molecolare. Negli studi evolutivi sono stati spesso utilizzati come markers per la costruzione di filogenesi a causa della elevata velocità evolutiva tipica di questa zona e la sua presenza praticamente ubiquitaria (Hilu and Liang 1997; Kelchner 2002).

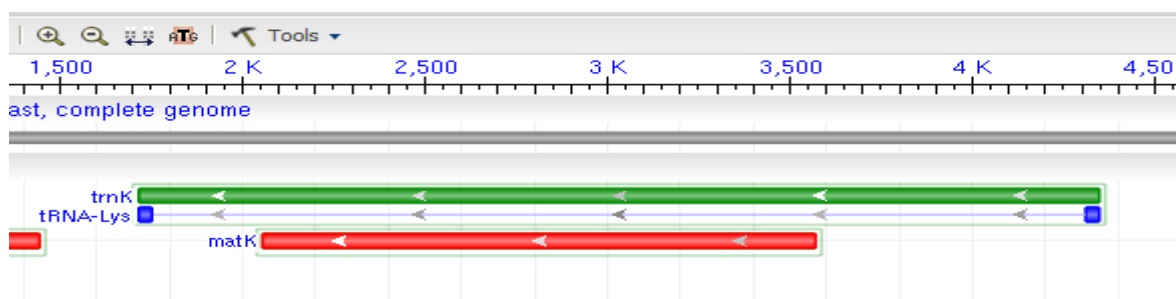


Fig. 8: Il gene *trnK* (verde) è costituito da 2 piccoli esoni (blu) e un grande introne (linea sottile grigia) al suo interno, ma non in tutta la sua lunghezza, c'è *matK* ORF indicato in rosso

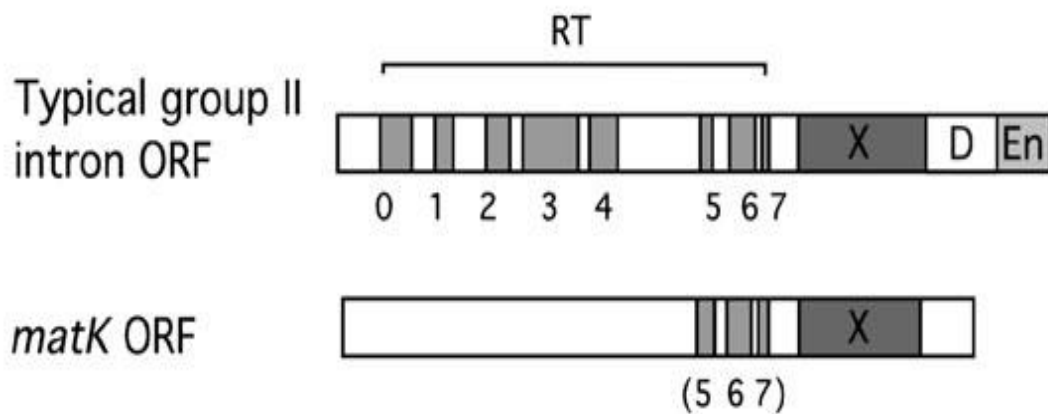


Fig 9: Struttura canonica dei domini delle ORF negli introni i tipo II e nella matK ORF

rbcL

Il gene *rbcL*, codifica per la subunità grande dell'enzima ribulosio-1,5-difosfato carbossilasi/ossigenasi (RuBisCO). Costituisce un caposaldo della fotosintesi ed è responsabile della trasformazione del carbonio inorganico in composti organici. Viene definita come la proteina più abbondante nel mondo e costituisce fino al 50% delle proteine solubili nelle piante. Tale abbondanza è necessaria a causa delle sue basse performance catalitiche, infatti confonde il reale substrato fotosintetico, CO₂, con il prodotto O₂, determinando una notevole dispersione di energia. Tuttavia esistono delle differenze nella cinetica della Rubisco nelle diverse specie e indipendenza di differenti pressioni ambientali.

Il gene è costituito da un solo esone (Fig. 10), rappresenta uno dei geni più conservati ed è stato uno dei marcatori più utilizzati in studi filogenetici, sebbene in un recente lavoro venga indicato come soggetto a selezione positiva nella maggior parte delle piante terrestri (Kapralov MV et al. 2007).

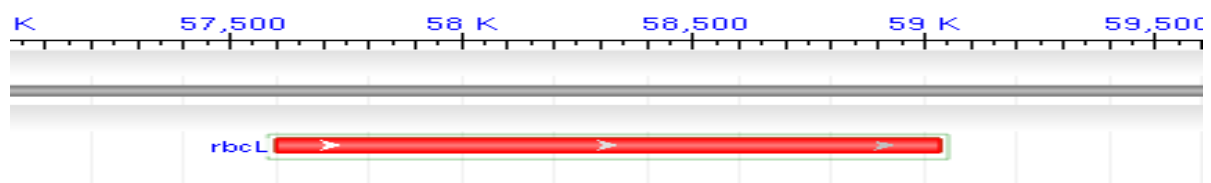


Figura 10: Esone unico del gene *rbcL*

rpoC1

Il gene cloroplastico *rpoC1* insieme a *rpoC2*, costituisce la regione *rpoC* che si trova localizzata nella regione LSC (Large Single Copy) del genoma plastidiale nella maggior parte delle Angiosperme (Hudson et al 1988). Questi 2 geni, insieme a *rpoB* codificano le tre subunità della cpRNA polimerasi.(Fig.11)

Il gene *rpoC1* è interrotto da un singolo introne nella maggior parte, ma non in tutte, le piante terrestri. La perdita dell'introne è stata osservata in diversi taxa tra cui Poaceae (Katayama and Ogiwara, 1993), Cactaceae subfamily Cactoideae (Wallace and Cota, 1996) e altri.

A causa della natura conservativa del genoma cloroplastico la perdita di tale introne rappresenta un evento raro che viene interpretato come un robusto segnale di relazione dei lignaggi in cui avviene (Downie et al., 1991).

Anche questo gene è stato ampiamente utilizzato per studi filogenetici, sebbene con limitati successi nella risoluzione di livelli tassonomici più bassi.(Downie et al 1996).

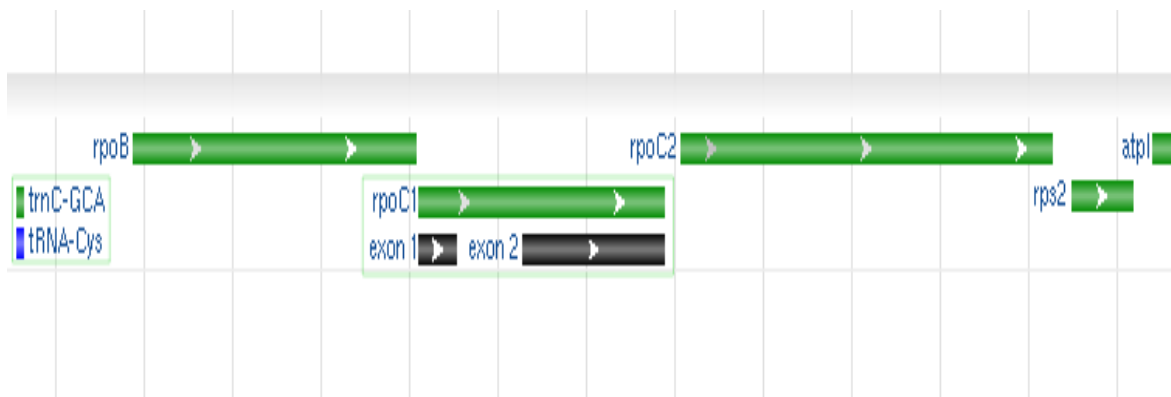


Fig. 11: Sono riportati in verde i 3 geni *rpoC1*, *rpoC2* e *rpoB*.

Nel gene *rpoC1* i due esoni in nero separati da un introne.

trnH-psbA

Gli studi su questa regione cominciarono con Aldrich (Aldrich *et al.*, 1988) che mostrò come fosse caratterizzata da inserzioni e delezioni (indels) anche tra specie vicine. Studi successivi hanno mostrato il suo valore a livello sistematico (Sang *et al.*, 1997) evidenziando anche la sua alta variabilità rispetto a *matK* e *trnL-trnF*. La sua utilità è stata anche mostrata da Hamilton (1999b) e Holdregger and Abbott (2003) che la hanno utilizzata anche per studi intraspecie. Purtroppo la sua alta variabilità rende difficile il suo allineamento a livelli più alti (intergenere) (Laurales: Renner, 1999, Saxifragaceae: Soltis et al 2001, Lecythidaceae: Hamilton et al., 2003). Diversi studi hanno evidenziato come *trnH-psbA* (Fig.12) contenga un alto numero di

siti variabili (Hamilton et al., 2003), accompagnato ad una lunghezza relativamente bassa (mediamente di 465 bp) (Shaw et al. 2005). Queste caratteristiche, insieme al fatto di essere facilmente amplificabile e sequenziabile nei diversi *taxa*, ne hanno determinato il suo largo utilizzo negli studi barcode, tuttavia, come accennato, tale marcatore non risulta attualmente come barcode ufficiale delle piante. Il PlantWG, ha infatti ritenuto che la presenza di omopolimeri generi un drastico calo della qualità dell'elettroferogramma e quindi la successiva necessità di un grosso lavoro manuale di assemblaggio tra forward e reverse, con conseguente perdita dell'automatismo insito nella metodica barcoding. Resta inoltre il problema dei metodi di analisi dei dati al sopra di genere o famiglia che, secondo quanto sperimentato da alcuni autori può tuttavia essere superato attraverso l'utilizzo e la costruzione di una “*supermatrix*” (Kress et al. 2009).

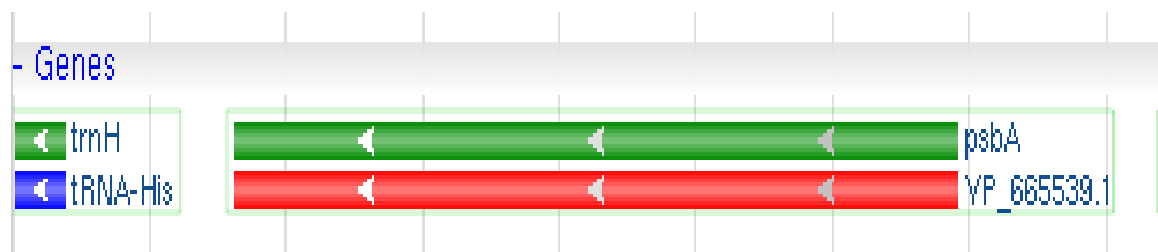


Fig.12: In verde sono mostrati i due geni *trnH* e *psbA* adiacenti

Il protocollo molecolare

Estrazione del DNA

E' stato utilizzato un kit commerciale specificamente preposto per l'estrazione del DNA (*DNAeasy Plant Minikit, Qiagen*) che si caratterizza per la mancanza del fenolo e del cloroformio, sostanze altamente tossiche, e per la rapidità di esecuzione delle varie operazioni che portano ad un'ottima purificazione dell'acido nucleico.

Il kit comporta l'uso di vari tamponi e di due diverse colonnine "*Microspin*" per la purificazione dell'acido nucleico. La procedura di estrazione è stata condotta attraverso numerose fasi di seguito riassunte:

Il materiale vegetale utilizzato è stato o liofilizzato presso la Banca Centrale del DNA Forestale di Cittaducale (RI) o frantumato in piccoli mortai con azoto liquido a cui, successivamente, è stato aggiunto il tampone di lisi per favorire l'ulteriore rottura delle cellule e la fuoriuscita del materiale citoplasmatico e nucleare; è stato aggiunto l'enzima RNAsi per evitare contaminazioni di RNA con DNA e l'omogenato è stato posto ad incubare a 65 °C per 10 minuti. La reazione di lisi è stata bloccata con un tampone che determina la precipitazione delle proteine e dei polisaccaridi. La purificazione dell'estratto è stata eseguita con vari passaggi su colonnine filtrando i residui cellulari mediante l'uso di una microcentrifuga da banco. Infine, il DNA è stato raccolto in una microprovetta mediante centrifugazione con un tampone di eluizione a bassa concentrazione salina. È possibile conservare il DNA così estratto nella soluzione utilizzata per l'eluizione per breve tempo a 4°C o, per lungo tempo, a -20°C.

L'amplificazione del frammento desiderato

È stata condotta mediante la tecnica della "PCR" utilizzando un *thermal cycler* (termocicizzatore) Gene Amp PCR System 2400 della Perkin Elmer. Aliquote definite dei DNA sono state fatte reagire con un preparato commerciale liofilizzato contenente tutti gli elementi necessari alla reazione di amplificazione (enzima TAQ polimerasi, tampone, deossinucleotidi) a cui sono stati aggiunti segmenti specifici di DNA di pochi nucleotidi denominati *primers*, che hanno permesso l'innescio della amplificazione della regione del DNA che interessa grazie ad una loro specifica temperatura di appaiamento. La tabella 12 mostra le coppie utilizzate in questo studio.

Locus	Primer Name	Sequence	Reference
rbcL	RbcL-fw. rbcL-rev.	ATGTCACCACAAACAGAGACTAAAGC CTTCTGCTACAAATAAGAATCGATCTC	(Kress and Eriksson 2007)
rbcL(2)	rbcL2-fw. rbcL2-rev.	ATGTCACCACAAACAGAAAC TCGCATGTACCTGCAGTAGC	(Kress et al. 2005)
rpoC1	rpoC1-fw rpoC1-rev	GTGGATACACTTCTTGATAATGG TGAGAAAACATAAGTAAACGGGC	(www.kew.org/barcoding/protocols)
trnH-psbA	trnH-psbA2_fw trnH-psbA2_rev	CGCGCATGGTGGATTACAAATCC GTTATGCATGAACGTAATGCTC	(Shaw et al. 2007)
matK	2.1 5	CCTATCCATCTGGAAATCTTAG GTTCTAGCACAAGAAAGTCG	(www.kew.org/barcoding/protocols)
matK_Quercus	matKq-fw matKq-rev	TGCATTATTGCGGTTCTTTTC CGCGTGCACTACTTTGTGT	This study
MatK2_Quercus	MatK2q-fw MatK2q-rev	GTACTTGATGCGGGAAATGC CAATGATTGCAAATCCTTCTGA	This study

Tabella 12 -Primers utilizzati

L'amplificazione, condotta nel termocicizzatore, prevedeva un ciclo di reazioni a catena che hanno consentito la separazione delle due eliche di DNA, l'appaiamento dei primers alle regioni omologhe e la produzione di nuove copie dei filamenti nucleari; ogni ciclo è stato ripetuto più volte. L'amplificazione si è conclusa con 10 minuti a 72°C, in modo da terminare i frammenti eventualmente rimasti incompleti. Il volume finale è stato di 25µl di cui 50 ng di DNA genomico, 20 ng di primers per reazione, una dose di puReTaq Ready-To-Go PCR Beads (Amersham) ed acqua quanto basta per raggiungere il volume desiderato, per raggiungere una concentrazione finale di 200 µM di ciascun nucleotide e 1,5 mM di MgSO₄. Per verificare il successo della PCR e delle relative amplificazioni, è stata effettuata un' elettroforesi su gel d'Agarosio. La polvere di Agarosio è stata sciolta a caldo in tampone TBE 1X, con l'aggiunta di bromuro di etidio, una sostanza a struttura aromatica che intercalandosi nella doppia elica del DNA la rende visibile ai raggi ultravioletti. La soluzione così ottenuta è stata fatta gelificare su un apposito vassoio e con l'ausilio di un "pettine" sono stati prodotti dei pozzetti, in ognuno di essi sono stati caricati: 6 µl di H₂O, 2 µl di Blu di Bromofenolo e 2 µl di DNA derivato dalla miscela di amplificazione. Il gel solidificato è stato posto in una cella elettroforetica contenente il buffer TBE 1X ed applicando agli elettrodi una differenza di potenziale elettrico di 80 V (d.d.p.) per circa 30 min. Al termine della corsa elettroforetica è

stato possibile osservare, mediante un trasilluminatore a raggi UV, le bande di DNA, il loro grado di pulizia e di integrità.

La purificazione del DNA amplificato

I prodotti delle PCR sono stati purificati seguendo il protocollo del manuale di utilizzo del kit commerciale GFX PCR DNA and *Gel Band Purification Kit* (GE Healthcare). Il campione (25µl) è stato stabilizzato con un tampone in grado di denaturare le proteine e/o dissolvere l'agarosio. Esso, inoltre, applicato nella colonna microspin, fa sì che il DNA, dopo successive centrifugazioni, si leghi alle membrane. Un seguente lavaggio rimuove i sali e le altre contaminazioni dalle membrane legate al DNA. Il campione così purificato viene eluito con un tampone e lasciato incubare a temperatura ambiente per un paio di minuti. Dopo aver eseguito un'ultima centrifugazione, si preleva il DNA che sarà conservato in freezer a -20°C.

Il sequenziamento

Le sequenze purificate sono state sottoposte a reazioni di sequenziamento presso i laboratori della *Eurofins MWG Operon* (Germania). La sequenza nucleotidica di ciascun frammento è stata determinata su entrambe le eliche, utilizzando in ciascuna reazione il corrispettivo primer in posizione 5'.

La lettura dei tracciati grezzi, detta anche identificazione delle basi (*base-calling*) ad oggi viene effettuata tramite un software che legge automaticamente le basi, allinea le sequenze simili e fornisce una base intuitiva per il lavoro di correzione di bozze (*editing*). Mentre leggono i tracciati, questi programmi assegnano punteggi di probabilità all'accuratezza d'identificazione di ciascuna base. Le quattro tracce corrispondenti ai quattro spettri di fluorescenza sono fuse in un unico file, mantenendo le caratteristiche grafiche dei picchi. Al passaggio successivo il computer calcola il punto in cui si attende di trovare un picco, usando un procedimento di calcolo basato sulla distanza media tra i picchi in alcune regioni della sequenza. Gli algoritmi migliori sono in grado di effettuare aggiustamenti per l'aumento della distanza tra i picchi che avviene con l'avanzamento della corsa elettroforetica e per le variazioni locali nel contenuto di GC. In seguito, l'algoritmo rileva i valori massimi locali per ciascuna delle quattro serie di tracciati e controlla se ciascun picco ricade nello spazio previsto rispetto ai picchi adiacenti. Poiché non tutti i picchi sono della stessa altezza, per decidere se un picco è reale gli algoritmi usano di solito un valore soglia, che calcola la grandezza relativa di un minimo e massimo locale. (Figura 13)

I risultati delle reazioni di sequenza sono stati visualizzati sotto forma di elettroferogrammi con il programma CHROMAS 2.1 (Technelysium), in cui la successione dei diversi picchi rappresenta la sequenza nucleotidica reale del DNA studiato.

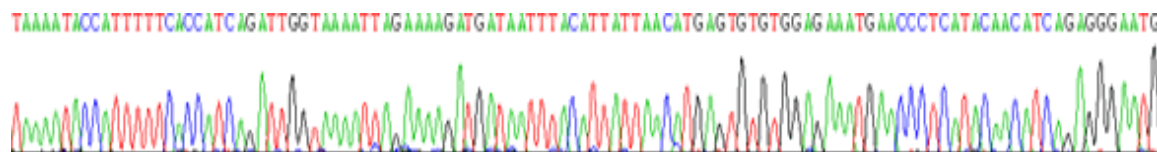


Figura 13: -Esempio di elettroferogramma di sequenza-

Le prime 30/50 basi di un elettroferogramma sono normalmente indistinguibili dal “rumore di fondo”, a causa della migrazione anomala di brevi frammenti di DNA che contengono agglomerati di coloranti. Inoltre i tracciati diventano progressivamente meno uniformi con l’avanzare della corsa: aumentano gli effetti della diffusione molecolare e contemporaneamente diminuiscono le differenze relative di massa tra frammenti successivi.

Poiché gli algoritmi automatici sovente commettono alcuni errori nell’identificazione delle basi, si è proceduto manualmente alla correzione delle sequenze.

Tutte le sequenze sono state verificate, controllate e assemblate con il programma DNAMAN 2.3 (Lynnon-Biosoft).

Si è proceduto alla correzione nel caso in cui l’elettroferogramma della sequenza fornito dall’*Eurofins* considerasse la coda di un picco precedente come se fosse una base, nel caso in cui non fosse stata considerata l’eterozigosità in alcuni nucleotidi, nel caso di sbagliato conteggio di una serie numerosa di basi uguali (omopolimeri).

L'analisi dei dati

Come evidenziato nella parte introduttiva non esiste una metodologia standard per l'analisi dei dati e quindi la valutazione della percentuale di successo può avvenire attraverso approcci di vario tipo.

Nel campionamento di tipo floristico le analisi sono state eseguite seguendo un approccio basato sulla similarità. Infatti, poiché il dataset è costituito da specie appartenenti a generi e famiglie diverse, non è stato possibile utilizzare metodi che richiedessero come base un multiallineamento globale.

Quindi attraverso l'utilizzo di BLUSTClust 2.2.20 (<http://ftp.ncbi.nlm.nih.gov/blast/>), per ogni marcatore singolo o in combinazione, sono stati generati gli aplotipi. BLUSTClust consente di raggruppare sequenze nucleotidiche e proteiche secondo diversi livelli di stringenza dei parametri di controllo.

Per le nostre analisi è stata utilizzata la seguente linea di comando :

```
blastclust -i infile -o outfile -p F -L1 -b T -S100
```

che consente di raggruppare le sequenze esattamente identiche. In questo modo sono state individuate le sequenze singole (che successivamente nella tabella saranno indicate con **Hs** e indicano che la specie è stata discriminata con successo, poiché non esistono campioni intraspecie) e i gruppi di sequenze identiche (indicati nella tabella riassuntiva con **Ht** che indicano le specie che condividendo la stessa sequenza sono state raggruppate insieme).

Il campionamento taxon-based all'interno del genere *Quercus* è stato invece multiallineato con MUSCLE (Edgar 2004) con i parametri di default. Il multiallineamento è stato successivamente visualizzato e controllato attraverso l'utilizzo del programma Seaview (Galtier et al. 1996). Il pacchetto DnaSP 4.5 (Rozas et al. 2003) è stato quindi utilizzato per la generazione della lista degli aplotipi visualizzati successivamente con Network.

Infine attraverso MEGA 4.0 (Tamura et al. 2007) è stata calcolata la variabilità inter e intra specie utilizzando il modello evolutivo Kimura 2-parametri, e costruito il dendrogramma Neighbour-Joining (NJ).

E' stato inoltre eseguito il test del bootstrap (1000 repliche) per valutare la robustezza dei nodi.

5 RISULTATI

Campionamento floristico

Come detto in precedenza, questo tipo di approccio fornisce sia indicazioni su aspetti tecnico-molecolari (risposta primers e amplificazione in generale), sia una prima indicazione del comportamento e variabilità dei diversi loci nei taxa analizzati, ma senza fornire nessun dettaglio su aspetti intraspecifici.

La tabella 13A mostra i risultati ottenuti per i quattro marcatori utilizzati: rbcL, rpoC1, trnH-psbA, e matK. Possiamo osservare come per rpoC1 e trnH-psbA una sola coppia di primers sia stata sufficiente per amplificare tutto il dataset; per rbcL è stato necessario lo sviluppo di una ulteriore coppia per procedere all'amplificazione nelle Fagaceae. Come atteso il locus matK è risultato essere il più problematico poiché solo il 56% del dataset è stato amplificato dalla coppia di primers, mentre in Castanea, Quercus, Ulmus, Rosa, Rubus, Pinus, Cupressus, e Juniperus non si è riscontrata nessuna amplificazione, e in Populus e Acer una banda molto debole.

Gli elettroferogrammi ottenuti dopo il sequenziamento erano chiaramente leggibili e anche le regioni omopolimeriche presenti nel locus trnH-psbA non hanno disturbato in modo rilevante la qualità del tracciato.

La tabella 13B mostra le percentuali di successo dei marcatori analizzati sia singolarmente che nelle varie combinazioni nelle famiglie analizzate. Le analisi sono state effettuate con BlastClust e per ogni marcatore sono riportati gli aplotipi singoli (**Hs**) e quelli condivisi (**Ht**) e ricordo che, poiché esiste una sola sequenza per specie, la presenza di un aplotipo singolo indica il successo nella discriminazione.

Una prima valutazione della percentuale di successo indica il locus trnH-psbA come il migliore con il 73% seguito da rpoC1 48% e rbcL 40%, inoltre nessuna delle combinazioni di più loci riesce ad ottenere percentuali maggiori. L'osservazione della tabella ci consente di notare come le due regioni rbcL e rpoC1 abbiano comportamenti diversi nei diversi taxa, infatti rpoC1 mostra in Fraxinus, Phillyrea e Rubus buone percentuali di successo migliori di rbcL, mentre in Castanea fallisce addirittura nel discriminare a livello di genere (unico caso in tutto il dataset).

Il conteggio della percentuale di successo può risultare sovrastimato dall'inclusione dei generi rappresentati da una sola specie, la cui esclusione porta infatti una riduzione dei valori che

diventano 61% per trnH-psbA, 25% rpoC1 e 13% rbcL. Tali risultati sono tuttavia decisamente più bassi di quelli ottenuti da altri autori e l'osservazione dei generi presenti con più specie ci aiuta a capire la causa. I risultati ottenuti nel genere *Quercus* infatti, sono decisamente negativi: solo 3 specie su tredici vengono discriminate con successo e considerando che una delle tre specie è una specie naturalizzata (*Q. rubra*) si intuisce la criticità della situazione.

A questo punto, riconsiderando i conteggi escludendo il genere *Quercus*, le performance dei marcatori diventano: trnH-psbA 82% , rpoC1 34% , rbcL 17% .

Region	rbcL	rpoC1	trnH-psbA	matK
% PCR success	100	100	100	56%
% sequencing success	100	100	100	—
Total no. primer pair used	2*	1	1	1
Sequence length; mean (range)	670	554	539.7 (347–686)	—

Tab. 13A: Schema riassuntivo delle caratteristiche dei loci utilizzati nel campionamento floristico

Tab. 13B: Schema riassuntivo degli aplotipi generati nell'approccio floristico

[illegible]

Output generati dal programma BlastClust

IV a: Output blastclust locus rbcL: 31 haplotypes

- *Quercus cerris*, *Quercus crenata*, *Quercus frainetto*, *Quercus pubescens*, *Quercus robur*, *Quercus suber*, *Quercus macrolepis*, *Quercus petraea*, *Quercus trojana*.
- *Acer monspessulanum*, *Acer obtusatum*, *Acer peronai*.
- *Fraxinus angustifolia*, *Fraxinus excelsior*, *Fraxinus ornus*.
- *Phillyrea angustifolia*, *Phillyrea latifolia*, *Phillyrea media*.
- *Quercus coccifera*, *Quercus ilex*, *Quercus morisii*.
- *Acer lobelii*, *Acer platanoides*.
- *Populus alba*, *Populus nigra*.
- *Pinus halepensis*, *Pinus pinea*.
- *Rubus caesius*, *Rubus ulmifolius*.
- *Salix alba*, *Salix triandra*.
- *Acer campestre*.
- *Alnus glutinosa*.
- *Acer negundo*.
- *Acer psedoplatanus*.
- *Corylus avellana*.
- *Craetegus monogyna*.
- *Castanea sativa*.
- *Cupressus sempervirens*.
- *Fontanesia phylliraeoides*.
- *Fagus sylvatica*.
- *Hedera helix*.
- *Humulus lupulus*.
- *Juniperus phoenicea*.
- *Larix decidua*.
- *Ligustrum vulgare*.
- *Olea europaea*.
- *Pinus nigra*.
- *Prunus spinosa*.
- *Quercus rubra*.
- *Rosa canina*.
- *Ulmus minor*.

IV b. Output blastclust locus rpoC1: 33 haplotypes

- *Castanea sativa*, *Quercus Suber*, *Quercus cerris*, *Quercus frainetto*, *Quercus pubescens*, *Quercus robur*, *Quercus crenata*, *Quercus macrolepis*, *Quercus petraea*, *Quercus trojana*.
- *Acer campestre*, *Acer lobelii*, *Acer platanoides*.
- *Pinus halepensis*, *Pinus nigra*, *Pinus pinea*.
- *Quercus coccifera*, *Quercus ilex*, *Quercus morisii*.
- *Acer monspessulanum*, *Acer obtusatum*.
- *Populus alba*, *Populus nigra*.
- *Phillyrea latifolia*, *Phillyrea media*.
- *Salix alba*, *Salix triandra*.
- *Alnus glutinosa*.

- *Acer negundo*.
- *Acer peronai*.
- *Acer pseudoplatanus*.
- *Corylus avellana*.
- *Craetegus monogyna*.
- *Cupressus sempervirens*.
- *Fontanesia phylliraeoides*.
- *Fraxinus angustifolia*.
- *Fraxinus excelsior*.
- *Fraxinus ornus*.
- *Fagus sylvatica*.
- *Hedera helix*.
- *Humulus lupulus*.
- *Juniperus phoenicea*.
- *Larix decidua*.
- *Ligustrum vulgare*.
- *Olea europaea*.
- *Phillyrea angustifolia*.
- *Prunus spinosa*.
- *Quercus rubra*.
- *Rosa canina*.
- *Rubus caesius*.
- *Rubus ulmifolius*.
- *Ulmus minor*.

IV c. Output blastclust locus trnH-psbA:44 haplotypes

- *Quercus pubescens*, *Quercus robur*, *Quercus petraea*.
- *Quercus cerris*, *Quercus crenata*, *Quercus suber*.
- *Phillyrea latifolia*, *Phillyrea media*.
- *Quercus macrolepis*, *Quercus trojana*.
- *Quercus ilex*, *Quercus morisii*.
- *Salix alba*, *Salix triandra*.
- *Pinus nigra*.
- *Pinus pinea*.
- *Pinus halepensis*.
- *Larix decidua*.
- *Fontanesia phylliraeoides*.
- *Cupressus sempervirens*.
- *Olea europaea*.
- *Quercus rubra*.
- *Phillyrea angustifolia*.
- *Quercus frainetto*.
- *Castanea sativa*.
- *Juniperus phoenicea*.
- *Alnus glutinosa*.
- *Fraxinus angustifolia*.
- *Fraxinus excelsior*.
- *Quercus coccifera*.

- *Ligustrum vulgare*.
- *Acer negundo*.
- *Acer lobelii*.
- *Acer platanoides*.
- *Acer peronai*.
- *Hedera helix*.
- *Acer campestre*.
- *Acer monspessulanum*.
- *Fagus sylvatica*.
- *Acer obtusatum*.
- *Acer pseudoplatanus*.
- *Corylus avellana*.
- *Fraxinus ornus*.
- *Rubus ulmifolius*.
- *Rosa canina*.
- *Humulus lupulus*.
- *Rubus caesius*.
- *Prunus spinosa*.
- *Craetegus monogyna*.
- *Pinus alba*.
- *Pinus nigra*.
- *Ulmus minor*.

IV d. Output blastclust locus *rbcL+rpoC1*: 36 haplotypes

- *Quercus cerris*, *Quercus crenata*, *Quercus frainetto*, *Quercus pubescens*, *Quercus robur*, *Quercus suber*, *Quercus macrolepis*, *Quercus petraea*, *Quercus trojana*.
- *Quercus coccifera*, *Quercus ilex*, *Quercus morisii*.
- *Acer lobelii*, *Acer platanoides*.
- *Acer monspessulanum*, *Acer obtusatum*.
- *Pinus alba*, *Pinus nigra*.
- *Pinus halepensis*, *Pinus pinea*.
- *Phillyrea latifolia*, *Phillyrea media*.
- *Salix alba*, *Salix triandra*.
- *Acer campestre*.
- *Alnus glutinosa*.
- *Acer negundo*.
- *Acer peronai*.
- *Acer pseudoplatanus*.
- *Corylus avellana*.
- *Craetegus monogyna*.
- *Castanea sativa*.
- *Cupressus sempervirens*.
- *Fraxinus angustifolia*.
- *Fraxinus excelsior*.
- *Fraxinus ornus*.
- *Fontanesia phylliraeoides*.
- *Fagus sylvatica*.

- *Hedera helix*.
- *Humulus lupulus*.
- *Juniperus phoenicea*.
- *Larix decidua*.
- *Ligustrum vulgare*.
- *Olea europaea*.
- *Phillyrea angustifolia*.
- *Pinus nigra*.
- *Prunus spinosa*.
- *Quercus rubra*.
- *Rosa canina*.
- *Rubus caesius*.
- *Rubus ulmifolius*.
- *Ulmus minor*.

IV e. Output blastclust locus rbcL+trnH-psbA: 44 haplotypes

- *Quercus pubescens*, *Quercus robur*, *Quercus petraea*.
- *Quercus cerris*, *Quercus crenata*, *Quercus suber*.
- *Phillyrea latifolia*, *Phillyrea media*.
- *Quercus macrolepis*, *Quercus trojana*.
- *Quercus ilex*, *Quercus morisii*.
- *Salix alba*, *Salix triandra*.
- *Pinus nigra*.
- *Pinus pinea*.
- *Pinus halepensis*.
- *Larix decidua*.
- *Fontanesia phylliraeoides*.
- *Cupressus sempervirens*.
- *Olea europaea*.
- *Quercus rubra*.
- *Phillyrea angustifolia*.
- *Quercus frainetto*.
- *Castanea sativa*.
- *Juniperus phoenicea*.
- *Alnus glutinosa*.
- *Fraxinus angustifolia*.
- *Fraxinus excelsior*.
- *Quercus coccifera*.
- *Ligustrum vulgare*.
- *Acer negundo*.
- *Acer lobelii*.
- *Acer platanoides*.
- *Acer peronai*.
- *Hedera helix*.
- *Acer campestre*.
- *Acer monspessulanum*.
- *Fagus sylvatica*.
- *Acer obtusatum*.

- *Acer pseudoplatanus*.
- *Corylus avellana*.
- *Fraxinus ornus*.
- *Rubus ulmifolius*.
- *Rosa canina*.
- *Humulus lupulus*.
- *Rubus caesius*.
- *Prunus spinosa*.
- *Craetagus monogyna*.
- *Populus alba*.
- *Pinus nigra*.
- *Ulmus minor*.

IV f. Output blastclust locus rpoC1+trnH-psbA: 44 haplotypes

- *Quercus pubescens*, *Quercus robur*, *Quercus petraea*.
- *Quercus cerris*, *Quercus crenata*, *Quercus suber*.
- *Phillyrea latifolia*, *Phillyrea media*.
- *Quercus macrolepis*, *Quercus trojana*.
- *Quercus ilex*, *Quercus morisii*.
- *Salix alba*, *Salix triandra*.
- *Pinus nigra*.
- *Pinus pinea*.
- *Pinus halepensis*.
- *Larix decidua*.
- *Fontanesia phylliraeoides*.
- *Cupressus sempervirens*.
- *Olea europaea*.
- *Quercus rubra*.
- *Phillyrea angustifolia*.
- *Quercus frainetto*.
- *Castanea sativa*.
- *Juniperus phoenicea*.
- *Alnus glutinosa*.
- *Fraxinus angustifolia*.
- *Fraxinus excelsior*.
- *Quercus coccifera*.
- *Ligustrum vulgare*.
- *Acer negundo*.
- *Acer lobelii*.
- *Acer platanoides*.
- *Acer peronai*.
- *Hedera helix*.
- *Acer campestre*.
- *Acer monspessulanum*.
- *Fagus sylvatica*.
- *Acer obtusatum*.
- *Acer psedoplatanus*.

- *Corylus avellana*.
- *Fraxinus ornus*.
- *Rubus ulmifolius*.
- *Rosa canina*.
- *Humulus lupulus*.
- *Rubus caesius*.
- *Prunus spinosa*.
- *Craetegus monogyna*.
- *Populus alba*.
- *Populus nigra*.
- *Ulmus minor*.

IV g. Output blastclust locus rbcL+rpoC1+trnh-psbA: 44 haplotypes

Quercus pubescens, *Quercus robur*, *Quercus petraea*.

Quercus cerris, *Quercus crenata*, *Quercus suber*.

Phillyrea latifolia, *Phillyrea media*.

Quercus macrolepis, *Quercus trojana*.

Quercus ilex, *Quercus morisii*.

Salix alba, *Salix triandra*.

Pinus nigra.

Pinus pinea.

Pinus halepensis.

Larix decidua.

Fontanesia phylliraeoides.

Cupressus sempervirens.

Olea europaea.

Quercus rubra.

Phillyrea angustifolia.

Quercus frainetto.

Castanea sativa.

Juniperus phoenicea.

Alnus glutinosa.

Fraxinus angustifolia.

Fraxinus excelsior.

Quercus coccifera.

Ligustrum vulgare.

Acer negundo.

Acer lobelii.

Acer platanoides.

Acer peronai.

Hedera helix.

Acer campestre.

Acer monspessulanum.

Fagus sylvatica.

Acer obtusatum.

Acer psedoplatanus.

Corylus avellana.

Fraxinus ornus.
Rubus ulmifolius
Rosa canina.
Humulus lupulus.
Rubus caesius.
Prunus spinosa.
Craetegus monogyna.
Populus alba.
Populus nigra.
Ulmus minor.

Campionamento taxon-based

I risultati della precedente esplorazione generale ci ha consentito di scegliere il genere *Quercus* sul quale è stato quindi effettuato un secondo tipo di studio definito “*taxon-based*”, che prevede un campionamento intraspecie più dettagliato (2-5 individui) con lo scopo di esplorare in maniera esaustiva il genere stesso.

A causa delle variabilità molto bassa riscontrata nelle Fagaceae, il marcatore *rpoC1* è stato escluso da questo tipo di analisi. I marcatori utilizzati sono quindi *rbcL*, *trnH-psbA*, *matK_Quercus* (che corrisponde al frammento barcode ufficiale) e *matK2_Quercus* (che corrisponde ad un frammento a monte di questo da noi selezionato); La tabella 14, illustra alcuni semplici parametri della variabilità (siti polimorfici, numero aplotipi, distanza media e range inter e intra specie dei 4 marcatori).

Possiamo notare come la variabilità della regione codificante *rbcL* sia decisamente molto bassa con tre soli siti polimorfici, mentre le altre regioni mostrano una variabilità decisamente maggiore. Tuttavia, la produzione grafica dei Network (Figura 14a, 14b) ci consente di visualizzare in modo semplice e chiaro come la variabilità sia principalmente concentrata tra i sottogeneri piuttosto che tra le specie.

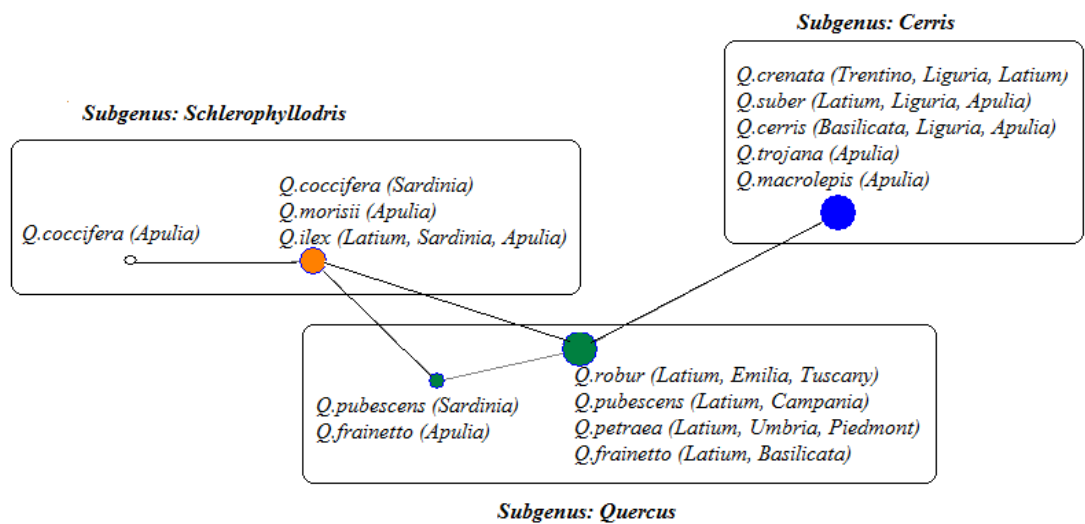
Region	rbcL	trnH-psbA	matK_Quercus	matK2_Quercus	trnH-psbA + matK2_Quercus
Alignment length	743	577	828	696	1273
Polymorphic sites	3	16	10	13	29
PICs*	2	15	9	12	27
Indels*	0	9(1-5bp)	0	2(1-3bp)	11(1-5)
Mean intraspecies K2P distance (range)	0.0003 (0-0.001)	0.0006 (0-0.004)	0.0004 (0-0.004)	0.0007 (0-0.003)	0.0008 (0-0.003)
Mean interspecies K2P distance (range)	0.001 (0-0.003)	0.0103 (0-0.03)	0.003 (0-0.011)	0.006 (0-0.016)	0.008 (0-0.019)
Species #	12	12	12	12	12
Total haplotypes #	5	8	4	8	11
Single haplotypes #	1	2	1	3	4
Samples #	30	30	30	30	30

Tab. 14 : Principali caratteristiche della variabilità dei marcatori nel campionamento
taxon-based

*PIC: Caratteri informativi per il criterio delle parsimonia;

* indels: numero e lunghezza.

Haplotype Distribution with *rbcL*



Haplotype distribution with *matK_Quercus*

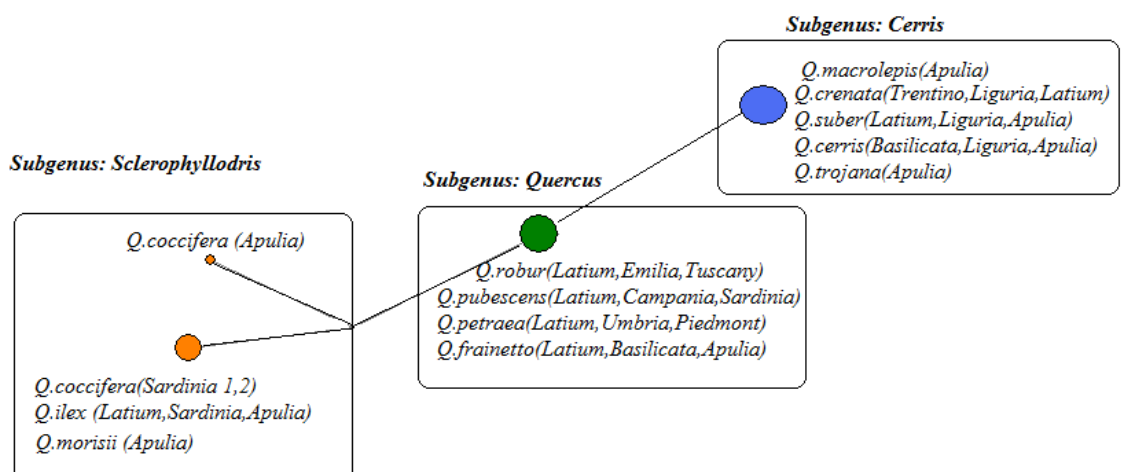


Fig.14a: Network relativi ai marcatori considerati singolarmente che individuano gli aplotipi generati con *rbcL* e *matK_Quercus*

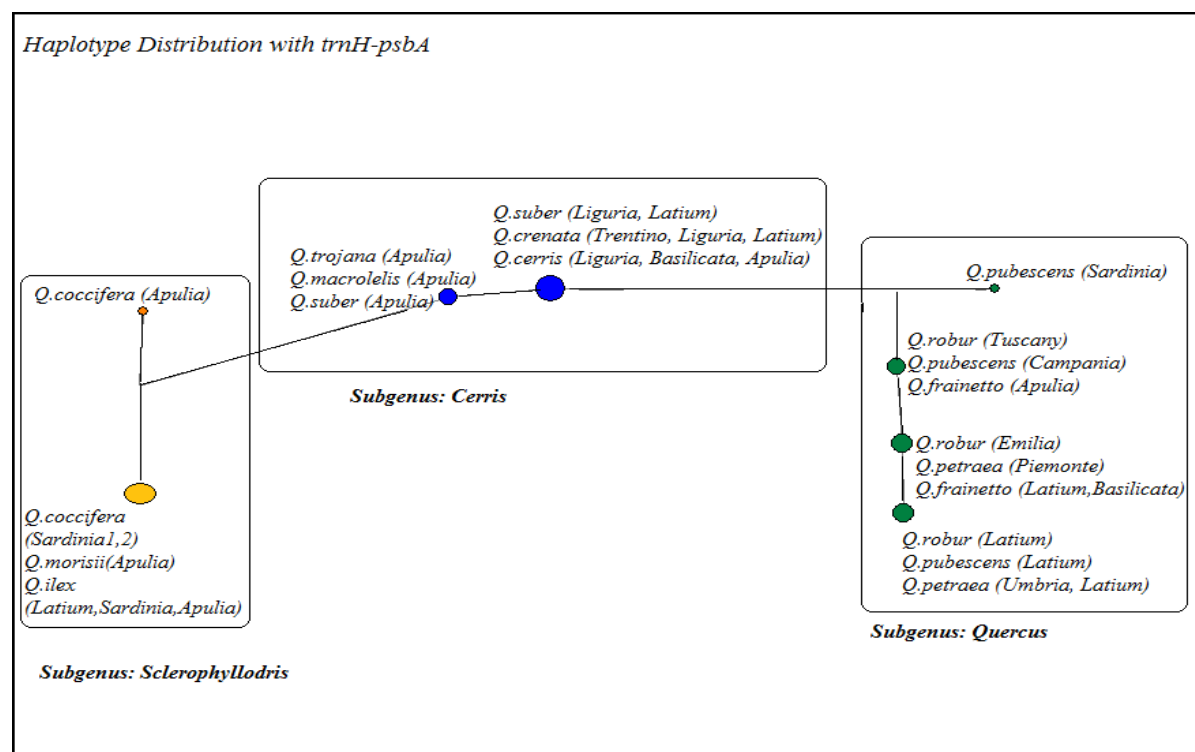
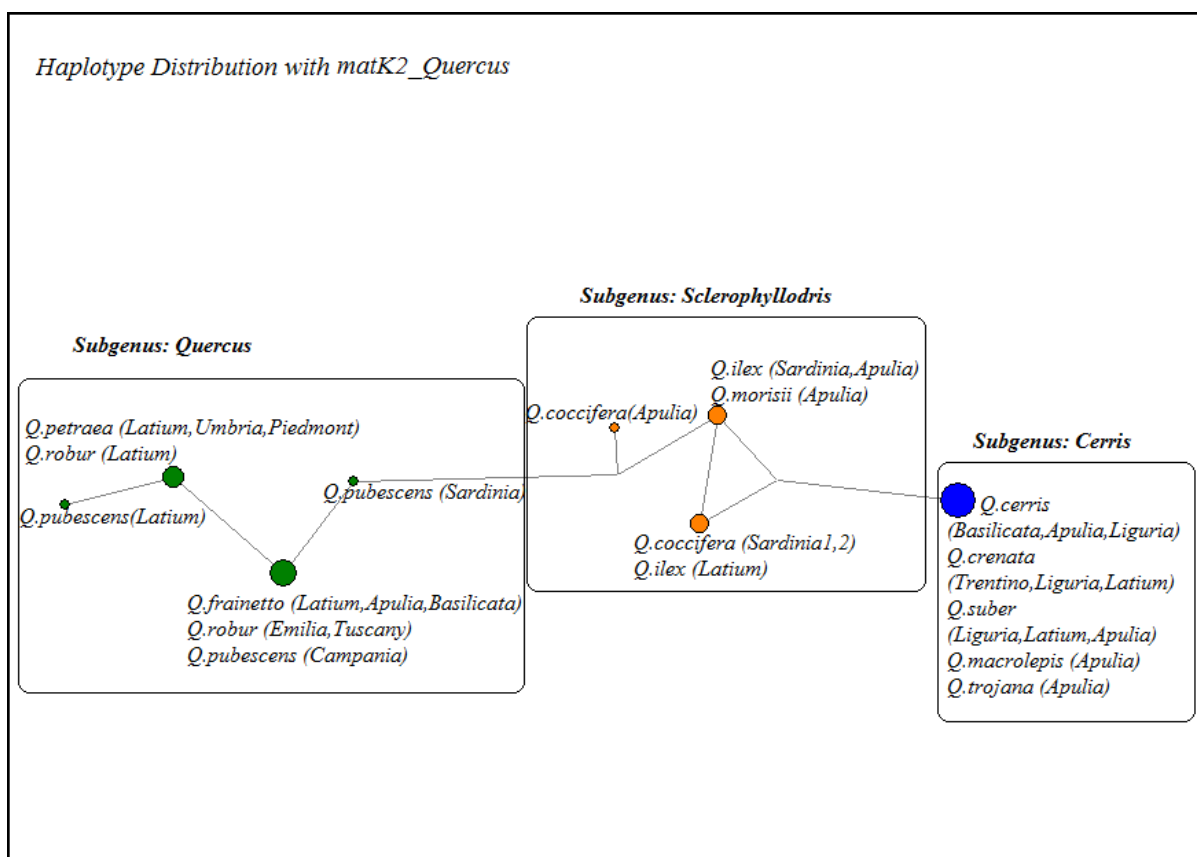


Fig.14b : Network relativi ai marcatori considerati singolarmente che individuano gli aplotipi generati con matK_Quercus e trnH-psbA

I quattro Network precedentemente riportati sono relativi ai marcatori considerati singolarmente.

In tutti i casi possiamo notare come specie appartenenti allo stesso sottogenere condividano la stessa sequenza (aplotipo); questa situazione comune a tutti i marcatori risulta più evidente in *rbcL* e *matK_Quercus*, dove all'interno del sottogenere le sequenze formano quasi un unico gruppo indifferenziato. Negli altri due marcatori *trnH-psbA* e *matK2_Quercus* sebbene all'interno dei sottogeneri si evidenzino più gruppi, questi corrispondono comunque a specie diverse che mostrano sequenza identica.

La combinazione dei due marcatori *trnH-psbA* e *matK2_Quercus* mostrava in tabella 14 un livello di variabilità piuttosto alto con 29 siti polimorfici, tuttavia sia l'analisi degli aplotipi e del Network, sia la produzione del dendrogramma hanno evidenziato caratteristiche analoghe a quelle evidenziate con i locus singoli, come visibile nelle figure 15 e 16. In particolare dall'osservazione del Network vediamo come il sottogenere *Cerris* sia separato in due gruppi di cui uno formato da *Q. trojana*- *Q. macrolepis*- *Q. suber* che condividono la stessa provenienza geografica (Puglia). Nel sottogenere *Scherophyllodris* solo la *Q. cocciifera* proveniente dalla Puglia viene riconosciuta come unica e non condivide la sequenza con altre specie, mentre nel sottogenere *Quercus* sono tre (le due *Q. pubescens* provenienti dal Lazio e dalla Sardegna e la *Q. petraea* del Piemonte) le specie identificate come singole.

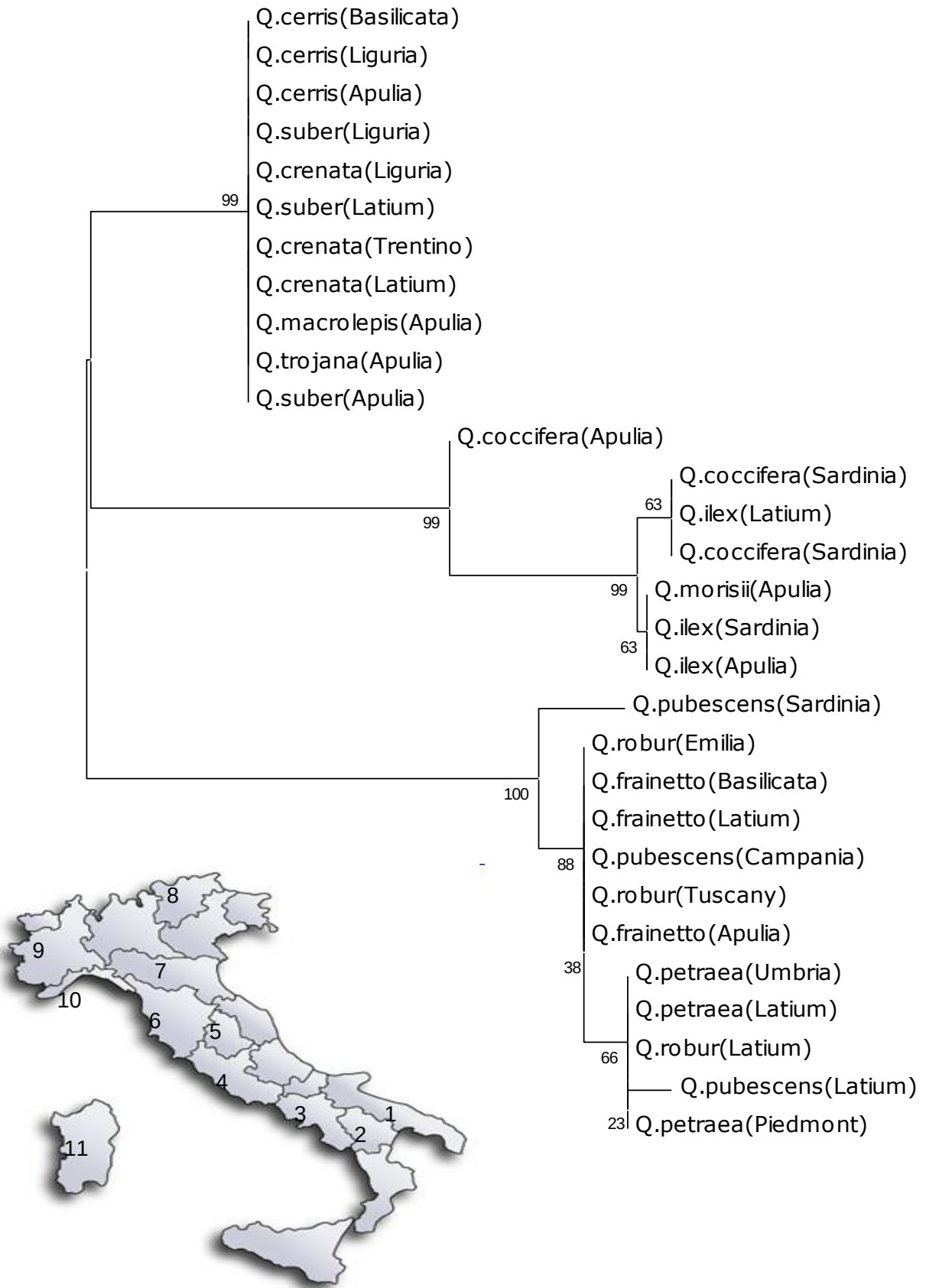


Fig.15: Dendrogramma ottenuto con i 2 locus trnH-psbA+matK2_Quercus concatenati.
La cartina mostra inoltre le regioni di provenienza del campionamento taxon-based

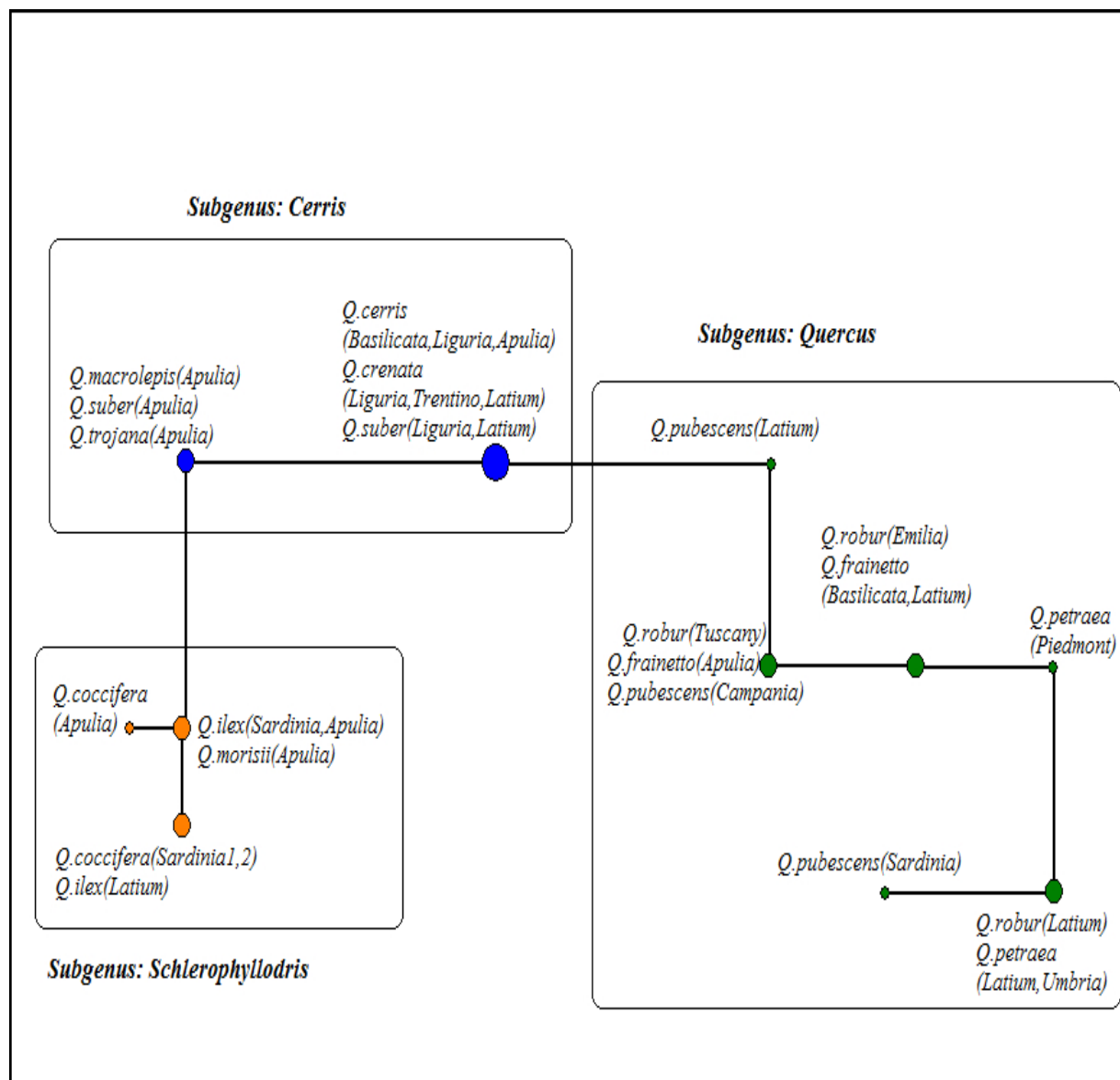


Fig. 16: Distribuzione aplotipi ottenuta con i 2 locus trnH-psbA+matK2_Quercus concatenati

6 DISCUSSIONE

Una delle principali caratteristiche della flora forestale italiana è quella di avere un numero relativamente basso di specie che tuttavia coprono un ampio range tassonomico. In particolare la flora italiana comprende circa 90 specie (il più alto numero in Europa) che appartengono a 40 generi e 22 famiglie (Gellini et al 1997). Circa la metà dei generi sono rappresentati da una singola specie (Cupressus, Picea, Larix, Fagus, Castanea, Corylus, Olea, etc.) e un quarto delle famiglie sono rappresentate da generi monospecifici (Taxaceae: Taxus baccata; Laricaceae: Larix decidua; Juglandaceae: Juglans regia; Platanaceae: Plantanus orientalis; Lauraceae: Laurus nobilis; Aquifoliaceae: Ilex aquifolius; Moraceae: Ficus carica). Dei rimanenti generi, quattro (Acer, Pinus, Prunus, Sorbus) possiedono un numero di specie compreso tra 5 e 9, infine solo due generi (Quercus, Salix) sono presenti con più di 10 specie.

Il campionamento floristico è stato effettuato quindi in modo da essere rappresentativo di tale situazione tipica del territorio italiano.

I risultati ottenuti nell'ambito dello studio sulla biodiversità della **dendroflora italiana**, hanno evidenziato come l'utilizzo di matK non sia consigliabile a causa delle basse percentuali di successo ottenuti nell'amplificazione, mentre i risultati ottenuti con la combinazione di una regione codificante, rbcL in particolare, e una non codificante trnH-psbA si sono dimostrati veramente incoraggianti. Infatti, escludendo il genere Quercus (il cui comportamento devia in maniera importante dal resto del dataset) le percentuali di successo di discriminazione si attestano a circa 82%, confermando quanto proposto da Kress (Kress et al 2007) che evidenzia come l'utilizzo dei due soli marcatori rbcL e trnH-psbA potrebbe essere sufficiente come barcode nelle piante. Questi dati possono essere considerati come un importante punto di partenza per studi di vario tipo, come ecologici, biodiversità generale, o certificazione in campo vivaistico, ma è necessario comunque sottolineare il bisogno di ulteriori campionamenti che vadano ad esplorare la variabilità intraspecie di tutti i taxon impiegati in questo studio. Solo in questo modo sarà possibile valutare in maniera certa le percentuali di successo nella discriminazione delle specie e le possibili problematiche ad esse collegate.

L'approccio floristico, sebbene non dettagliato, è comunque in grado di evidenziare situazioni che in qualche modo si discostano dal comportamento medio, sia in senso positivo sia in senso negativo, come nel caso del genere Quercus. Infatti, lo studio taxon-based intrapreso nella seconda parte ha evidenziato i limiti del metodo barcoding in questo genere.

La caratteristica dell'approccio taxon-based prevede come sappiamo un campionamento intraspecie dettagliato volto a studiare dinamiche più sottili che legano l'approccio barcode

alla genetica di popolazione. I dati ottenuti, soprattutto se derivanti da un approccio multi-locus, possono infatti costituire un primo segnale dei processi di microevoluzione relativi alle popolazioni in esame (Dasmahapatra et al. 2006 ; Hajibabaei et al. 2007).

In particolare, nel nostro caso, i campioni intraspecie provenivano dal territorio italiano dove spesso popolamenti di specie diverse condividono le stesse aree geografiche consentendo di verificare eventuali condizioni di ibridizzazione simpatica.

Lo studio relativo al genere *Quercus* attraverso i marcatori codificanti *rbcL* e *rpoC1* si è dimostrato molto difficile, infatti entrambi mostrano un tasso evolutivo molto lento capace solo di rivelarci relazioni ancestrali. In particolare *rpoC1*, di cui non sono riportati i dati nel dettaglio ma che è comunque stato esplorato parzialmente anche con campioni intraspecie, mostra un tasso evolutivo talmente basso che, come evidenziato nello studio floristico, fallisce nel discriminare *Castanea* a livello di genere. Il campione di *Castanea sativa* infatti possiede una sequenza identica al gruppo di campioni appartenenti al sottogenere *Cerris*; questo risultato sebbene da considerare negativo in un contesto strettamente correlato al DNA barcoding, ci fornisce la conferma di un antenore castanoideo per il genere *Quercus* come già evidenziato da altri autori con marcatori nucleari (Oh et Manos 2008).

Il locus *rbcL* ci consente di vedere qualcosa in più rispetto a *rpoC1*, e con i suoi 3 siti polimorfici riesce a separare il genere *Quercus* nei tre classici sottogeneri *Cerris*, *Quercus*, e *Sclerophyllodris*. Questo risultato, sebbene insufficiente se considerato isolato, fornisce invece una buona base nell'approccio multilocus necessario nel regno vegetale. Infatti questa regione codificante facilmente allineabile consentirebbe, all'interno di un database globale, un primo passaggio per arrivare a livello di sottogenere.

Il frammento *matK_Quercus*, (che corrisponde alla regione classica del CBOL per la quale è stato necessario lo sviluppo dei primers specifici per il genere *Quercus*) mostra, come visibile dalla tabella, una maggiore variabilità generale. Questa, tuttavia, non migliora la situazione e attraverso il Network possiamo osservare praticamente lo stesso successo di discriminazione che si ferma a livello di sottogenere.

Il frammento da noi sviluppato, *matK2_Quercus*, presenta una maggiore variabilità e osservando il Network vediamo come all'interno dei sottogeneri *Quercus* e *Sclerophyllodris*, sia stato in grado di generare una ulteriore separazione tra le specie, mentre in *Cerris* il gruppo resta compatto. Tale separazione però, non può essere considerata come un maggiore successo di discriminazione, poiché nei gruppi formati ci sono specie differenti che possiedono sequenza identica.

Osserviamo una situazione analoga sia con il marcatore non codificante trnH-psbA che riesce a separare anche il sottogenere *Cerris* in due gruppi differenti.

La combinazione dei due marcatori matK2-*Quercus* + trnH-psbA, pur possedendo buoni livelli di variabilità conferma lo stesso comportamento all'interno dei sottogeneri dove individui di specie differenti condividono lo stesso aplotipo. Tale condizione nota come **polifilia** è stata evidenziata in letteratura in campioni mitocondriali di specie animali con percentuali di circa il 20% (Funk et al 2003). Non esistono in letteratura studi paralleli su dati relativi al cloroplasto, tuttavia seguendo anch'esso una eredità di tipo uniparentale, è possibile ipotizzare un comportamento analogo. Le cause che possono essere alla base di tale fenomeno sono veramente molteplici e spesso non è possibile interpretarle in modo certo, tuttavia relativamente al metodo del DNA barcoding questa condizione non consente una identificazione univoca della specie.

La polifilia può essere dovuta ad una non adeguata informazione di partenza che può generare artefatti filogenetici: la base potrebbe essere un frammento troppo corto proveniente da un gene che evolve in maniera troppo lenta o una regione non codificante con un elevato tasso di mutazione che giunto a saturazione mostra elevata omoplasia. Tuttavia è stato osservato che nella maggior parte casi non sono questi artefatti la causa prevalente della polifilia riscontrata.

Un'importante causa risulta invece essere quella legata alla tassonomia corrente che potrebbe aver fallito nell'identificare i limiti genetici delle entità biologiche. Questa non corretta delimitazione della specie può essere generata per due diversi motivi: in un primo caso la tassonomia potrebbe aver sottostimato l'ampiezza della specie, e quelle che erano delle semplici varianti (anatomiche, ecologiche o comportamentali) interne alla specie sono state invece ascritte come specie diverse (sottostima della variabilità intraspecie-> genera eccesso di false specie); nel secondo caso viceversa, tratti e varianti sono stati valutati erroneamente come non significativi per essere attribuiti ad una nuova specie e raggruppate sotto uno stesso nome come nel caso delle specie criptiche (sovrastima della variabilità intraspecie-> genera un numero inferiore di specie).

Anche l'ibridizzazione interspecie, cioè l'occasionale incrocio tra specie diverse, può generare polifilia. La principale conseguenza di tale fenomeno è l'introggressione (detta anche ibridizzazione introggressiva, introggressione, flusso genico interspecie), ossia alleli di una specie possono penetrare nel pool genico di un'altra a seguito di un primo incrocio tra le specie seguito dal successivo reincrocio degli ibridi con la popolazione parentale. Gli effetti di tale fenomeno sono particolarmente evidenti nei genomi ad eredità uniparentale dove l'assenza di ricombinazione e la trasmissione "in blocco" della molecola amplificano tale segnale.

L'introggressione potrebbe essere in parte riconosciuta attraverso la valutazione di un albero costruito con marcatori cloroplastici contrapposto ad uno proveniente da dati nucleari che evidenzia in maniera corretta e consistente la specificità dei taxon. Un ulteriore segno di introggressione è la condivisione simpatrica (localizzata geograficamente) di aplotipi di specie diverse, che invece se campionate da aree geografiche lontane potrebbero invece mostrare monofilia.

Tuttavia, sfortunatamente, la polifilia dovuta a introggressione diventa più difficile da riconoscere tanto più lontano nel tempo sia avvenuto il flusso genico, poiché anche specie che hanno recentemente cessato tale flusso, possono aver acquisito nuove mutazioni post-introggressione, generando aplotipi comunque molto correlati e filogeneticamente annidati. Comunque se l'ultimo evento di flusso genico è avvenuto lontano nel tempo gli aplotipi introgressi che ancora rimangono saranno con molta probabilità filogeneticamente basali con un successivo smistamento dei polimorfismi a partire dall'evento di introggressione, e inoltre sarà meno probabile riconoscere una associazione geografica delle popolazioni introgresse.

Ci sono anche altri fattori che possono complicare la situazione come per esempio le situazioni in cui il flusso riguarda più specie o gli aplotipi introgressi si fissano non lasciando traccia della loro origine eterospecifica. Questo processo può essere facilitato quando il fenomeno di introggressione avviene in una piccola popolazione.

Infine anche l'evento "incomplete lineage sorting" rappresenta una possibile causa di polifilia. La casuale distribuzione degli alleli del progenitore al momento della speciazione avrà come risultato che alcuni individui di una specie figlia possiederanno alcuni alleli che saranno più simili a quelli di un'altra specie piuttosto che alla propria. Questa condizione è naturale nelle nuove specie che spesso appaiono essere filogeneticamente polifiletiche, ed è destinata a portare a monofilia quando, nel corso del tempo per deriva si perderanno alcuni vecchi alleli e nuovi alleli si formeranno in modo che tutti gli individui intraspecie rifletteranno la situazione post-speciazione. La progressione temporale di questo processo è ovviamente funzione della taglia effettiva della popolazione (N_e), e poiché è noto che i genomi aploidi e ad eredità uniparentale hanno una taglia che è un quarto rispetto al nucleare, è atteso che siano i loci nucleari quelli più affetti da questo problema. Tuttavia, nel caso di taxa che si siano evoluti molto rapidamente è possibile che anche i genomi mitocondriali e cloroplastici possano presentare questo problema.

Purtroppo, sebbene siano stati elaborati alcuni metodi, non esiste un metodo che sia universalmente accettato per distinguere l'introggressione da "incomplete sorting".

Proprio questa condizione di polifilia appare come il risultato più evidente dello studio taxon-based in *Quercus*, ma come intuibile da quanto detto sopra, riconoscere le cause che la hanno generata risulta essere praticamente non possibile. Tuttavia osservando il Network generato dall'unione dei due marcatori (Figura 15), la distribuzione degli aplotipi condivisi all'interno del sottogenere, sembra riflettere una localizzazione geografica delle specie, il che potrebbe indicare la presenza di ibridizzazione e conseguente introgressione.

Questa condizione ci riporta anche al concetto biologico di specie e al fatto che interfecondità e isolamento riproduttivo non sono condizioni esclusive del tipo tutto o nulla, tanto che alcuni specialisti pur sostenendo solidamente il concetto biologico di specie, non insistono sul fatto che le popolazioni debbano essere isolate al 100% per poter essere qualificate come specie, ma riconoscono che ci possa essere qualche piccola infiltrazione genetica tra le specie soprattutto se strettamente imparentate e simpatriche. L'ibridazione simpatica, in particolare, sembra essere più comune nel mondo vegetale che negli animali, tanto che diversi botanici sono stati a lungo riluttanti ad adottare il concetto biologico di specie.

E' stato inoltre ipotizzato che l'aumento del tasso di ibridizzazione possa essere una risposta genetica naturale delle piante nelle situazioni di rapidi cambiamenti ambientali che siano indotti o meno dall'uomo (Petit et al. 2004a). In particolare tali perturbazioni potrebbero essere state molto forti nelle zone boschive dell'area Mediterranea a causa della loro gestione e utilizzo estremo (pascolo eccessivo, tagli intensivi, dissodamenti, incendi) e quindi aver favorito lo sviluppo di questa risposta biologica così marcata in sottogruppi di specie simpatriche.

Tuttavia, come già detto nell'introduzione, anche la tassonomia del genere risulta essere problematica e controversa, tanto da non poter escludere che in alcuni casi i limiti di specie possano essere stati definiti in maniera non corretta.

Perciò i risultati ottenuti, pur fornendo importanti informazioni e conferme sulle complesse relazioni evolutive e filogeografiche relative a questo genere, determinano un fallimento nell'applicazione del metodo barcode in senso stretto. Infatti, poiché il barcode utilizza come riferimento la tradizionale tassonomia la cui concezione di specie risulta essere come sappiamo molto statica e chiusa, non siamo in grado di discriminare con successo al di sotto del sottogenere. Quindi il risultato finale dell'approccio Barcode applicato al genere *Quercus* attraverso marcatori cloroplastici è indiscutibilmente negativo.

Restano tuttavia due possibili considerazioni. La prima riguarda la possibilità di esplorare il genere *Quercus* con marcatori nucleari come il già citato I.T.S., o altri marcatori neutrali, come del resto previsto dal CBOL e dal WPG per quei gruppi tassonomici, come appare essere

Quercus, particolarmente problematici. Tale esplorazione è in realtà già stata avviata dal nostro gruppo, ma i risultati sono ancora in fase di elaborazione.

La seconda considerazione riguarda aspetti più filosofici legati al DNA barcoding.

La nascita di questo nuovo approccio è stata non semplice e come detto nell'introduzione sono state molte le polemiche sollevate da più parti, ma come del resto è giusto che sia in un dibattito scientifico sano e libero; nonostante ciò a distanza di circa dieci anni, il progetto Barcode è ancora attivo, anzi in crescita, e potremmo dire “vincente” sugli scettici.

Risulta del tutto comprensibile come fosse necessario per il DNA barcoding in una fase iniziale appoggiarsi ad una struttura solida come quella della tassonomia classica che, sebbene fosse la branca da cui provenivano la maggior parte degli oppositori, costituiva un punto fermo molto forte rispetto a tutti i nuovi protocolli implementati nel metodo. Possiamo anche aggiungere che la classica tassonomia ha costituito e costituisce senza dubbio un ottimo punto di partenza per lo studio della biodiversità, tuttavia forse è possibile pensare, anche grazie ai dati molecolari raccolti in questi anni anche attraverso i loci barcode, ad una visione tassonomica più plastica e dinamica e sicuramente più conforme alla realtà naturale, e quindi ipotizzare una fase futura del progetto Barcode in cui il ruolo centrale non sia attribuito alla tassonomia classica. Infatti è probabile l'utilizzo di nuova tassonomia, la cosiddetta *DNA taxonomy*, che attraverso l'integrazione dei dati molecolari allo scheletro della corrente tassonomia, potrebbe consentire, soprattutto relativamente a gruppi particolari, una nuova delimitazione delle specie forse più sensibile nella comprensione dell'evoluzione della biodiversità.

BIBLIOGRAFIA

- Aldrich J. B. W. Cherney E. Merlin L. Christopherson 1988 The role of insertions/deletions in the evolution of the intergenic region between *psbA* and *trnH* in the chloroplast genome. *Current Genetics* 14: 137-146
- Azuma H. J. G. García-Franco V. Rico-Gray L. B. Thien 2001 Molecular phylogeny of the Magnoliaceae: the biogeography of tropical and temperate disjunctions. *American Journal of Botany* 88: 2275-2285
- CBOL Plant Working Group (2009) A DNA barcode for land plants. Proceedings of the National Academy of Sciences of the USA, 106, 12794–12797.
- Baldauf, S.L. and Palmer, J.D. (1990) Evolutionary transfer of the chloroplast *tufA* gene to the nucleus. *Nature*. 344:262-265
- Barbrook, A.C., Lockhart, P.J. and Howe, C.J. (1998) Phylogenetic analysis of plastid origins based on *secA* sequences. *Current Genetics*. 34:336-341
- Burger WC (1975) The species concept in *Quercus*. *Taxon*, 24, 45–50.
- CBOL Plant Working Group (2009) [A DNA barcode for land plants](#). Proceedings of the National Academy of Sciences, USA, **106**: 12794-12797
- Coyne J. A. , Orr H. A. (2004) : Speciation. Sinauer associates
- Curtis, S. and Clegg, M.T. (1984) Molecular evolution of chloroplast DNA sequences. *Mol. Biol. Evol.* 1:291-301
- Dasmahapatra KK, Mallet J (2006). DNA barcodes: recent successes and future prospects. *J Heredity* 97: 254-255.
- Delwiche, C.F., Kuhsel, M. and Palmer, J.D. (1995) Phylogenetic analysis of *tufA* sequences indicates a cyanobacterial origin of all plastids. *Molecular Phylogenetics and Evolution*. 4:110-128
- Downie S.R., E. Llanas and D.S. Katz-Downie, Multiple independent losses of the *rpoC1* intron in angiosperm chloroplast DNA's, *Syst. Bot.* **21** (1996), pp. 135–151
- Downie S.R., R.G. Olmstead, G. Zurawski, D.E. Soltis, P.S. Soltis, J.C. Watson and J.D. Palmer, Six independent losses of the chloroplast DNA *rpl2* intron in dicotyledons: molecular and phylogenetic implications, *Evolution* **45** (1991), pp. 1245–1259
- Ebach, M. C., and C. Holdrege. 2005. DNA barcoding is no substitute for taxonomy. *Nature* 434:697.
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 19, 1792–1797.
- Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG, Husband BC, Percy DM, Hajibabaei M, and Barrett SCH (2008). Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS ONE* 3(7): e2802.
- Funk DJ, Omland KE (2003) Species-level monophyly and paraphyly: frequency, causes and consequences with insights from animal mitochondrial DNA. *Annual Review of Ecology Evolution and Systematics*, 34, 397–423.
- Futuyma Douglas J. (2005): Evolution. Sinauer Associates
- Galtier N, Gouy M, Gautier C (1996) SeaView and Phylo_win, two graphic tools for sequence alignment and molecular phylogeny. *Computer Applications in the Biosciences*, 12, 543–548.
- Gantt, J.S., Baldauf, S.L. Calie, P.J., Weeden, N.F. and Palmer, J.D. (1991) Transfer of *rpl22* to the nucleus greatly preceded its loss from the chloroplast and involved the gain of an intron. *EMBO J.* 10:3073-3078
- Gellini R, Grossoni P (1997) Botanica Forestale. CEDAM, Padova, Italy.

- Hamilton M. B. 1999b Tropical tree gene flow and seed dispersal: deforestation affects the genetic structure of the surviving forest fragments. *Nature* 401: 129-130
- Hamilton M. B. J. M. Braverman D. F. Soria-Hernanz 2003 Patterns and relative rates of nucleotide and insertion/deletion evolution at six chloroplast intergenic regions in New World species of the Lecythidaceae. *Molecular Biology and Evolution* 20: 1710-1721
- Hajibabaei M, Singer GAC, Hebert PDN, Hickey DA (2007) DNA barcoding: how it complements taxonomy, molecular phylogenetics and populations genetics. *Trends in Genetics*, 23, 167–172.
- Hausner *et al.*, Origin and evolution of the chloroplast *trnK* (*matK*) intron: a model for evolution of group II intron RNA structures, *Mol. Biol. Evol.* **23** (2006), pp. 380–391.
- Hebert PDN, A Cywinska, SL Ball and JR deWaard. 2003a. Biological identifications through DNA barcodes. *Proc. R. Soc. B* 270: 313-321
- Hebert PDN, S Ratnasingham and JR deWaard. 2003b. Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. B* 270: 596-599.
- Hebert PDN, Stoeckle MY, Zemplak TS, Francis CM (2004a). Identification of birds through DNA barcodes. *PLoS Biol* 2: e312
- Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004b). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *PNAS* 101: 14812-14817.
- Hebert PDN and TR Gregory. 2005. The promise of DNA barcoding for taxonomy. *System. Biol.* 54: 852-859
- Holdregger R. R. J. Abbott 2003 Phylogeography of the Arctic-Alpine *Saxifraga oppositifolia* (Saxifragaceae) and some related taxa based on cpDNA and ITS sequence variation. *American Journal of Botany* 90: 931-936
- Hudson G.S., T.A. Holton, P.R. Whitfield and W. Bottomley, Spinach chloroplast *rpoBC* genes encode three subunits of the chloroplast RNA polymerase, *J. Mol. Biol.* **200** (1988), pp. 639–654
- Hilu, K. W., and H. Liang. 1997. The *matK* gene: sequence variation and application in plant systematics. *Am. J. Bot.* 84:830–839.
- Hollingsworth ML, Clark A, Forrest LL, Richardson JE, Pennington RT, Long D, Cowan R, Chase MW, Gaudeul M, Hollingsworth PM (2009) [Selecting barcoding loci for plants: evaluation of seven candidate loci with species level sampling in three divergent groups of land plants](#) *Molecular Ecology Resources* **9**: 439-457
- Feliner GN, Rossello JA (2007) Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Molecular Phylogenetics and Evolution*, 44, 911–919.
- Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* 3, 294-297.
- Kapralov MV, Filatov DA. 2007. Widespread positive selection in the photosynthetic Rubisco enzyme. *BMC Evol. Biol.*;7:73
- Katayama H. and Y. Ogihara, Structural alterations of the chloroplast genome found in grasses are not common in monocots, *Curr. Genet.* **23** (1993), pp. 160–165.
- Kelchner, S. A. 2002. Group II introns as phylogenetic tools: structure, function, and evolutionary constraints. *Am. J. Bot.* 89:1651–1669.

- Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005) Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences, USA*, 102, 8369–8374
- Kress WJ, Erickson DL (2007). A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS ONE* 6: 1-10.
- Kress J, Erickson DL (2009) Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. *Proceeding of the National Academy of Sciences, USA* , **106**:18621-18626
- Martin, W. and Hermann, R.G. (1998) Gene transfer from organelles to the nucleus: how much, what happens and why? *Plant Physiology*. 118:9-17
- Mayr Ernst: Animal species and evolution. Harvard University Press, Cambridge 1963
- Mayr Ernst :Systematic and the origin of species. Coloumbia University Press, New York, 1942
- Meier, R. G. Zhang*, F. Ali*. The use of mean instead of smallest interspecific distances exaggerates the size of the 'barcoding gap' and leads to misidentification. **Systematic Biology**, **57**: 809-813.
- Meyer CP, Paulay G (2005). DNA barcoding: error rates based on comprehensive sampling. *PLoS Biol* 3: e422.
- Mower J, Touzet P, Gummow J, Delph L, Palmer J (2007) Extensive variation in synonymous substitution rates in mitochondrial genes of seed plants. *BMC Evolutionary Biology*, 7, 135.
- Newmaster SG, Fazekas AJ, Ragupathy S (2006). DNA barcoding in land plants: evaluation of *rbcL* in a multigene tiered approach. *Can J Bot* 84: 335-341.
- Oh Sang-Hun , Manos Paul S. (2008). Molecular phylogenetics and cupule evolution in Fagaceae as inferred from nuclear CRABS CLAW sequences. *Taxon* vol. 57, n°2, pp. 434-451
- Palmer, J. D. (1990) Contrasting modes and tempos of genome evolution in land plant organelles. *Trends Genet.* 6, 115–120.
- Palmer, J. D. (1991) Plastid chromosomes: structure and evolution; in *Cell Culture and Somatic Cell Genetics in Plants*, Vol.7A, The Molecular Biology of Plastids, Vasil, I. K. and Bogorad, L. (eds.), pp. 5–53, Academic Press, San Diego.
- Pennisi E (2007). Taxonomy. Wanted: a barcode for plants. *Science* 318: 190–191.
- Race, H., Hermann, R.G. and Martin, W. (1999) Why have organelles retained genomes? *TIG*. 15:364-370
- Rach J, De Salle R, Sarkar IN, Schierwater B, Hadrys H (2008). Character-based DNA barcoding allows discrimination of genera, species and populations in Odonata. *Proc R Soc B* 275: 237-247.
- Renner S. S. 1999 Circumscription and phylogeny of the Laurales: evidence from molecular and morphological data. *American Journal of Botany* 86: 1301-1315
- Rubinoff, D., S. Cameron, and K. Will. 2006. Are plant DNA barcodes a search for the Holy Grail? *Trends in Ecology and Evolution*.21:1-2.
- San Filippo, J., and A. M. Lambowitz. 2002. Characterization of the C-terminal DNA-binding/DNA endonuclease region of a group II intron-encoded protein. *J. Mol. Biol.* 324:933–951.
- Sang T. D. J. Crawford T. F. Stuessy 1997 Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* 84: 1120-1136
- Shearer TL, Coffroth MA (2008) Barcoding corals: limited by interspecific divergence, not intraspecific variation. *Molecular Ecology Resources*, 8, 247–255.

Soltis D. E. R. K. Kuzoff M. E. Mort M. Zanis M. Fishbein L. Hufford J. Koontz M. K. Orroyo 2001 Elucidating deep-level phylogenetic relationships in Saxifragaceae using sequences for six chloroplastic and nuclear DNA regions. *Annals of the Missouri Botanical Garden* 88: 669-693

Sugiura, M. (1989) The chloroplast chromosomes in land plants. *Annu. Rev. Cell Biol.* 5, 51–70.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, 24, 1596–1599.

Turmel, M., Otis, C. and Lemieux, C. (1999) The complete chloroplast DNA sequence of the green alga *Nephroselmis olivacea*: insights into the architecture of ancestral chloroplast genomes. *PNAS USA*. 96:10248-10253

Wakasugi, T., Tsudzuki, J., Ito, S., Nakashima, K., Tsudzuki, T., et al. (1994) Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. *Proc. Natl. Acad. Sci. USA* 91, 9794–9798.

Wallace R.S. and J.H. Cota, An intron loss in the chloroplast gene *rpoC1* supports a monophyletic origin for the subfamily Cactoideae of the Cactaceae, *Curr. Genet.* **29** (1996), pp. 275–281

Wheeler, Q. D. 2004. Taxonomic triage and the poverty of phylogeny. *Phil. Trans. R. Soc. Lond. B* 359:571–583.

Whitworth TL, Dawson RD, Magalon H, Baudry E (2007) DNA barcoding cannot reliably identify species of the blowfly genus *Protocalliphora* (Diptera: Calliphoridae). *Proceedings of the Royal Society B: Biological Sciences*, 274, 1731–1739.

Will, K. W., B. D. Mishler, and Q. D. Wheeler. 2005. The perils of DNA barcoding and the need for integrative taxonomy. *Syst. Biol.* 54:844–851.

Will, K. W., and D. Rubinoff. 2004. Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics* 20:47–55.

Woese CR Fox GE (1977). Phylogenetic Structure of the prokaryotic domain: the primary kingdoms. *PNAS* 74(11): 5088-5090.

.

Ringraziamenti

Arrivare a questa pagina significa aver finito, finalmente, la stesura della tesi ma anche chiudere un periodo della propria vita.

In questi anni ci sono state tante persone con cui ho interagito e con le quali ho condiviso molto.

Per prima cosa vorrei ringraziare la prof. Attimonelli e il prof. Schirone, che hanno reso possibile questo bellissimo progetto di ricerca e che in modi diversi hanno contribuito alla mia formazione.

Poi Marco Simeone, brillante ricercatore che, sebbene fosse a chilometri di distanza, ho sempre sentito molto vicino sia come collega che come amico.

Ringrazio i miei genitori, a cui ho dedicato questo mio lavoro, che sempre mi hanno sostenuto nelle scelte che ho fatto e come un porto sicuro mi aspettano sempre a braccia aperte.

Mia sorella Teresa che alcuni anni fa mi ha dato lo stimolo a riprendere la mia vita di biologa.

Ringrazio poi il mio compagno Teodoro per tutti i chilometri fatti per stare insieme e i miei figli Olga e Angelo che mi hanno accompagnato in questa avventura fuori sede.

Ora passiamo ai colleghi di ogni ordine e grado che hanno condiviso con me tutta la polvere e le formiche della nostra stanza: grazie a tutti, ognuno di voi mi ha dato qualcosa che va molto oltre i tanti multiallineamenti condivisi.....

Infine un pensiero a chi non c'è più, ma sono certa mi protegge dal cielo.