



UNIVERSITA' degli STUDI della TUSCIA di VITERBO
Dipartimento di Produzione Vegetale
Dottorato in
"ORTOFLOROFRUTTICOLTURA" XVIII ciclo (AGR 07)

**APPLICAZIONE DI TECNICHE MOLECOLARI AVANZATE NEL
MIGLIORAMENTO GENETICO DEL PESCO [*Prunus persica* (L.)
BATSCH] PER I CARATTERI LEGATI ALLA QUALITA' DEL
FRUTTO.**

Dottoranda:

Dott.ssa Elisa Vendramin

Tutor:

Dott.ssa Roberta Quarta

Coordinatore:

Chiar.mo Prof. Francesco Saccardo

A Michele per il suo amore ed il suo sostegno.

“...et pomis, quae barbara Persis miserat, ut fama est, patriis armata venenis. At nunc expositi parvo discrimine leti. Ambrosios praebent sucos, oblita nocendi. Quin etiam eiusdem gentis de nomine dicta exiguo properant mitescere. Persica malo.”

Columella, De Re Rustica

RINGRAZIAMENTI

Durante questi tre anni tante persone mi hanno aiutata e sostenuta, sia professionalmente che umanamente. E' quindi per me un piacere, più che un obbligo, ringraziare tutti coloro che mi sono stati accanto.

Il mio primo pensiero va alla dott.ssa Quarta, Roberta, che mi ha dato la grande possibilità di svolgere questo dottorato, mi ha sempre aiutata e sostenuta con fiducia ed affetto.

Non posso non ringraziare il fantastico duo delle mappe, la dott.ssa Dettori, Mary ed il dott. Verde, Iggy, che con il loro acume e spirito critico mi hanno sempre stimolata a pormi nuove domande.

Un grazie lo devo anche al direttore dell'Istituto Sperimentale per la Frutticoltura, Prof. Fideghelli, ed al coordinatore del dottorato, Prof. Saccardo; per la fiducia accordatami.

Un ringraziamento speciale va alla mia amica e collega Sabrina per aver diviso con me il suo tempo, la sua conoscenza e le lunghe giornate in laboratorio. Grazie anche a Jessica, per il suo aiuto e la sua disponibilità

Vorrei ringraziare tutte le mie "amiche di provetta", per avermi "sopportata" durante questo periodo: Alessia, Claudia, Daniela, Manuela, Maria Grazia, Melania, Sonia e Simona.

Non posso non ringraziare i miei genitori che, ancora una volta, mi sono stati vicino e che continuano a credere in me incondizionatamente. Un grazie a tutta la mia famiglia, che pur lontana mi è sempre vicino. Vorrei anche ringraziare la mia nuova famiglia, Paola, Pietro, Francesco e Sandro, per il loro affetto.

Un pensiero va anche ai miei nonni che, pur non essendo fisicamente presenti, mi hanno lasciato in dono il loro amore per la conoscenza che mi ha sempre guidato in questi anni.

Infine un grazie a tutti coloro che ho incontrato lungo il cammino.

INTRODUZIONE

1	<u>IL PESCO.....</u>	3
2	<u>I MARCATORI MOLECOLARI.....</u>	6
2.1	MARCATORI MOLECOLARI DERIVATI DAL DNA.....	7
2.1.1	<i>Marcatori basati sulle tecniche di Ibridazione</i>	7
2.1.2	<i>Marcatori basati sulla reazione a catena della polimerasi (PCR)</i>	10
2.2	MARCATORI MOLECOLARI DERIVATI DAL TRASCRITTOMA E DALLE REGIONI ESPRESSE.....	16
2.2.1	<i>cDNA o mRNA come fonte di Marcatori</i>	17
2.2.2	<i>Banche dati EST come fonte di Marcatori</i>	18
2.2.3	<i>Sequenze di Geni noti come fonte di Marcatori</i>	24
3	<u>LE MAPPE D'ASSOCIAZIONE NEL GENERE PRUNUS</u>	27
4	<u>MATURAZIONE E CONSISTENZA DELLA POLPA.....</u>	33
4.1	L'ETILENE E LA REGOLAZIONE GENICA DELLA MATURAZIONE	33
4.1.1	<i>Recettori per l'Etilene e traduzione del segnale durante la maturazione.....</i>	36
4.2	GENI COINVOLTI NELL'AMMORBIDIMENTO DELLA POLPA	37
4.3	PESCHE MELTING, NON-MELTING E STONYHARD.....	38
5	<u>SCOPO DEL LAVORO.....</u>	40
6	<u>MATERIALI E METODI</u>	42
6.1	COSTRUZIONE MAPPA D'ASSOCIAZIONE	42
6.1.1	<i>Materiali Vegetali.....</i>	42
6.1.2	<i>Caratteri Morfologici</i>	42
6.1.3	<i>Marcatori Microsatelliti (Simple Sequence Repeats - SSRs)</i>	43
6.1.4	<i>Amplified fragment length polymorphisms (AFLPs).....</i>	44
6.1.5	<i>Analisi dei Dati.....</i>	46
6.2	BULKED SEGREGANT ANALYSIS (BSA)	46
6.2.1	<i>Materiali Vegetali e Costituzione dei Bulk.....</i>	47
6.2.2	<i>Random Amplified Polymorphic DNA (RAPDs)</i>	47
6.2.3	<i>Amplified Fragment Length Polymorphisms (AFLPs)</i>	48
6.3	EXPRESSED SEQUENCE TAGS (ESTs) DA FRUTTO DELLA CV YUMYEONG	49
6.3.1	<i>Estrazione RNA</i>	49
6.3.2	<i>Costruzione della cDNA Library.....</i>	50
6.3.3	<i>Sequenziamento Cloni cDNA Library</i>	52
6.3.4	<i>Annotazione delle sequenze.....</i>	53
6.3.5	<i>Caratterizzazione Marcatori Molecolari Microsatelliti derivati dall' EST</i>	55
6.4	ANALISI DELL'ESPRESSIONE DELLE ENDOPOLIGALACTURONASI (ENDOPG) E DELL' 1-AMINOCICLOPROPANO-1-CARBOSSILATO OSSIDASI (ACO).....	56
6.4.1	<i>Estrazione RNA Mesocarpo frutto Stadio S4</i>	56

6.4.2	<i>Analisi Northern</i>	57
6.4.3	<i>Ibridazione</i>	57
6.4.4	<i>ReverseTrascriptase-PCR</i>	59
7	<u>RISULTATI</u>	60
7.1	COSTRUZIONE DELLA MAPPA D'ASSOCIAZIONE O'H x Y	60
7.1.1	<i>Marcatori Microsatelliti</i>	60
7.1.2	<i>Marcatori AFLP (Amplified Fragment Length Polymorphisms)</i>	62
7.1.3	<i>Mappa di Associazione</i>	64
7.2	BULKED SEGREGANT ANALYSIS PER L'INDIVIDUAZIONE DI UN MARCATORE ASSOCIATO AL CARATTERE 'STONY HARD'	68
7.3	INDIVIDUAZIONE DI EST CORRELATE ALLA QUALITÀ DEL FRUTTO	70
7.3.1	<i>Individuazione e caratterizzazione di EST-SSR</i>	72
7.3.2	<i>Analisi dell'Espressione delle Endopoligacturonasi (EndoPG) e dell' 1-aminociclopropano-1-carbossilato ossidasi (ACO)</i>	77
8	<u>DISCUSSIONE</u>	79
8.1	COSTRUZIONE DELLA MAPPA D'ASSOCIAZIONE Y x O'H	79
8.1.1	<i>Marcatori Molecolari</i>	79
8.1.2	<i>Mappa D'Associazione</i>	80
8.2	BULKED SEGRAGANT ANALYSIS PER L'INDIVIDUAZIONE DI UN MARCATORE ASSOCIATO AL CARATTERE 'STONY HARD'	81
8.3	INDIVIDUAZIONE DI EST CORRELATE ALLA QUALITÀ DEL FRUTTO	82
8.3.1	<i>Assemblaggio delle sequenze ed individuazione di EST-SSR</i>	82
8.3.2	<i>La consistenza della polpa: melting, nonmelting e stony hard</i>	83
9	<u>CONCLUSIONI</u>	85
10	<u>BIBLIOGRAFIA</u>	87

INTRODUZIONE

1 IL PESCO

Poche specie da frutto si sono diffuse così rapidamente nel mondo adattandosi a molteplici ambienti come il pesco [*Prunus persica* (L.) Batsch] diventando una delle colture più importanti, la cui produzione mondiale raggiunge i 15 milioni di tonnellate all'anno superata solo da quella del melo (62 milioni di tonnellate) (FAOSTAT data, 2005).

Il pesco è una specie autogama, appartiene alla famiglia delle *Rosacee*, genere delle *Prunoidee* e sottogenere *Amygdales*.

Il centro d'origine del pesco è la Cina dove veniva utilizzato come albero ornamentale, lasciato crescere allo stato selvatico o semicultivato. Alcuni esploratori americani dei primi del '900 (Hedrick 1917) riportano che fra i peschi selvatici era possibile identificare tutti i fenotipi che poi, nel tempo, sono stati oggetto del miglioramento genetico: pesche a polpa gialla, bianca, rossa, aderenti e spicche (comune nelle piante selvatiche). Questa eterogeneità delle forme ne conferma l'origine cinese secondo il principio di Vavilov per il quale la maggior variabilità all'interno di una specie si riscontra o presso o vicino il centro d'origine della specie stessa.

Un'altra importante regione di speciazione è il Turkestan (Kazakistan, Uzbekistan, Turkmenistan, Tagikistan e Kirghizistan, dell'Afghanistan e Cina) dove la presenza del pesco è stata documentata molto prima che questa coltura fosse nota in Europa Occidentale.

Il pesco arrivò nel bacino del Mediterraneo dalla Persia come documentano scrittori greci e romani, e come indica il nome stesso. I greci lo introducono come coltura fra il 400 ed il 300 A.C. e, come riportato da Plinio, i Romani nel primo secolo D.C.. Furono sempre i Romani che diffusero il pesco all'interno del loro impero fino in Francia ed in Inghilterra, sebbene i Mori lo avessero introdotto precedentemente in Spagna portandolo dal nord africa. Benché la tecnica dell'innesto fosse già conosciuta dai greci il pesco veniva propagato attraverso i semi che erano di più facile trasporto. Dopo alcune generazioni ed il relativo adattamento il pesco cominciò a differenziarsi: le pesche a polpa bianca erano tipiche della Francia, mentre quelle a polpa gialla erano caratteristiche della Spagna. All'inizio del '500 il pesco venne introdotto in Messico, in Sud ed in Nord America durante la colonizzazione del nuovo mondo da parte degli Spagnoli, dove la sua diffusione si deve anche ai nativi americani che lo trasportarono e lo piantarono durante le loro migrazioni.

Il maggior impulso economico nella coltura del pesco si deve all'avvento dell'era industriale, alla possibilità di conservare enormi quantità di cibo e alla globalizzazione dei mercati. Nasce la necessità di trasportare i frutti non più solo a livello regionale, ma nazionale e

internazionale. Proprio a questo scopo l'attività di miglioramento genetico, dalla fine del '900 agli anni novanta, si è focalizzata sull'ottenimento di varietà altamente produttive, di elevata pezzatura, di bell'aspetto, ma spesso insapori che potessero resistere alle manipolazioni, caratteristiche che soddisfano più i rivenditori che i consumatori.

Oggi gli obiettivi del miglioramento genetico non sono più questi, infatti negli ultimi 15 anni si è imposto il problema della qualità dei prodotti agro-alimentari. Questo crescente interesse nei confronti della qualità si deve, da un lato, ad una presa di coscienza da parte dei consumatori e dell'altro all'esigenza sempre più preponderante di un'agricoltura di tipo ecocompatibile o sostenibile. L'obiettivo odierno è quello di ottenere produzioni con elevate caratteristiche organolettiche e crescenti standard di salubrità sia per il consumatore sia per l'ambiente.

Uno dei principali problemi che concerne la qualità è quello di definirla in maniera scientifica. Esistono diverse definizioni e modelli uno dei quali è stato proposto da Peri nel 1997. Egli distingue fra "Qualità Strutturale" e "Qualità Funzionale": la prima si riferisce alle caratteristiche proprie o intrinseche del prodotto (per esempio il contenuto in zuccheri di un frutto), la seconda si riferisce alle funzioni ad esso correlate (come il gusto dolce). Un'altra distinzione apportata da Peri è quella fra "l'Efficacia" di un prodotto alimentare intesa come la capacità di soddisfare le esigenze d'uso del consumatore e la sua "Efficienza" come capacità di soddisfare le esigenze del produttore. Le aziende più innovative tendono a considerare complessivamente questi aspetti e basano i loro approcci produttivi sul TQM (Total Quality Management), la gestione totale della qualità.

Altri autori non operano questo tipo di distinzione, ma focalizzano il loro interesse sulle diverse "tipologie" di qualità dei prodotti agro-alimentari e sulle caratteristiche di ciascuna tipologia (Eccher-Zerbini, 1989). Essendo il consumatore il principale fruitore della qualità è su di lui che viene posta principalmente l'attenzione. Anche le norme di certificazione sulla qualità UNI e ISO individuano come soggetto referente il consumatore; in questo caso la qualità viene definita come: l'insieme delle proprietà e delle caratteristiche di un prodotto che ne determinano la capacità di soddisfare esigenze espresse o implicite.

Nel caso del frutto all'interno del concetto di qualità si devono inserire e valutare (Di Vaio et al., 2003):

- Le caratteristiche Organolettiche: l'insieme delle proprietà di un prodotto percepibili dai nostri sensi come il sapore l'odore e le sensazioni visive. Nel caso del pesco hanno enorme importanza il gusto (contenuto in zuccheri ed acidi organici), la consistenza della polpa, l'odore e l'aroma;

- Le caratteristiche Nutrizionali: proprietà che tengono conto della quantità e della qualità dei nutraceutici presenti nella frutta come vitamine ed antiossidanti. Il ruolo importante giocato da questi composti nella salvaguardia della salute è stato ormai documentato da alcuni studi scientifici (Di Mascio et al., 1989; Gaziano e Hennekens, 1993; du Toit et al., 2001);
- Gli standard Sanitari: questo aspetto coincide con quello di salubrità del prodotto, pertanto i frutti immessi sul mercato devono essere privi di sostanze tossiche, secondo le normative vigenti.

Sono numerosi i parametri che influenzano la qualità di un frutto, in prima istanza la varietà utilizzata, poi la tecnica colturale impiegata ed infine uno dei più importanti è l'epoca di raccolta che è strettamente connessa con la maturazione. Durante la maturazione del frutto avvengono delle notevoli modifiche fisiologiche le principali sono:

- Il cambiamento del colore: dal verde del frutto immaturo si vira al giallo fino al rosso a causa di una produzione eterogenea di antociani e carotenoidi a seconda delle cultivar;
- L'accumulo di zuccheri: in particolare si osserva un aumento del di saccarosio ed una diminuzione del sorbitolo;
- La diminuzione degli acidi organici: nel caso specifico del pesco si verifica un incremento del rapporto acido malico/ acido citrico;
- L'evoluzione delle sostanze volatili: diminuzione degli alcoli e di aldeidi ed incremento dei lattoni, responsabili dell'aroma fruttato (Robertson et al., 1990);
- L'ammorbidimento della polpa: dovuto al processo di degradazione delle pectine ed aumento della succosità ad opera delle poligattoronasi (PG) e dell'endo- β -1,4-glucanasi (EG) (Bonghi *et al.*, 1998).

La problematica più importante strettamente legata all'epoca di maturazione e alla qualità dei frutti è la pratica della raccolta commerciale anticipata, al fine di favorire la trasportabilità del prodotto. Questa pratica penalizza la qualità finale dei prodotti, poiché tutti i fenomeni correlati alla maturazione già citati avvengono nelle fasi finali dello sviluppo. Quindi è importante trovare un compromesso fra la maturazione fisiologica e quella commerciale al fine di ottenere delle caratteristiche organolettiche ottimali che consentano di mantenere i frutti integri fino al consumo.

All'interno di questo quadro generale un ruolo importante è quello che possono giocare le biotecnologie sia come mezzo di comprensione dei meccanismi biologici coinvolti nella maturazione del frutto, sia come supporto al miglioramento genetico classico come nel caso

della selezione genetica assistita con i marcatori molecolari (MAS, Marker Assisted Selection).

2 I MARCATORI MOLECOLARI

Tradizionalmente, la caratterizzazione e l'identificazione delle specie appartenenti al genere *Prunus* e delle cultivar relative si è basata sui caratteri morfologici e fisiologici. Questo tipo di marcatori può essere rilevato solo quando il carattere viene espresso e risente delle influenze ambientali.

Negli ultimi vent'anni i marcatori molecolari hanno efficacemente integrato gli approcci di miglioramento genetico classico (Weising K. et al., 1995). I marcatori molecolari possono essere sia proteine che acidi nucleici (DNA, RNA); in questo capitolo verranno presi in rassegna i principali marcatori basati sulla dissezione degli acidi nucleici.

I polimorfismi genetici sono definiti come differenze tra individui nella sequenza del DNA in una specifica regione del genoma, queste variazioni possono essere evidenziate attraverso l'uso dei marcatori molecolari.

Le caratteristiche principali che definiscono un marcatore molecolare sono:

1. Polimorfismo: individuazione di variazioni tra individui in regioni specifiche del genoma.
2. Codominanza: possibilità di riconoscere gli eterozigoti dagli omozigoti.
3. Assenza di effetti Epistatici: assenza di interazioni interlocus.
4. Neutralità: il marcatore molecolare non deve avere effetto a livello fenotipico.
5. Indipendenza dall'ambiente esterno: l'identificazione del marcatore molecolare non deve essere influenzata dalle variazioni delle condizioni ambientali.
6. Semplicità, Velocità e Basso Costo della metodologia.
7. Riproducibilità in laboratori diversi.
8. Contenuto informativo: riguarda l'informatività dei marcatori molecolari e cioè la loro capacità di distinguere e comparare il più alto numero di alleli.
9. Multiplex Ratio: è il numero di LOCI differenti che può essere simultaneamente analizzato in un esperimento.

Ogni marcatore è caratterizzato da una combinazione unica di vantaggi e svantaggi, quindi la scelta di quello più idoneo è dettata dalle peculiarità dell'applicazione.

Ci sono diversi fattori da considerare quando si sceglie un Marcatore Molecolare:

- ❖ Quantità di DNA per produrre i marcatori molecolari;
- ❖ Possibilità di accedere a collezioni di marcatori per una determinata specie o necessità di sviluppare nuovi marcatori;

- ❖ Il costo del marcatore;
- ❖ Possibilità di utilizzare sistemi alternativi al radioattivo.

Le principali applicazioni dei marcatori molecolari sono innumerevoli: Creazione di mappe genetiche, Mappatura di mutanti e caratterizzazione di trasformanti, Fingerprinting, Tassonomia molecolare ed evolutiva, Studi su malattie di origine genetica, Genetica di popolazione, Caratterizzazione del germoplasma.

I marcatori molecolari possono essere catalogati in numerosi modi in questo caso verranno presi prima in esame quelli che si basano sul DNA e poi quelli che provengono dal trascrittoma (l'insieme degli RNA messaggeri di una cellula).

2.1 Marcatori Molecolari derivati dal DNA

I marcatori molecolari basati sul DNA sono stati i primi ad essere utilizzati, a loro volta possono essere divisi fra marcatori che utilizzano tecniche di ibridazione e marcatori che si basano sulla reazione a catena della polimerasi (PCR).

2.1.1 Marcatori basati sulle tecniche di Ibridazione

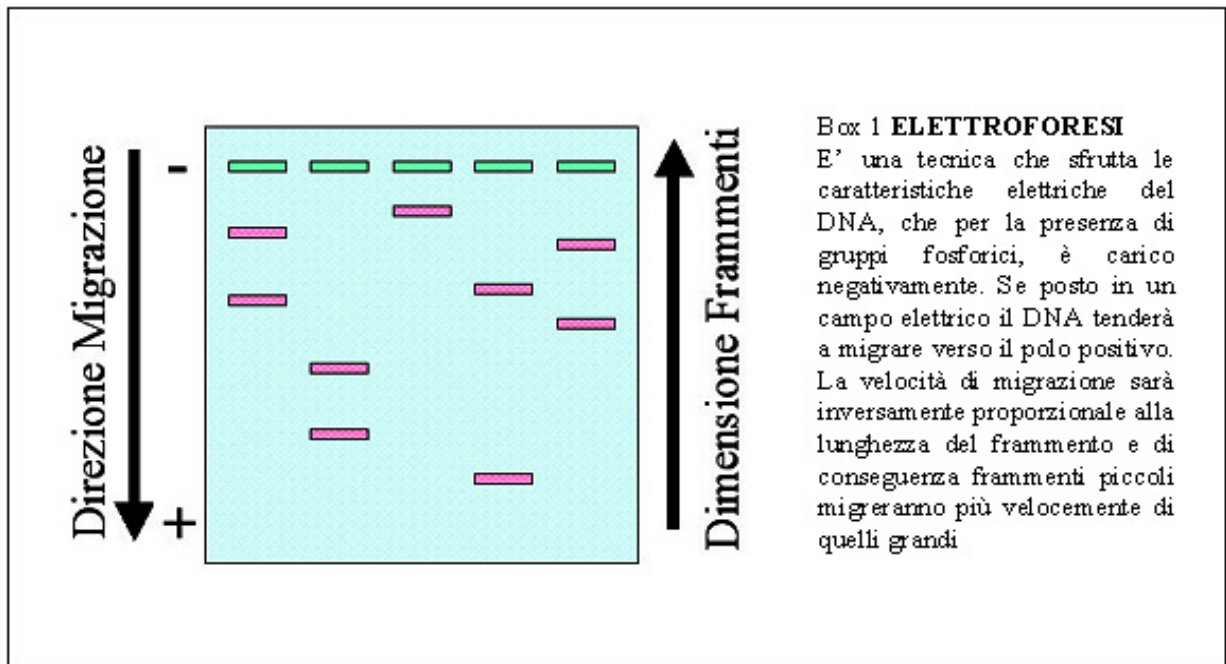
L'ibridazione è la prima delle tecniche che sfruttano la caratteristica del DNA di formare dei duplex; la molecola del DNA, nonostante la sua complessità strutturale, è in grado di appaiarsi perfettamente a partire da singoli filamenti.

L'ancoraggio dei singoli filamenti su di un supporto, come una membrana di nitrocellulosa, previene il riassociarsi dei frammenti gli uni con gli altri, e favorisce l'associazione con filamenti singoli di DNA o RNA. Questa possibilità, in combinazione con la separazione dei frammenti, ottenuti tramite digestione enzimatica, in un campo elettrico (Box 1) ha dato il via alla nascita della prima classe di marcatori molecolari: i Restriction Fragment Length Polymorphism (RFLP) (Bostein et al., 1980). Questi marcatori si trasmettono in maniera mendeliana semplice e derivano dai riarrangiamenti del DNA che avvengono durante i fenomeni evolutivi: mutazioni puntiformi, inserzioni/ delezioni o errati crossing-over a livello della sequenza di riconoscimento degli enzimi di restrizione (Figura 2.1).

In questa tecnica gioca un ruolo fondamentale l'impiego di enzimi di restrizione (di tipo II) e di sonde molecolari.

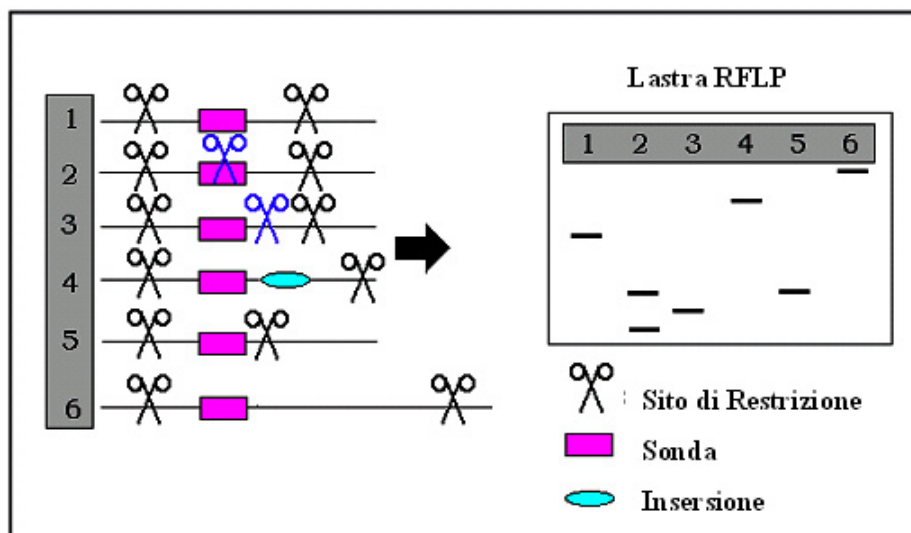
La specificità degli enzimi di restrizione, capaci di tagliare il DNA in specifiche regioni, è tale che la sostituzione di una sola base all'interno della sequenza di riconoscimento impedisce all'enzima di effettuare il taglio (le mutazioni a livello di queste sequenze non sono rare, anzi sono uno dei motori dell'evoluzione); ed è proprio questa caratteristica che permette di evidenziare polimorfismi. Il vero problema non è tanto generare dei frammenti di restrizione

(ottenuti dalla digestione totale del DNA genomico), ma di distinguere quelli di interesse e di seguirne il destino nei diversi individui; questo problema viene risolto attraverso l'uso di Sonde Molecolari.



Le sonde molecolari impiegate per ottenere marcatori RFLP sono dei frammenti di dimensioni variabili, possibilmente presenti nel genoma in singola copia o con basso numero di copie, generalmente ottenuti a partire da librerie genomiche o librerie a cDNA. Le sonde di cDNA derivano dall'mRNA di un individuo, che viene convertito in cDNA (DNA copia) tramite la trascrittasi inversa; queste sonde rappresentano tutto il pool genetico che viene trascritto nella cellula e di conseguenza possono essere anche utilizzate per l'individuazione diretta di geni.

Figura 2.1 Individuazione di Polimorfismi attraverso gli RFLP



Le sonde di DNA genomico sono generalmente individuate entro librerie genomiche opportunamente costruite. Porzioni discrete del patrimonio genetico di un individuo non vengono espresse, questo perché la polimerasi non riesce ad agire a causa dell'elevato livello di metilazione in queste regioni. Utilizzando enzimi di restrizione, quali PstI, che non sono in grado di tagliare regioni metilate di DNA si ottengono librerie genomiche definite arricchite per sequenze a singola copia.

La scelta della coppia sonda di DNA/enzima di restrizione è cruciale e mediante opportune prove viene identificata la combinazione più idonea, quella che mette in evidenza il maggior numero di polimorfismi (Tanskley et al., 1989)

La funzionalità degli RFLP per l'analisi dei genomi vegetali è stata subito riconosciuta ed applicata con successo ad un elevato numero di specie (Beckman & Soller, 1983), gli RFLP sono stati efficientemente impiegati per eseguire analisi di segregazione e costruire mappe genetiche di specie appartenenti al genere *Prunus* (Rajapakse, S. et al., 1995; Joobeur et al., 1998; Wang D. et al., 2000; Verde et al., 2005).

I marcatori RFLP presentano il notevole vantaggio, rispetto ad altri tipi di marcatori, di poter individuare 'loci ancora' che rappresentano punti di riferimento all'interno di una mappa genetica. Infatti, una sonda di DNA, soprattutto sonde a singola copia, individuerà sempre lo stesso locus RFLP indipendentemente dalla combinazione di incrocio utilizzata per ottenere la mappa. Nonostante i numerosi vantaggi, questa tecnica presenta alcuni inconvenienti: è laboriosa, ha tempi di lavoro prolungati, necessita di un'elevata quantità di DNA di partenza e spesso per la visualizzazione dei frammenti vengono utilizzati isotopi radioattivi.

Per superare alcune di queste difficoltà una valida alternativa è fornita dagli STS (Sequence Tagged Site), in questo caso un marcatore RFLP viene convertito per essere utilizzato nella reazione a catena della polimerasi (PCR, tecnica che verrà elucidata in seguito). Questo tipo di marcatori sono molto utili quando si trovano bande RFLP strettamente associate a caratteri morfologici (Hartl L. et al., 1993), poiché possono essere impiegati più agevolmente nella selezione genetica assistita dai marcatori (MAS).

Un'altra classe di marcatori molecolari che si basa sulle tecniche di ibridazione è quella delle Expressed Sequence Tag (EST), questa terminologia venne utilizzata per la prima volta da Adams nel 1991. Le EST sono un validissimo strumento per l'identificazione di nuovi geni ed un importante supporto per la mappatura dei genomi (Cooke et al., 1996). Per quel che riguarda il genere *Prunus* sono state impiegate per la costruzione di una mappa di trascrizione e per la mappatura fisica del pesco. (Horn et al., 2005)

Il concetto di EST deriva dalla possibilità di ottenere del DNA copia (cDNA). Quando un gene viene espresso, la sua sequenza di DNA viene trascritta in RNA messaggero (mRNA), queste molecole vengono utilizzate come template durante la sintesi proteica. L'espressione dei geni è tessuto, tempo e ambiente specifica. Grazie alla trascrittasi inversa, la fragile e instabile molecola di RNA messaggero può essere trasformata in cDNA. I frammenti di cDNA così ottenuti sono clonati e vanno a costituire delle "Librerie di cDNA". Le EST sono dei segmenti clonati di cDNA che vengono stati parzialmente sequenziati per alcune centinaia di paia di basi. Questo tipo di sequenziamento viene definito single-pass in opposizione al double-pass, i cloni di cDNA vengono sequenziati a partire da una delle due estremità e non si ottengono quindi una sequenza forward ed una reverse come invece accade in genere.

Lo sviluppo di questa tecnica è stato possibile in seguito alla creazione di strumentazioni specifiche che hanno consentito il sequenziamento su larga scala ed alla nascita di una nuova branca della biologia la Bioinformatica che ha permesso l'analisi dell'enorme quantità di dati prodotti; normalmente un progetto EST prevede il sequenziamento di centinaia se non migliaia di cloni.

Dagli anni novanta ad oggi milioni di EST sono state depositate in decine di banche dati pubbliche e private, le potenzialità di queste informazioni sono enormi e necessitano di un capitolo apposito che verrà trattato in seguito.

2.1.2 Marcatori basati sulla reazione a catena della polimerasi (PCR)

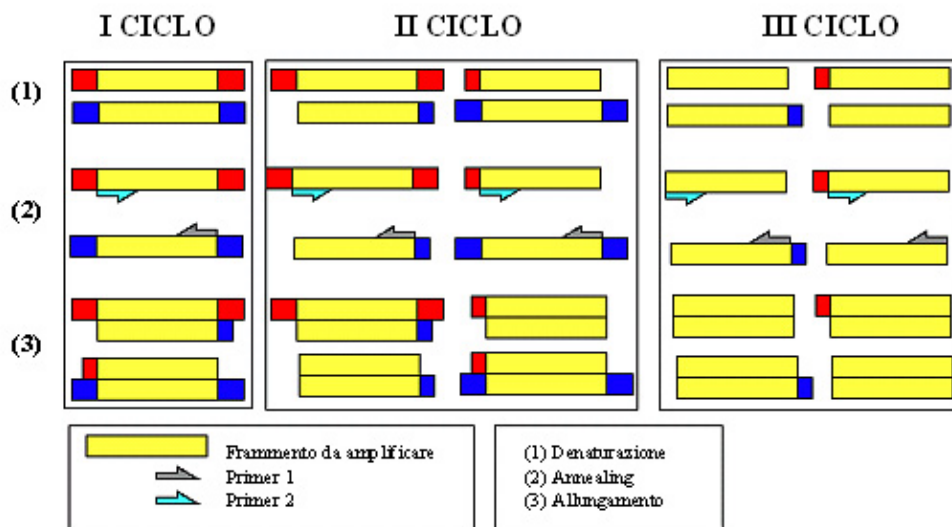
Uno degli eventi tipici nelle biotecnologie fu l'invenzione da parte di Mullis della reazione a catena della polimerasi PCR (1984) che permette di amplificare selettivamente degli specifici tratti di DNA.

Il punto di partenza di una reazione di PCR è una miscela che contiene: il DNA con la sequenza che si vuole amplificare, due primer che fiancheggiano la sequenza d'interesse, una DNA polimerasi che a partire da un piccolo tratto di DNA a doppio filamento sintetizzi il tratto desiderato, dNTP (deossinucleotidi) e ioni Mg^{++} . I primer sono degli oligonucleotidi sintetici, della lunghezza massima di 50 bp, che vengono disegnati in base alla sequenza che si vuole amplificare. Poiché normalmente il DNA è a doppio filamento, la prima fase della PCR è la DENATURAZIONE del DNA sufficiente per far avvenire la separazione dei due filamenti della doppia elica (Figura 2.2).

La seconda fase è detta di ANNEALING, consiste in un abbassamento della temperatura fino a raggiungere un valore (temperatura di annealing o di associazione) che è correlato alla lunghezza, composizione e temperatura di fusione (T_m) dei primer.

La terza fase è quella di ALLUNGAMENTO in cui una DNA polimerasi sintetizza un nuovo filamento, a partire dall'estremità 3' del primer, utilizzando i dNTP presenti nella miscela di reazione. Questa fase si esegue sfruttando la DNA polimerasi di un batterio termofilo il *Thermus aquaticus* che ha il suo optimum di attività a 72°C e che non viene degradata alle temperature elevate impiegate durante la fase di denaturazione. Il tempo di allungamento è direttamente proporzionale alla lunghezza del frammento che si vuole amplificare. Queste tre fasi vengono ripetute un numero elevato di volte, ed è proprio questa ciclicità che permette di amplificare il frammento di DNA secondo una progressione geometrica.

Figura 2.2 Reazione a Catena della Polimerasi (PCR)



Solo quattro anni dopo la pubblicazione del lavoro di Mullis, Williams ed il suo gruppo (1990) descrissero l'uso di corti oligonucleotidi (8-10bp) con sequenza arbitraria da utilizzare come primer per l'ottenimento, tramite PCR, di marcatori molecolari denominati RAPD (Random Amplified Polymorphic DNA) distribuiti casualmente nel genoma.

Le applicazioni furono immediate: Rafalski et al. (1991), Waugh e Powell (1992) utilizzarono marcatori RAPD per il mappaggio ed il miglioramento genetico delle piante; Hadrys et al (1992) ne illustrarono le potenzialità in Ecologia molecolare. Questa velocità nell'applicazione della tecnica è dovuta alla semplicità della PCR, all'impiego di oligonucleotidi sintetici con sequenza casuale ed all'esigua quantità di DNA necessaria, notevolmente inferiore rispetto a quella per gli RFLP.

I polimorfismi RAPD sono evidenziati tramite elettroforesi (una separazione dei frammenti in base alla loro dimensione in campo elettrico).

Gli svantaggi di questi marcatori sono principalmente due: la dominanza e la riproducibilità. I marcatori RAPD vengono valutati in base alla presenza/assenza di singole bande e quindi non

consentono di distinguere genotipi eterozigoti dagli omozigoti (dominanza), questa situazione causa una diminuzione dell'informazione fornita. Inoltre l'impiego di oligonucleotidi corti (10bp) e di temperature di annealing molto basse (35-37°C) determinano problemi di riproducibilità tra laboratori differenti in quanto cambiamenti, anche minimi, nelle condizioni della reazione di PCR comportano la variazione del pattern dei prodotti di amplificazione. Ci sono infine problemi correlati con l'interpretazione dei risultati, in quanto prodotti di amplificazione differenti ma con peso molecolare simile migrano in maniera identica nel gel. Quindi la presenza in due individui di una stessa banda non significa che questi posseggano il medesimo frammento di DNA.

Anche in questo caso, quindi, ci sono dei pro e dei contro, di conseguenza la scelta di questa metodica dipende dal tipo di informazioni che si desidera ottenere.

Per quel che riguarda il genere *Prunus* questa tecnica è stata ampiamente utilizzata per l'identificazione varietale (Lu et al., 1996), per la stima della diversità genetica (Warbouton M.L. et al., 1996; Bartolozzi F. et al., 1998) e per la costruzione di mappe (Rajapakse S. et al., 1995; Dirlewanger E. et al., 1998; Joobeur T. et al., 2000).

I difetti dei marcatori RAPD sono stati superati da Michelmore e Martin (Michelmore R.W. et al., 1991; Martin G.B. et al., 1991). Questi ricercatori hanno introdotto la tecnica degli SCAR (Sequenze Characterized Amplified Region) che trasforma delle bande RAPD, in sequenze target.

Le bande RAPD di interesse vengono sequenziate e poi utilizzate per disegnare dei primer specifici (22-24 nucleotidi). I marcatori SCAR sono più riproducibili dei RAPD anche se rimangono dei marcatori prevalentemente dominanti. Sono stati impiegati con successo nella MAS per l'identificazione del gene *Ma1* per la resistenza al root-knot nematode in *Prunus cerasifera* Ehr. (Lecoals A.C. et al., 1999) e l'identificazione di genotipi fondenti (*melting flesh*) in pesco (Jun J.H. et al., 2002).

Alcune classi di marcatori molecolari sono la combinazione di due o più metodiche come i CAP (Cleavage Amplified Polymorphic Sequence) e gli AFLP (Amplified Fragment Length Polymorphism).

I CAP sono il risultato di una digestione enzimatica applicata a dei prodotti di PCR. Questo tipo di marcatore è codominante e mette in evidenza i polimorfismi nella sequenza all'interno del sito di restrizione. I CAP sono stati utilizzati per la prima volta per mappare mutazioni in *Arabidopsis thaliana* (Konieczny A. et al., 1993; Jarvis P. et al., 1994).

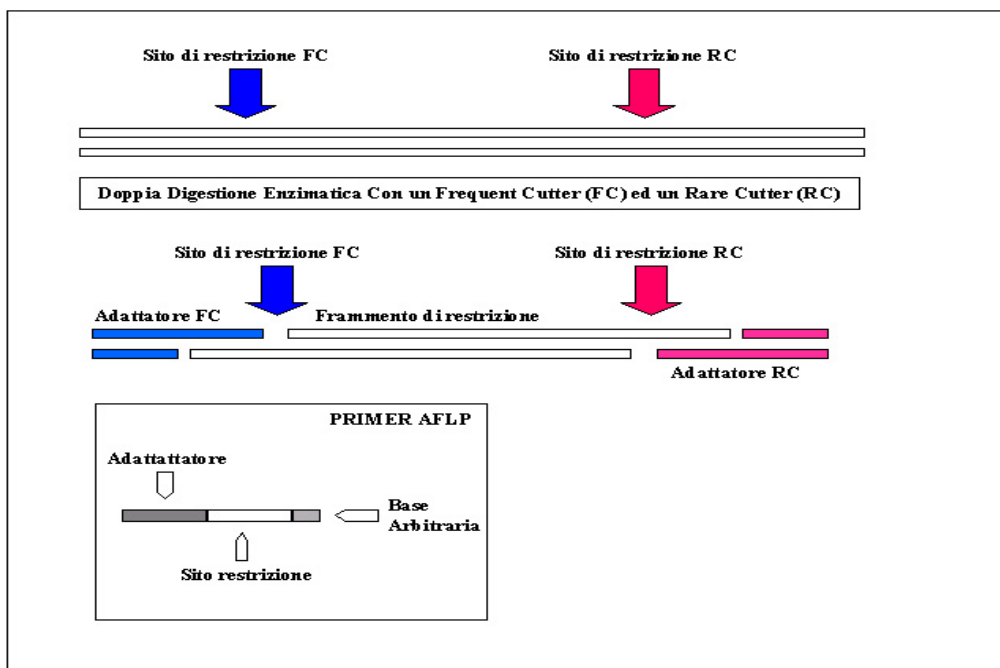
Gli AFLP (Amplified Fragment Length Polymorphism) sono stati ideati da Vos del 1995, questa classe di marcatori molecolari unisce insieme la tecnica degli RFLP e quella dei RAPD.

La Tecnica degli AFLP si basa sull'amplificazione di subsiti di frammenti di restrizione genomici; il DNA viene digerito utilizzando due enzimi di restrizione contemporaneamente:

- ✓ Un rare cutter: enzima con sito di taglio esanucleotidico;
- ✓ Un frequent cutter: enzima con sito di taglio tetranucleotidico.

La tecnica AFLP fa in modo di amplificare solo una parte dei frammenti che si generano dopo la digestione enzimatica, e per la precisione quelli che hanno ad un'estremità il sito di restrizione di un frequent cutter e dall'altra quello del rare cutter.

Figura 2.3 Ottenimento di frammenti di restrizione da utilizzare per analisi AFLP



Ai frammenti ottenuti vengono ligati degli adattatori a doppio filamento al fine di generare uno stampo di DNA da utilizzare per l'amplificazione. La sequenza degli adattatori e dei siti di restrizione adiacenti serve come sito di legame per il primer per la successiva amplificazione dei frammenti di restrizione, i siti di restrizione degli enzimi vengono distrutti dalla ligazione degli adattatori.

All'estremità 3' del primer viene posizionato un nucleotide arbitrariamente scelto, questo ha la funzione di selezionare e quindi amplificare solo un subset dei frammenti di restrizione prodotti durante la digestione. Dopo questa prima amplificazione, detta pre-selettiva, viene effettuata un'altra reazione di PCR con il medesimo primer utilizzando il prodotto della pre-amplificazione come template, ma all'estremità 3' del primer vengono posizionati due o tre nucleotidi arbitrariamente scelti. In questo modo si opera un'ennesima selezione all'interno

dei frammenti prodotti inizialmente dalla doppia digestione. I prodotti ottenuti vengono quindi visualizzati su un gel di poliacrilammide gel che ha un elevato potere di risoluzione. Questo tipo di marcatori sono altamente informativi anche se sono di tipo dominante. Se si considera pari ad 1 la totalità dei frammenti prodotti dalla digestione che abbiano all'estremità un sito di restrizione per il frequent cutter ed all'altra un sito per il rare cutter, la probabilità che contemporaneamente siano presenti 3 basi arbitrarie da un lato e 3 dall'altro è pari a $(\frac{1}{4})^6$. Nonostante questo valore apparentemente irrisorio una sola combinazione di primer AFLP può produrre fino a 30 bande polimorfiche, questo a causa dell'elevato numero di frammenti prodotti durante la digestione enzimatica del DNA genomico.

Gli AFLP sono stati ampiamente utilizzati per la discriminazione varietale (Powell et al., 1996; Hurtado 2002) e per la saturazione di mappe d'associazione (Verde et al., 2005).

Uno dei passi più importanti nella genetica fu la scoperta che il 30-90% del genoma delle maggior parte delle specie è costituito da regioni ripetute di DNA le quali, in natura, sono altamente poliformiche. Queste regioni contengono molti loci con centinaia di alleli che differiscono fra loro per la lunghezza e/o la sequenza e sono dispersi casualmente all'interno del genoma. Le regioni ripetute giocano un ruolo importante nell'assorbimento delle mutazioni, infatti solo le mutazioni che vengono ereditate giocano un ruolo attivo nei processi evolutivi.

Una di queste categorie di sequenze ripetute è rappresentata dai microsattelliti, termine coniato per la prima volta da Litt e Luty nel 1989, appartengono alla categoria dei marcatori molecolari basati sulla PCR e negli ultimi 10 anni hanno avuto una crescita esponenziale come testimoniano il numero sempre crescente di pubblicazioni a riguardo.

In base alla lunghezza del motivo ripetuto vengono distinte diverse tipologie di sequenze ripetute, i microsattelliti o Sequenze Semplici Ripetute (SSR) sono motivi mono-, di-, tri-, tetra-, penta-, esa-, eptanucleotidici ripetuti in tandem; furono scoperti casualmente durante lo studio di librerie genomiche umane (Hamada et al., 1982), in seguito furono anche individuate nel genoma di altri organismi eucarioti e procarioti (Tautz et al., 1984).

Gli SSR sono distribuiti "random" all'interno del genoma sia nelle regioni codificanti che in quelle non codificanti (Tautz et al., 1984) e mostrano un'elevata variabilità sia per quanto riguarda la dimensione che per la composizione nucleotidica (Jeffreys et al., 1985, Tautz et al., 1989, Kashi et al., 1997).

Il meccanismo di origine del polimorfismo generato dai microsattelliti non è ancora del tutto noto, ma si ritiene che sia dovuto ad eventi di scivolamento durante la replicazione del DNA (Schlötterer & Tautz, 1992). Sutheland e Richards (1995) hanno dimostrato che le mutazioni

dinamiche del tri-nucleotide CCG (presente nel gene *FMRI*, *fragile X mental retardation* sul cromosoma X) portano ad un aumento repentino del numero di ripetizioni e sono strettamente associate alla Sindrome dell'X Fragile. Quando lo stesso evento di mutazione interessa la tripletta AGC all'interno del gene della miotonina sul cromosoma 19, si verifica la Distrofia Mitotica.

Il significato adattativo o casuale, il meccanismo di evoluzione e la funzione dei microsatelliti non sono chiari, ma si ritiene che le diverse frazioni di DNA ripetuto possano avere differenti funzioni: funzione strutturale nei cromosomi poiché blocchi di DNA satellite si trovano nei centromeri, dove contraggono rapporti con le fibre del fuso mitotico, e altri con diversa sequenza a livello dei telomeri, dove partecipano a specifici meccanismi che rendono possibile la replicazione del DNA fino all'estremità dei cromosomi; funzioni di controllo della struttura secondaria del DNA e, di conseguenza, della sua replicazione e trascrizione (Curtoni et al., 1991).

Le sequenze microsatelliti sono state utilizzate come marcatori molecolari poiché mostrano un elevato polimorfismo e sono codominanti. L'elevato valore informativo, direttamente collegato all'alto numero di alleli per locus e la semplicità della tecnica della PCR, fanno degli SSR i marcatori genetici ideali.

D'altro canto bisogna considerare le difficoltà di generare questi marcatori rispetto ad altri (Rafalski & Tinget, 1993) in quanto necessitano di primer specifici disegnati in base alla sequenza delle zone fiancheggianti la regione microsatellite.

L'isolamento di sequenze SSR a partire da DNA genomico può avvenire in diversi modi (Zane et al., 2002):

- Screening di librerie genomiche con sonde specifiche marcate
- PIMA Pcr Isolation of Microsatelliti Arrayas (Lunt et al. 1999)
- Screening di librerie genomiche arricchite per sequenze microsatelliti.

Tutte queste metodiche sono laboriose ed economicamente impegnative, oggi un'altra fonte di microsatelliti proviene dall'elevato numero di EST prodotte per numerose specie animali e vegetali. In questo caso, le sequenze microsatelliti sono disponibili in banche dati pubbliche e possono essere associate a geni con funzione nota.

All'interno del genere *Prunus* sono stati isolate e pubblicate più di 300 coppie di primer microsatelliti (Cipriani et al., 1999; Sosinski et al., 2000; Testolin et al., 2000; Dirlwanger et al., 2002; Aranzana et al., 2002; Yamamoto et al., 2002; Lopes et al., 2002; Wang et al., 2002; Decroocq et al., 2003; Georgi et al., 2003; Mnejja et al., 2004; Vaughan et al., 2004; Verde et al., 2004; Testolin et al., 2004; Messina et al., 2004; Mnejja et al., 2005; Verde et al., 2005),

in oltre questi marcatori sono stati ampiamente applicati per la costruzione di mappe d'associazione (Joobeur et al., 2000; Dettori et al., 2001; Dirlewnger et al., 2004a), per la discriminazione varietale (Testolin et al., 2000), per la stima della diversità genetica fra pesco e mandorlo (Martinez-Gomez, 2001) e per studi comparativi fra mappe all'interno delle rosacee e per la MAS (Dirlewnger et al., 2004a).

2.2 Marcatori Molecolari derivati dal Trascrittoma e dalle Regioni Espresse

Negli ultimi anni lo sviluppo di nuovi marcatori molecolari ha riguardato non più il DNA, ma la porzione trascritta del genoma. Questo è stato reso possibile in seguito al verificarsi di due condizioni principali: la disponibilità di numerosi cloni di cDNA provenienti da diverse specie vegetali e l'accumularsi di un elevato numero di EST all'interno di banche dati pubbliche (<http://www.tigr.org>: the Insitute of Genome Research; <http://www.ncbi.nlm.nih.gov>: National Center of Biotechnology Information; <http://www.ebi.ac.uk>: European Informatic Insitute).

I cloni di cDNA possono essere utilizzati come sonde RFLP per individuare dei cDNA-RFLP o per sviluppare dei cDNA-AFLP. Allo stesso modo, le banche dati EST possono essere analizzate per l'identificazione di sequenze SSR e SNP (Single Nucleotide Polymorphism), per poter poi disegnare delle coppie di primer locus-specifici. Un'altra classe di marcatori è stata sviluppata disegnando dei primer in base alle sequenze fiancheggianti un frammento EST amplificabile (Amplicone); questa tipologia di marcatori è definita come polimorfismi EST che possono individuare variazioni sia nella lunghezza che nella sequenza.

I marcatori molecolari basati sul trascrittoma possono evidenziare polimorfismi dovuti alla lunghezza e alla composizione nucleotidica all'interno della porzione espressa del genoma, e sono pertanto più informativi rispetto ai marcatori molecolari genomici che sono stati ampiamente impiegati in passato. Questi marcatori sono stati utilizzati, inizialmente, con successo per la costruzione di mappe di trascrizione in diverse specie vegetali. Lo sviluppo dei GTM (gene target marker), localizzati all'interno di geni d'interesse, è stato possibile utilizzando le informazioni ottenute da diversi progetti di ricerca basati sulla genomica funzionale e strutturale. In alcuni casi è stato anche possibile ottenere dei FM (Functional Marker) che sono casualmente correlati a variazioni fenotipiche utili per il miglioramento genetico (Andersen & Lübberstedt, 2003).

Sia la classe dei marcatori che si basano sull'ibridazione sia quelli basati sulla PCR sono stati sviluppati sfruttando il trascrittoma, in questo caso possono essere suddivisi in base alla fonte dalla quale derivano: cDNA, EST, o sequenze di geni noti.

2.2.1 cDNA o mRNA come fonte di Marcatori

In molte specie vegetali una notevole quantità di marcatori molecolari sono stati sviluppati con successo da mRNA e cDNA derivanti da differenti tessuti i quali, in alcuni casi, sono stati sottoposti a specifici stimoli o stress.

A questa categoria appartengono i marcatori cDNA-RFLP in cui cDNA vengono utilizzati come sonde RFLP. Le sonde sono reperite da differenti fonti, possono essere dei frammenti di cDNA che individuano geni specifici e che vengono utilizzati all'interno di generi e specie correlati, oppure sono dei prodotti di PCR clonati, che derivano, dall'amplificazione con primer specifici disegnati sulla base di sequenze di cDNA o genomiche. Un'altra fonte di sonde RFLP possono essere gli stessi geni che si intende studiare. Questa tipologia di marcatori è stata impiegata per la costruzione di mappe d'associazione (Thamarus K.A. et al., 2002).

Quando gli AFLP utilizzano come template il cDNA vengono definiti cDNA-AFLP, in questo caso la tecnica è del tutto simile a quella degli AFLP sviluppati sul DNA genomico, le uniche differenze consistono nella scelta degli enzimi restrizione, a volte si può utilizzare anche un solo enzima (Habu et al., 1997). Questa tecnica è stata utilizzata con successo in patata per visualizzare frammenti differenzialmente espressi in tessuti diversi (Bachem et al., 1996, 1998). Per l'RNAF-DD (RNA Fingerprinting Differential Display). La strategia prevede l'utilizzo di un set di coppie di primer appositamente disegnate; in ogni coppia un primer è ancorato alla coda di adenosine terminali e ha 3 basi arbitrarie che selezionano solo una parte dei messaggeri, mentre l'altro è un primer arbitrario complementare. Queste amplificazioni selettive dell'mRNA permettono l'identificazione di geni differenzialmente espressi (Liang et al., 1992); in grano tenero questa tecnica ha permesso la mappatura fine delle regioni ricche in geni (GRR) nel cromosoma 1 (Sandhu et al., 2002).

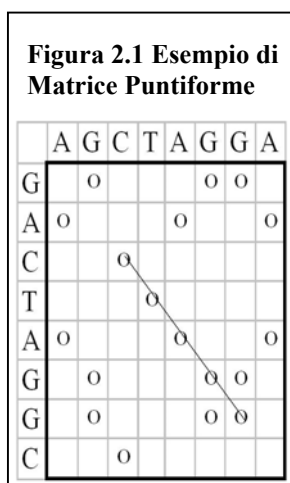
Un'ultima tecnica che utilizza cDNA come fonte di marcatori è la SAGE (Serial Analysis of Gene Expression) in cui il DNA copia viene utilizzato per creare un libreria genomica arricchita per sequenze microsatelliti (Hayden et al., 2002). Coppie di primer ancorati in 5' grazie alla presenza di una sequenza SSR (CT_n) vengono utilizzati per l'amplificazione creando dei concatenameri contenenti più di un singolo locus microsatellite. Il sequenziamento di questi frammenti porta all'isolamento di diversi loci SSR con una sola reazione di sequenza.

2.2.2 Banche dati EST come fonte di Marcatori

Recentemente si è assistito all'espansione delle banche dati EST, questo fenomeno ha consentito di sviluppare una nuova generazione di marcatori quali EST-SSR, SNP, InDel.

La gestione di banche dati di sequenza sempre più grandi ha reso necessaria l'applicazione di software appositamente studiati, è nata così una nuova branca della Biologia: la Bioinformatica. Le prime banche dati di sequenza risalgono agli anni '80 come quella sviluppata da Goad per il GeneBank Los Alamos National Laboratory (LANL) e quella sviluppata in Europa la European Molecular Biology Laboratory (EMBL), la DNA DataBank of Japan (DDBJ), sviluppata in Giappone nel 1984. Dal 1992 la GeneBank è stata assorbita all'interno del National Center of Biotechnology Information (NCBI). Oggi queste tre banche dati si sono consorziate all'interno dell'NCBI per facilitare l'accesso all'informazione da parte del mondo scientifico. Per capire l'importanza di queste collezioni di sequenza basti pensare che in GeneBank nel dicembre 1997 erano presenti 1.26×10^9 basi e che nel 2000 erano diventate 1.0×10^{10} . Le caratteristiche primarie di una banca dati sono sostanzialmente due: la possibilità di aggiornamento rapido ed efficiente e la facilità di confronto di una sequenza con tutte quelle contenute nella banca.

L'aumento così consistente del numero di sequenze disponibili ha posto il problema dell'automazione dell'analisi. Infatti mentre è pensabile leggere una sequenza manualmente non è immaginabile, per questioni di tempo, leggerne 100 o 1000. Il sequenziamento del DNA consiste nell'ordinare un set di picchi (A,T,C,G) che provengono da un gel di sequenza, questo procedimento è soggetto ad errori strettamente correlati con la qualità stessa del gel. Questo tipo di problema venne affrontato da diversi scienziati producendo una serie di programmi che svolgono un'analisi automatizzata della qualità di ogni picco per poi poter assegnare correttamente la base correlata. Uno di questi programmi denominato PHRED fu



ideato da Ewing (Ewing et Green, 1998; Ewing et al., 1998), questo software, tramite un apposito algoritmo, assegna dei valori di qualità ideali ad ogni picco e poi confronta questi valori con quelli reali. Se il valore reale non si discosta in modo significativo da quello ideale viene assegnata la base, in caso contrario il picco viene definito come sconosciuto. La produzione di sequenze affidabili è una necessità imprescindibile, infatti queste rappresentano la base per poter poi procedere con confronti ed assemblaggi.

Un'altra caratteristica fondamentale delle banche dati è che queste

siano facilmente consultabili per effettuare dei confronti fra sequenze. Per fare ciò, inizialmente, sono state costruite delle matrici puntiformi (DOT MATRIX Gibbs and McIntyre, 1970) in cui le due sequenze venivano scritte una in verticale e l'altra in orizzontale, quindi in ogni intersezione (due basi uguali) veniva disegnato un punto, alla fine veniva tracciata una linea sulla diagonale che evidenziava la similarità delle due sequenze (Figura 2.1). Questo tipo di matrice era funzionale per fare allineamenti singoli, ma non riusciva ad evidenziare inserzioni e delezioni, eventi che hanno un'importanza biologica rilevante. Per queste ragioni si è passati ad allineamenti dinamici (Needleman et Wunsch, 1970) in cui due sequenze aminoacidiche o nucleotidiche, vengono confrontate mantenendone una fissa e muovendo l'altra di una base alla volta analizzando tutti i possibili appaiamenti. In questo caso l'obiettivo è quello di determinare l'appaiamento migliore a seconda del punteggio ottenuto (1 similarità, 0 differenza); l'allineamento con il punteggio migliore viene definito allineamento ottimale. Questi due tipi di matrice definiscono un allineamento Globale, ma biologicamente ci sono delle regioni che si allineano perfettamente ed altre regioni che invece mostrano maggior difformità, queste regioni conservate hanno un'importanza funzionale e strutturale elevata e permettono di fare delle predizioni sul loro ruolo biologico.

Da ciò deriva la necessità di riuscire ad individuare degli allineamenti locali, piuttosto che trovare degli allineamenti globali, Waterman e Smith nel 1981 svilupparono un algoritmo in grado di individuare allineamenti locali. Questo tipo di statistica riusciva anche ad evidenziare inserzioni e delezioni, fenomeni che hanno una forte rilevanza biologica nei fenomeni evolutivi ed adattativi. La similarità, fra due sequenze, viene definita come il rapporto fra gli appaiamenti esatti e la somma degli appaiamenti esatti (match) e quelli inesatti (mismatch) non tenendo conto dei gap. La distanza è invece il rapporto fra gli appaiamenti inesatti e la somma degli appaiamenti esatti (match) e di quelli inesatti (mismatch) e può essere anche definita come il numero delle sostituzioni necessarie per trasformare la sequenza a in b. E' una misura molto importante per la predizione delle distanze evolutive.

Per quel che riguarda l'allineamento locale la maggiore difficoltà sta nel capire quali siano gli appaiamenti significativi e quali no, infatti la similarità ottenuta è frutto del caso o è reale? L'analisi di due sequenze simili, per effetto del caso, rivela valori di similarità più alti della distribuzione normale; la distribuzione evidenzia una distorsione positiva che viene definita come valore estremo di distribuzione. Karlin e Altschul (1990,1993) idearono un algoritmo in grado di definire la probabilità che l'indice di similarità sia dovuto al caso piuttosto che alla reale uguaglianza delle due sequenze. Questa relazione è di estrema importanza quando si

confronta una sequenza contro un'intera banca dati, in questo caso la probabilità che due sequenze siano appaiate casualmente è di 10^{-7} e confrontando la sequenza con un database di 80.000 sequenze il maggiore valore atteso (definito EXPECT score) è $(10^{-7} \times 8 \times 10^4 = 8 \times 10^{-3})$ 0,008; un valore compreso fra 0,02-0,05 viene considerato significativo. Quando si ottengono dei valori di similarità significativi gli allineamenti vanno comunque sempre controllati e verificati. Ma la vera sfida era quella di ottenere degli allineamenti multipli (Johnson et Doolittle, 1986), che consentissero di confrontare una sequenza ignota con tutto quelle presenti nelle banche dati a cui è stata assegnata una funzione reale o putativa (sequenze annotate). Questa metodologia richiede computer estremamente potenti ed è solitamente basata sull'allineamento sequenziale delle coppie di sequenze più simili. Quando l'allineamento di un set di sequenze è avvenuto individuando una famiglia, le regioni altamente conservate (Gribskov, 1987) vengono identificate poiché possono essere peculiari di quella famiglia e possono essere importanti per evidenziare nuove sequenze appartenenti alla stessa. Generalmente le matrici utilizzate sono due la PROFILE e la Position-Specific Scoring Matrix (PSSM). Queste matrici sono state utilizzate per la creazione del programma FASTA (Pearson et Lipman, 1988) che permette una veloce interrogazione della banca dati alla ricerca di similarità di sequenza. Questo software consente questa operazione in tempi rapidi e possiede un'interfaccia di facile utilizzazione anche per i non addetti ai lavori. Un'evoluzione del FASTA è il BLAST ideato da Altschul nel 1990, questo programma è ancora più veloce ed è infatti il più utilizzato al mondo. Si riesce a confrontare una sequenza sconosciuta con milioni di altre sequenze presenti in banca dati in pochi minuti. Ci sono differenti opzioni quando si procede con un'analisi come questa, si può fare il confronto con solo sequenze nucleotidiche o con sequenze nucleotidiche tradotte o con sequenze amminoacidiche tradotte in nucleotidiche. Questi programmi a volte consentono anche di fare delle predizioni sulla struttura secondaria e terziaria della proteina correlata e sulla struttura dei geni individuando introni, esoni, origini di replicazione., ecc. In oltre sono anche nati degli strumenti che consentono di individuare siti di restrizione e probabili polimorfismi rendendo possibili degli esperimenti detti "*in silico*". La Keygene ha sviluppato una banca dati completa per *Arabidopsis thaliana* attraverso la quale è possibile ottenere un bandeggio cDNA-AFLP analizzando le EST presenti. Allo stesso modo Qin et al. (2001) hanno utilizzato il programma GenEST per correlare dei frammenti predetti, partendo da EST con quelli cDNA-AFLP ottenuti analizzando il nematode *Globodera rostochiensis*. Un altro programma ideato a questo scopo è AFLPinSilico (Rombauts et al., 2003) che è stato utilizzato con successo per riprodurre un bandeggio cDNA-AFLP *in silico* e per un'analisi di

restrizione di cloni BAC propedeutica alla costruzione di una mappa fisica ad alta risoluzione in *Arabidopsis* (Peters et al., 2001) utilizzando marcatori molecolari AFLP. Questa tecnica presenta due limitazioni principali: l'assenza di una copertura completa del trascrittoma tramite EST e la bassa qualità delle sequenze cDNA/EST presenti nei database pubblici. Queste limitazioni verranno sicuramente superate nel prossimo futuro con l'ottenimento di nuove sequenze.

L'accumularsi di migliaia di EST all'interno delle banche dati è stato un'ottima fonte per l'isolamento di loci microsatelliti. L'identificazione di SSR all'interno delle EST è particolarmente utile poiché questi motivi ripetuti permettono più facilmente l'isolamento di marcatori strettamente associati a geni d'importanza agronomica. Anche se i marcatori SSR derivati dalle EST sono generalmente meno polimorfici, sono però, più facilmente trasferibili alle specie correlate (Cordeiro et al., 2001), inoltre gli SSR disegnati all'interno delle 3' e 5' UTR (untranslated region) possono essere molto utili perché associati a caratteri agronomici e più polimorfici rispetto a quelli interni agli esoni. L'individuazione *in silico* di marcatori SSR è stata utilizzata con successo in diverse specie vegetali: canna da zucchero (Cordeiro et al., 2001), mais, riso, orzo, grano, sorgo (Kantety et al., 2002). Presso il Clemson University Genomics Institute (CUGI) è stato istituito il Genome Database for Rosaceae (GDR), in cui sono presenti 218.181 EST, all'interno di queste sono state individuate più di 1000 sequenze SSR (Jung et al. 2005). In albicocco ed in vite è stata anche testata la trasferibilità degli EST-SSR all'interno della specie (Decrooq et al., 2003), ed in kiwi sono stati utilizzati per costruire una mappa d'associazione (Fraser et al., 2004).

I polimorfismi singoli di sequenza (SNP, Single Nucleotide Polymorphism) sono un altro esempio di marcatori molecolari che possono essere dedotti dall'analisi delle banche dati EST. Uno SNP si definisce come una sostituzione, o in senso più generale anche inserzione o delezione (InsDel), di un singolo nucleotide all'interno di una sequenza rispetto ad una di riferimento; gli EST-SNP sono molto importanti poiché localizzandosi all'interno della regione espressa del genoma possono essere strettamente legati a geni d'interesse. Questo tipo di marcatori sono utili per la MAS, la caratterizzazione di alleli e il mappaggio genico. Gli SNP sono stati utilizzati per la prima nell'uomo e si sono rilevati efficienti per la mappatura fine dei geni (International SNP Map Working Group, 2001; Aerts et al., 2002, Chen et al., 2002) oltre che per l'individuazione di alleli legati a gravi patologie (Verhage et al., 2002; Sugimoto et al., 2002; Margiotti et al., 2002). In diverse specie erbacee gli SNP sono presenti con una frequenza tale da poter diventare dei marcatori efficaci, anche in questo caso gli EST-SNP sono meno polimorfici rispetto agli SNP genomici sviluppati all'interno del trascrittoma,

in mais la frequenza degli SNP in regioni non codificanti è di uno ogni 31 bp, mentre nelle regioni codificanti diventa di uno ogni 124 bp (Ching et al., 2002). In controtendenza è quello che è stato evidenziato, in soia, in 22 genotipi diversi in cui si è riscontrato uno SNP ogni 1,64 kb nelle regioni codificanti e solo uno SNP ogni 4,68 kb nelle regioni non codificanti (Zhu et al., 2003). SNP associati a caratteri di importanza agronomica sono stati individuati in riso, soia e cipolla (Gupta et al., 2002). Uno dei maggiori vantaggi nell'uso degli SNP è la possibilità di automatizzazione della metodica attraverso la primer extension (Giordano et al., 2001) che consente di ridurre il tempo dell'analisi e di aumentarne l'efficienza, mentre il maggior difetto di questa tecnologia è sicuramente il costo poiché necessita dell'uso di strumentazione ad alta definizione. Un altro problema connesso agli SNP può essere l'errore di sequenziamento che a volte si rivela maggiore della frequenza con cui compaiono i marcatori (Wencai et al., 2004).

Gli ESTP (Expressed Sequence Tag Polymorphism) sono dei marcatori in cui dei primer specifici vengono disegnati sulla base della sequenza di un EST di interesse per poi amplificare del DNA genomico. La lunghezza del prodotto di PCR viene rilevata attraverso una semplice corsa elettroforetica, ma anche la sequenza può essere investigata attraverso delle metodologie che si basano su questo principio: come la DGGE (Denaturing Gradient Gel Electrophoresis) e l'SSCP (Single Strand Conformational Polymorphism).

La DGGE: questa tecnica sfrutta il principio per cui la temperatura di denaturazione del DNA (T_m) dipende dalla sua sequenza nucleotidica ed è peculiare di ogni frammento. Variazioni di un singolo nucleotide (mutazioni o polimorfismi) modificano la temperatura di denaturazione del frammento e di conseguenza la sua mobilità. La variazione nella T_m di un dato frammento di DNA può essere controllata su un gel di poliacrilammide a concentrazioni denaturanti crescenti poiché il raggiungimento della temperatura di denaturazione comporta una parziale apertura della doppia elica e quindi un rallentamento nella progressione del frammento sul gel. La reazione di PCR generalmente prevede l'utilizzo di un primer modificato con l'aggiunta di una coda stabilizzante ricca in G e C (GC clamp) che consente di creare artificialmente un dominio ad alta temperatura di denaturazione. L'amplificato del controllo e del campione da analizzare vengono miscelati in parti uguali, denaturati e lasciati riassociarsi lentamente. Una qualsiasi variazione di sequenza tra la molecola originale (wild type) e quella mutata porta alla formazione di eteroduplex (combinazione di due catene di DNA non perfettamente complementari, caratterizzata dalla presenza di un mismatch, un disappaiamento, dove c'è la mutazione e quindi la non perfetta complementarità). Nel caso in cui ci sia una mutazione puntiforme o un'altra nella sequenza, durante la fase di re-

annealing oltre agli omoduplex originali, si generano eteroduplex dalla combinazione di uno strand *wild type* e di uno mutato.

La mobilità di un singolo filamento di DNA su un gel varia considerevolmente a seconda della composizione nucleotidica è su questo fenomeno che di basa la tecnica SSCP. Questa metodologia viene prevalentemente utilizzata per analizzare polimorfismi in singoli loci, primer specifici vengono disegnati in base alla sequenza delle regioni fiancheggianti la regione bersaglio e vengono poi utilizzati per amplificare il DNA genomico. DNA a singolo filamento viene prodotto attraverso un'amplificazione asimmetrica via PCR, in cui un primer è aggiunto in eccesso rispetto all'altro. A causa della bassa concentrazione di uno dei due primer, la PCR amplifica solo uno dei due filamenti, quello che contiene il locus d'interesse. I frammenti vengono, quindi, confrontati tramite elettroforesi. La mobilità dei singoli filamenti è sequenza dipendente a causa della formazione a vari livelli di strutture secondarie dovute ad appaiamenti che portano alla formazione di "loop" ed al compattamento della molecola. Il polimorfismo evidenziato dai marcatori SSCP è di tipo codominante, la codominanza è il risultato di un polimorfismo all'interno della regione amplificata.

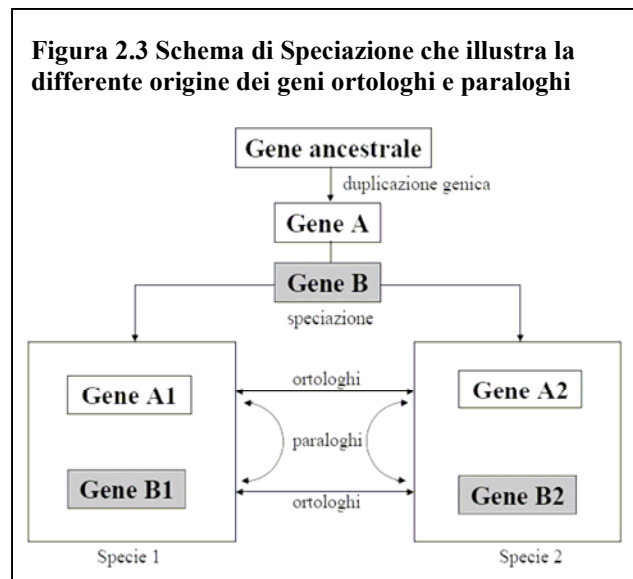
Entrambe queste tecniche non necessitano di informazioni sul tipo di polimorfismo che si sta investigando, sia questo una sostituzione, una delezione o un'inserzione.

Questo tipo di marcatori è stato utilizzato in *Pinus* e *Picea* per ancorare mappe d'associazione di specie differenti (Harry et al., 1998; Brown et al., 2001).

Recentemente grazie al completo sequenziamento del genoma di alcuni organismi ed all'aumento sempre più significativo di EST è stato possibile individuare un set di geni

ortologhi conservati (COS). I **Geni ortologhi** sono quei geni che si sono evoluti, in specie diverse, da un gene ancestrale comune, questi sono tanto più simili quanto più le specie sono evolutivamente vicine. Il fenomeno della speciazione porta alla divergenza dei geni e quindi delle proteine che essi codificano. L' α -globina di uomo e di topo, per esempio, hanno iniziato a divergere circa 80 milioni di anni fa,

quando avvenne la divisione che dette vita ai primati e ai roditori. I due geni sono da considerarsi ortologhi.



I **Geni paraloghi**, invece, sono geni originati dalla duplicazione di un unico gene nello stesso organismo, come per esempio l' α -globina e β -globina umana hanno iniziato a divergere in seguito alla duplicazione di un gene globinico ancestrale. I due geni sono da considerarsi paraloghi.

Confrontando le sequenze genomiche di *Arabidopsis* con la banca dati EST di pomodoro sono stati individuati 1025 COS, questi rappresentano dei geni funzionali che sono stati annotati e che mostrano un elevato grado di conservazione all'interno di tutte le dicotiledoni (Fulton et al., 2002). E' facile intuire come questi marcatori possano essere estremamente utili per la comparazione di mappe di specie piuttosto distanti fra loro e diventare uno strumento valido per stabilire le relazioni filogenetiche all'interno di specie e generi. Più di 550 COS sono stati già mappati in pomodoro ed in futuro lo saranno anche in altre specie.

2.2.3 Sequenze di Geni noti come fonte di Marcatori

Recentemente sono state sviluppate delle nuove classi di marcatori: i GTM (Gene Target Marker) e gli FM (Functional Marker). I GTM vengono sviluppati a partire da sequenze di geni con funzione nota, mentre gli FM derivano da sequenze di geni con funzione ignota o putativa, ma che evidenziano correlazioni strette con caratteri agronomici (Andersen and Lübberstedt, 2003). L'individuazione dei polimorfismi legati a queste classi può essere effettuata utilizzando i marcatori molecolari classici come RAPD, CAP, SSR, RFLP, SNP. La differenza sostanziale è che non sono più polimorfismi ignoti ma sono strettamente legati ad un gene o alla sua funzione. Questo tipo di marcatori sono stati ampiamente utilizzati nelle piante per individuare i geni che controllano diverse vie metaboliche, come in *Brassica napus*, in cui 70 geni di *Arabidopsis* sono stati amplificati con successo riuscendo ad identificare i diversi enzimi chiave nella via biosintetica dell'acido linoleico (Jourden et al., 1996; Brunel et al., 1999; Fourman et al., 2002).

Una delle classi più importanti di marcatori appartenenti a questa categoria sono gli RGA (Resistant Gene Analogue). All'interno dei geni che conferiscono la resistenza ai patogeni nella piante sono state individuate delle regioni conservate: LRR (Leucine Rich Repeat), NBS (Nucleotide Binding Site) e domini serina/treonina chinasi. Primer degenerati sono stati utilizzati per amplificare queste regioni consenso all'interno del DNA genomica e possono essere utilizzati come marcatori molecolari (Mohler et al., 2002). In *Cicer* sono stati utilizzati per individuare marcatori associati con la resistenza a *Fusarium* (Huettel et al., 2002). Una delle maggiori limitazioni di questo tipo di marcatori è che gli RGA rappresentano delle regioni non espresse e quindi non sempre possono essere correlate ad una funzione, ma si

sono dimostrate utili in quelle specie, come il grano, in cui solo 1-5% del genoma viene espresso (Quint et al., 2002).

Anche il DNA degli organelli (mitocondri, cloroplasti e ribosmi) è stato utilizzato sotto questa forma, infatti spesso la sequenza completa dei genomi organellari è disponibile in più specie rispetto a quella completa del genoma. Uno SNP cloroplastico in *Pinus radiata* è stato efficacemente utilizzato per distinguere gli abeti rossi, neri e bianchi (Germano and Klein, 1999). Anche il gene cloroplastico *rbcL* e la RUBISCO (Ribulosio bifosfato co-carbossilasi, enzima chiave della fotosintesi responsabile della carbossilazione) sono stati utilizzati come marcatori sia nelle piante inferiori che in quelle superiori (Shimomura et al., 2002; Savolainen and Chase, 2003).

La più importante classe che appartiene a questi marcatori funzionali è quella dei Microarray. I Microarray offrono per la prima volta la possibilità di una visione globale di ciò che avviene all'interno di un organo, di un tessuto o di una cellula in un dato momento metabolico e temporale, è come riuscire ad ottenere una fotografia istantanea. La tecnica dei Microarray non è altro che un'evoluzione della tecnica di ibridazione, ma due innovazioni chiave l'hanno resa possibile. La prima la si deve al gruppo di Pat Brown (1999) che fissò, roboticamente, 10.000 cDNA su un vetrino da microscopia. La seconda fu l'invenzione di una tecnica per la sintesi ad alta densità di oligonucleotidi ad opera di Steve Fodor (1991).

Un microarray è, quindi, un supporto solido sul quale vengono o fissati o sintetizzati dei corti oligonucleotidi i quali rappresentano sequenze target di differenti geni o alotipi dello stesso gene. Un singolo Microarray può contenere tutti i 100.000 geni dell'uomo. E' chiaro come la potenzialità di questo strumento siano elevatissime, anche se ci sono diverse difficoltà applicative quando utilizzato in specie in cui non è ancora nota tutta la sequenza del genoma. Per quel che concerne il campo vegetale attualmente sono attivi i diversi progetti per la costruzione di microarray per la mappatura di SNP in *Brassica oleracea* (Texas A&M University), per l'identificazione di geni coinvolti nella risposta della pianta agli stress in *Sorghum* (University of Delaware), per l'analisi degli effetti cellulari e molecolari prodotti da stimoli esterni alla pianta in *Arabidopsis* (North Carolina State University), per la comprensione dei meccanismi molecolari relativi allo sviluppo della pianta in *Prunus dulcis* (Michigan State University).

Pur essendo la costruzione del microarray un processo di routine, la vera difficoltà risiede nell'interpretazione dei dati e nella riproducibilità stessa dell'esperimento, anche in questo caso è la bioinformatica che interviene per cercare di normalizzare i risultati ottenuti. Quando il microarray fornisce dei dati che vengono considerati affidabili, come un certa frazione di

geni up-regolati o down-regolati in certi tipi cellulari o in determinati momenti dello sviluppo, questi devono essere validati. La validazione non è altro che l'analisi dell'espressione genica di ogni singolo gene utilizzando delle tecniche standard come il Northern blot, o avanzate come la PCR semi quantitativa (real time). Queste metodiche consentono di verificare l'espressione di uno o più geni in un dato tessuto ad un preciso momento e consentono anche di dare un valore numerico circa la quantità di trascritto.

In conclusione è importante notare come la biologia molecolare abbia fatto degli enormi balzi in avanti e come oggi ci siano sempre maggiori strumenti che permettono di dissezionare finemente i geni e la loro funzione; in pochi anni si sono evolute nuove scienze come la genomica, la trascrittomica e la proteomica che hanno come fine ultimo, quello di osservare i fenomeni biologici nella loro interezza e non fermarsi solo sullo studio del particolare che, anche se fondamentale, a volte si dimostra forviante.

3 LE MAPPE D'ASSOCIAZIONE NEL GENERE PRUNUS

Il genere *Prunus* comprende alcune delle più importanti specie da frutto come ciliegio (*Prunus cerasus* L.), albicocco (*Prunus armeniaca* L.), mandorlo (*Prunus dulcis* Batsch), susino (*Prunus domestica* L.) e pesco [*Prunus persica* (L.) Batsch]. Il miglioramento genetico in Pesco deve tener conto di alcune problematiche: la bassa eterogeneità genetica del germoplasma, utilizzato come base del miglioramento genetico, il periodo infertile e la conseguente differente espressione dei tratti fra le piante giovani e quelle adulte. In questo quadro generale la costruzione di mappe d'associazione diventa un passaggio obbligato per superare alcune di queste difficoltà. La costruzione di mappe "sature", che coprono l'intero genoma con un elevato numero di marcatori molecolari, è particolarmente utile in alcune applicazioni del miglioramento genetico assistito da marcatori (Tanksley et al., 1989).

L'ottenimento di mappe d'associazione diventa particolarmente importante per individuare marcatori strettamente associati a caratteri agronomici, spesso poligenici, permettendo l'individuazione dei QTL (Quantitative Trait Loci). Perché la mappa sia il più informativa possibile è importante utilizzare marcatori codominanti come RFLP, SSR o Isozimi. Un subset di marcatori selezionati da una mappa satura che coprono l'intero genoma ad intervalli regolari e che siano segreganti in un'altra popolazione può essere utilizzato per la costruzione di una mappa ridotta per analisi di segregazione di caratteri agronomici segreganti in quella popolazione. Una volta che la posizione di geni o QTL è nota si può saturare la regione di interesse per identificare marcatori strettamente associati ai caratteri per poi procedere con la MAS o con il clonaggio posizionale. (Tanksley et al., 1995).

Numerose mappe del genere *Prunus* sono state pubblicate negli ultimi dieci anni: in pesco (Chaparro et al., 1994, Rajapakse et al., 1995, Dirlwanger et al., 1998, Lu et al., 1998, Dettori et al., 2001, Foulogne et al., 2003, Verde et al., 2005), utilizzando ibridi pesco x mandorlo (Foolad et al., 1995, Joobeur et al., 1998, Juaregui et al., 2001, Bliss et al., 2002, Aranzana et al., 2003, Dirlwanger et al., 2004), in mandorlo (Viruel et al., 1995, Joobeur et al., 2000), in albicocco (Hurtado et al., 2002, Lambert et al., 2004) ed in ciliegio (Wang et al., 1998).

Nel 1998, nell'ambito di un progetto europeo, è stata ottenuta la mappa di riferimento del genere *Prunus* utilizzando una popolazione F2 derivante dall'autofecondazione di un ibrido pesco x mandorlo (Texas x EarlyGold, T x E). Questa mappa era costituita da 246 marcatori (11 isozimi e 235 RFLP). Gli RFLP utilizzati derivavano da librerie genomiche e cDNA di differenti specie (mandorlo, pesco, ciliegio, susino melo e *P.ferganensis*). La mappa era composta da 8 gruppi d'associazione, come atteso per le *Prunoidee* ($x = 8$), e copriva una

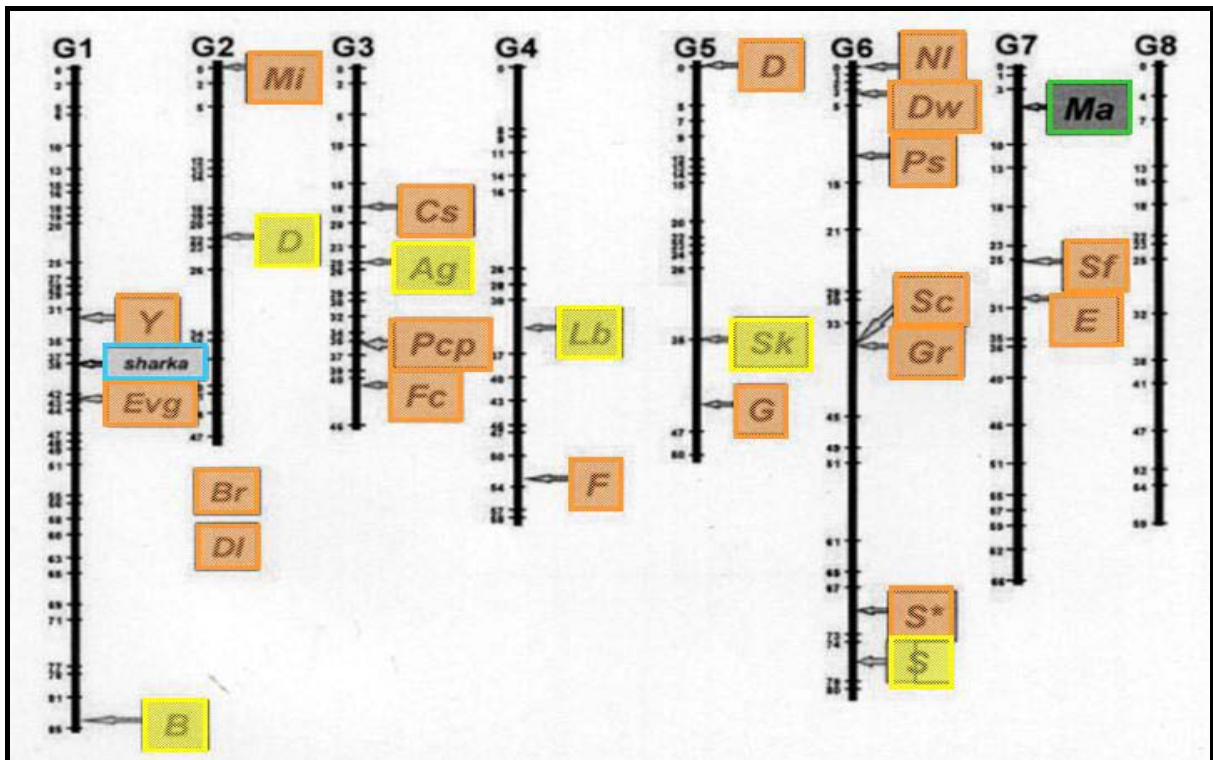
distanza totale di 491 cM con una distanza media fra i marcatori di 2 cM; erano presenti solo 4 intervalli maggiori di 10 cM. Nel 2003 la stessa popolazione F₂ (T x E) è stata analizzata con marcatori SSR (Aranzana et al., 2003); questo tipo di marcatori è codominante, altamente riproducibile (poiché basato sulla tecnica della reazione a catena della polimerasi) e molto polimorfico. Questi nuovi 96 loci SSR sono stati inseriti all'interno della mappa di riferimento T x E arrivando ad un totale di 342 marcatori che coprono una distanza di 522 cM con una distanza media di 2,2 cM fra marcatori. Una novità importante è stata quella dell'individuazione di un gruppo di microsatelliti, che con una distanza media di 25 cM, coprono tutto il genoma offrendo un'utile strumento nelle analisi di fingerprinting e genetiche della specie (Aranzana et al., 2003). Ma la saturazione vera propria della mappa è stata ottenuta nel 2004, la nuova mappa T x E è stata implementata con l'aggiunta di 220 nuovi marcatori (89 SSR, 5 STS, 126 RFLP derivanti da *Arabidopsis thaliana* con elevata omologia con geni di riso). La mappa a tutt'oggi è costituita da 562 marcatori, copre una distanza di 519 cM con una densità di un marcatore ogni 0,92 cM (Dirlewanger et al., 2004). Un numero elevato di marcatori RFLP e soprattutto SSR, utilizzati per la costruzione della mappa T x E, sono stati anche utilizzati per la costruzione di diverse mappe d'associazione all'interno del genere *Prunus*, questi loci in comune vengono detti "ancora" e sono molto importanti per studi di sintenia. Comparando la posizione dei loci ancora nella mappa T x E con le altre 13 costruite all'interno del genere *Prunus* è risultato chiaro che i genomi delle specie diploidi (2n = 16), pesco, mandorlo, albicocco, ciliegio, *P. davidiana*, *P. cerasifera*, *P. ferganensis*, sono sostanzialmente colineari. Le occasionali divergenze fra mappe di specie differenti sono attribuibili alla mappatura di loci RFLP e SSR duplicati che hanno due o più copie in differenti regioni del genoma. Quindi il genere *Prunus*, a livello del genoma, può essere considerato come una singola entità genetica. (Dirlewanger et al., 2004).

Il fine ultimo della costruzione delle mappe d'associazione è quello di localizzarvi caratteri agronomici d'interesse come la resistenza alle malattie, l'epoca di fioritura o i caratteri legati alla qualità della drupa. Nel corso degli ultimi dieci anni sono stati diversi i caratteri individuati all'interno delle mappe d'associazione, ottenute nelle diverse specie appartenenti al genere *Prunus*; grazie all'elevato livello di sintenia all'interno del genere e all'esistenza di una mappa di riferimento un numero considerevole di geni, studiati in popolazioni differenti, è stato integrato in una singola mappa (Figura 3.1, Dirlewanger et al., 2004). In molti casi la localizzazione di questi geni è stata stabilita a partire da mappe a bassa densità di marcatori, quindi la loro posizione potrà essere in futuro definita utilizzando le informazioni provenienti dalla implementazione delle mappe in *Prunus* al fine di individuare dei marcatori addizionali

posizionati nelle regioni d'interesse ed utilizzabili nelle procedure di selezione genetica assistita da marcatori (MAS).

Nella tabella 3.1 sono riportati i principali caratteri agronomici individuati, il gruppo d'associazione e la specie nella quale sono stati mappati.

Figura 3.1 Localizzazione dei 28 major gene mappati in differenti popolazioni di albicocco (celeste), pesco (arancione), mandorlo o mandorlo x pesco (giallo) e Mirabolano (verde) sulla mappa di riferimento del genere *Prunus* (Drlewanger et al., 2004)



Un esempio di MAS è quello dell'auto-incompatibilità gametofitica in mandorlo albicocco e ciliegio. L'Auto-incompatibilità (SI) è un meccanismo per cui il pistillo rigetta il polline che porta il medesimo allele, promovendo l'out-crossing e causando problemi nella produzione frutticola. Questo carattere è codificato da un locus altamente polimorfico (S/s) situato nella porzione distale del Gruppo d'Associazione 6 (G6) (Ballester et al., 1998) ed inizialmente è stato evidenziato con l'uso degli isozimi (Boskovic et al., 1996). Quando fu determinata la sequenza del gene S-RNase al locus S (Tao et al., 1997) furono sviluppati una serie di marcatori specie-specifici e addirittura allele-specifici che permettono una selezione precoce ed accurata dei più comuni alleli di auto-incompatibilità ed auto-compatibilità.

E' noto che i caratteri quantitativi (QTL) costituiscono la maggior parte della variabilità selezionata durante i processi di miglioramento genetico negli alberi da frutto come in molte delle specie coltivate. I caratteri correlati con la crescita, l'architettura della pianta, la produttività, l'epoca di fioritura, di maturazione e la qualità del frutto sono di tipo quantitativo

e sono stati analizzati con l'ausilio di marcatori molecolari in specie differenti. I QTL relativi all'epoca di fioritura, maturazione e alla qualità del frutto sono stati studiati in pesco (Etienne et al., 2002; Dirlewanger et al., 1998). Alcuni dei QTL coinvolti nell'ereditarietà delle componenti responsabili della qualità del frutto ed epoca di fioritura sono stati localizzati in regioni del genoma che coincidono con quelle in

Tabella 3.1 Descrizione dei 28 major gene che determinano caratteri morfologici ed agronomici in differenti specie di *Prunus* e che sono stati collocati all'interno della mappa di riferimento (Dirlewanger et al., 2004)

LG	Carattere	Specie	Citazione
G1	Colore della polpa (<i>Y</i>)	Pesco	Bliss et al. , 2002
	Resistenza Sharka (<i>sharka</i>)	Albicocco	Vilanova et al., 2003
	Evergrowing (<i>Evg</i>)	Pesco	Wang et al.,2002
	Colore del Fiore (<i>B</i>)	Mandorlo x Pesco	Juaregui et al., 1998
G2	Resistenza Root-Knot Nematode (<i>Mi</i>)	Pesco	Bliss et al. , 2002; Juaregui et al., 1998; Lu et al., 1998; Yamamoto et al., 2001 ; Claverie et al., 2004
	Durezza del Nocciolo (<i>D</i>)	Mandorlo	Arus et al., 1998
	Pillar habitus (<i>Br</i>)	Pesco	Scorza et al., 2002
	Fiore Doppio (<i>DI</i>)	Pesco	Chaparro et al., 1994
G3	Colore polpa attorno al nocciolo (<i>Cs</i>)	Pesco	Yamamoto et al., 2001
	Colore delle Antere (<i>Ag</i>)	Mandorlo x Pesco	Joobeur 1998
	Policarpellare (<i>Pcp</i>)	Pesco	Bliss et al. , 2002
	Colore del Fiore (<i>Fc</i>)	Pesco	Yamamoto et al., 2001
G4	Epoca di Fioritura (<i>Lb</i>)	Mandorlo	Ballester et al., 2001
	Adesione della polpa al Nocciolo (<i>F</i>)	Pesco	Dettori et al., 2001; Yamamoto et al., 2001
G5	Bassa Acidità (<i>D</i>)	Pesco	Etienne et al., 2002; Dirlewanger et al., 1998/1999
	Sapore del gheriglio (<i>Sk</i>)	Albicocco	Bliss et al. , 2002
	Pesco/Nettarina (<i>G</i>)	Pesco	Bliss et al. , 2002; Dirlewanger et al., 1998/1999
G6	Forma delle foglie (<i>Nl</i>)	Pesco	Yamamoto et al., 2001
	Altezza della pianta (<i>Dw</i>)	Pesco	Yamamoto et al., 2001
	Maschio Sterilità (<i>Ps</i>)	Pesco	Dirlewanger et al., 1998
	Forma del frutto (<i>S*</i>)	Pesco	Dirlewanger et al., 1998/1999
	Auto-incompatibilità (<i>S</i>)	Mandorlo Albicocco	Bliss et al. , 2002 ; Ballester et al., 1998 Vilanova et al., 2003
G6-	Colore della buccia (<i>Sc</i>)	Pesco	Yamamoto et al., 2001
G8	Colore delle foglie (<i>Gr</i>)	Pesco	Yamamoto et al., 2001 ; Juaregui et al., 1998; Claverie et al.,2004
G7	Resistenza Root-Knot Nematode (<i>Ma</i>)	Susino Mirabolano	Claverie et al.,2004
	Resistenza all'oidio (<i>Sf</i>)	Pesco	Dettori et al., 2001
	Ghiandole fogliari (<i>E</i>)	Pesco	Dettori et al, 2001

cui sono stati posizionati i major gene, come il gene D/d responsabile della bassa acidità in pesco ed il gene Lb/lb che determina l'epoca di fioritura in mandorlo (Ballester et al., 2001). Per l'individuazione dei QTL attraverso l'uso dei marcatori molecolari è stato anche utilizzato con successo l'approccio dei "geni candidati" in cui si cerca di correlare polimorfismi di geni con sequenza e funzione nota e tratti fenotipici. Questa metodologia è stata impiegata per mettere in evidenza le relazioni fra dei geni coinvolti in importanti vie metaboliche e major

gene o QTL. In pesco uno dei 12 geni candidati coinvolti nelle vie metaboliche degli zuccheri e degli acidi organici in frutto mappa nella medesima regione di un QTL. Questo gene *PRUpe* è una vacuolo pirofosfatasi che è coinvolta nella definizione di un gradiente elettrochimico all'interno del tonoplasto, e potrebbe essere implicata nell'accumulo del saccarosio all'interno del vacuolo.

Un'evoluzione nelle mappe d'associazione è rappresentata dalla tecnica della mappatura selettiva (Vision et al., 2000). Questa metodica è composta da due passaggi: il primo è la costruzione di una mappa d'associazione satura ed il secondo consiste nell'aggiunta di nuovi marcatori, con minor precisione, utilizzando un ristretto numero di individui altamente informativi. La selezione degli individui è basata sul numero e la posizione dei siti di ricombinazione; tutti i siti di ricombinazione di una popolazione, sono definiti da sottoinsieme di piante che viene chiamato *bin*. Utilizzando questa tecnica, ed analizzando solo 6 individui, sono stati recentemente localizzati 264 nuovi loci SSR sulla mappa di riferimento delle *Prunoidee* (Howard et al., 2005); questa procedura è in grado di fornire delle informazioni molto importanti, in tempi brevi ed a costi ridotti, a tutti quei gruppi che lavorano con le drupacee.

I risultati ottenuti in questi dieci anni dimostrano come la costruzione di mappe di associazione non sia solo auspicabile, ma sia un passaggio obbligato per i nuovi programmi di miglioramento genetico; in oltre l'elevata sentina delle specie del genere *Prunus* favorisce un efficiente trasferimento di informazioni all'interno della comunità scientifica e rende possibile applicare le conoscenze ottenute in campo commerciale e tecnico.

4 MATURAZIONE E CONSISTENZA DELLA POLPA

La maturazione del frutto è un processo complesso e geneticamente programmato che culmina con dei drastici cambiamenti per quello che riguarda la tessitura, il colore, l'aroma ed il sapore della polpa. Questi cambiamenti sono dovuti alla distruzione enzimatica della parete cellulare, all'idrolisi dell'amido, all'accumulo di zuccheri e alla scomparsa di acidi organici e di composti fenolici. A causa dell'importanza economica che questi eventi hanno sulla conservabilità e trasportabilità del frutto, negli anni sono stati oggetto di studio dal punto di vista biochimico e genetico (Alexander et al., 2002).

I frutti possono essere catalogati in due grandi gruppi a seconda del loro meccanismo di maturazione: frutti Climaterici, in cui la maturazione è accompagnata da un picco di produzione di Etilene e di CO₂; e quelli Non-Climaterici, per i quali non si osservano dei picchi di respirazione, né cambiamenti nella quantità di etilene sintetizzato. Sono frutti Climaterici il pomodoro, la banana, il melo, il pesco, il melone, il pero, il susino; la maturazione, in questi frutti, è strettamente dipendente con la sintesi di uno degli ormoni più importanti nelle piante, l'Etilene (Oeller et al., 1991; Picton et al., 1993; Theologis et al., 1993; Ayub et al., 1996).

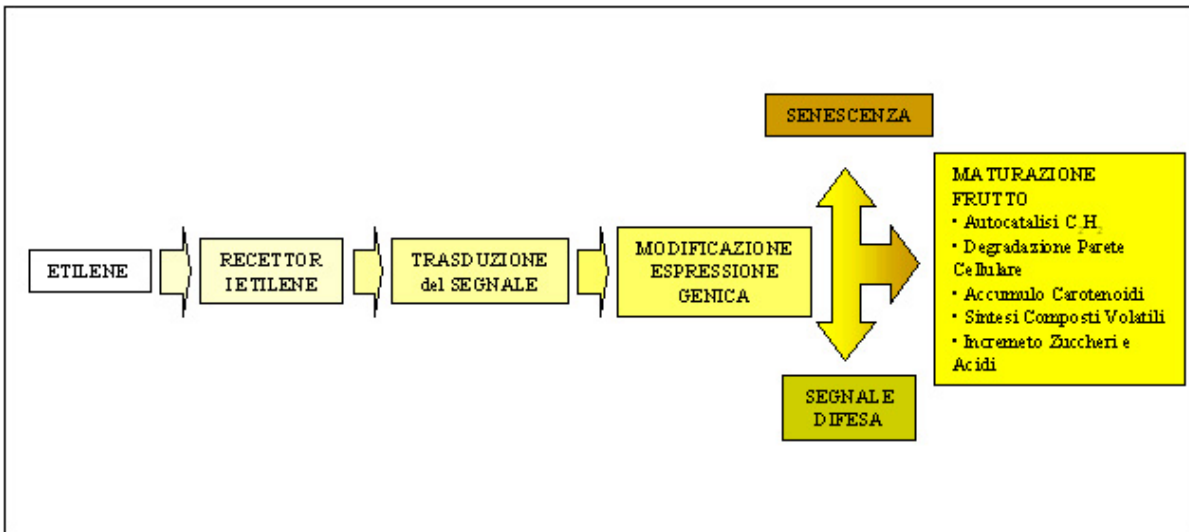
4.1 L'Etilene e la regolazione genica della maturazione

L'Etilene può essere prodotto in quasi tutti i tessuti delle piante superiori, anche se il tasso di produzione dipende sia dal tipo di tessuto sia dallo stadio di sviluppo nel quale si trova. Come già citato la produzione di etilene aumenta durante alcuni processi biologici come l'abscissione delle foglie, la senescenza dei fiori e la maturazione del frutto. Inoltre alcuni tipi di traumi o stress possono promuovere la biosintesi dell'etilene: danni da freddo, infezioni di patogeni, ecc. Essendo l'etilene un gas viene facilmente rilasciato dai tessuti diffondendosi fra gli spazi intercellulari fino all'esterno della pianta.

All'interno delle piante climateriche sono stati proposti due sistemi di regolazione della biosintesi dell'etilene. Il sistema 1 è attivo durante la normale crescita vegetativa, è etilene auto-inibito ed è responsabile della produzione dei livelli basali di etilene che si riscontrano in tutti i tessuti ivi compresi quelli delle piante non climateriche. Il sistema 2 agisce durante la maturazione, nei frutti climaterici, e nella senescenza di alcuni petali, in questo caso la produzione dell'etilene è autocatalitica. La maturazione inizia in una regione del frutto e poi l'etilene viene velocemente trasmesso ai tessuti circostanti dando il via ad un meccanismo a cascata.

La specie climaterica in cui la via biosintetica dell'etilene è meglio caratterizzata è il pomodoro, il *Lycopersicon esculentum*, che viene considerato una specie modello grazie al suo genoma di piccole dimensioni ed al breve tempo che intercorre fra una generazione e l'altra.

Figura 4.1 Rappresentazione schematica del ruolo dell'Etilene durante la maturazione del frutto



L'Etilene è sintetizzato a partire dalla metionina passando per due intermedi, la S-adenosil-L-metionina (AdoMet) e l'Acido 1-aminociclopropano 1-carbossilico (ACC). L'ACC è formato a partire dall'AdoMet attraverso l'azione dell'ACC Sintasi (ACS), quindi l'ACC viene convertito in Etilene dall'azione dell'ACC Ossidasi (ACO) (Kende, 1993) (Figura 4.1).

Dalla conversione dell' AdoMet la ACS produce anche la 5'-metiltioadenosina che attraverso il ciclo di Yang viene utilizzata per la sintesi della metionina; questa via alternativa di sintesi della metionina permette di attivare la biosintesi dell'etilene anche se il pool di metionina è ridotto.

E' noto che la produzione dell'etilene è soggetta sia ad un controllo positivo, sia a quello negativo (Kende, 1993). La regolazione positiva è propria dei frutti in maturazione e dei fiori senescenti , che se esposti ad etilene o propilene esogeni, mostrano un incremento della sintesi dell'etilene dovuto all'induzione dell'ACS e dell'ACO. Entrambi questi enzimi sono codificati da una ridotta famiglia multigenica e la loro espressione differenziale è sotto il controllo di differenti segnali ambientali, di fattori legati allo sviluppo e di ormoni (Kende, 1993; Zarebinski and Theologis, 1994; Barry et al., 2000; Llop-Tous et al., 2000).

In pomodoro sono stati individuati almeno 8 geni che codificano per l'ACS, e altri sono stati identificati in limone (non-climaterica), melone e banana (Yamamoto et al, 1995;Li et al., 1999). Dallo studio della struttura e sequenza aminoacidica dell'ACS si è dedotto che potrebbe agire formando dei dimeri (Tarum and Theologis,1998). In pomodoro si è cercato di comprendere quali delle quattro ACS presenti nel frutto (*LeACS1A*, *LeACS2*, *LeACS4*, *LeACS6*) agissero all'interno del sistema 1 o del sistema 2, in altre parole comprendere quali ACS sono sotto il diretto controllo dell'etilene. Due mutanti, il Never ripe (*Nr*), insensibile all'etilene a causa di una mutazione nel

dominio recettore (Lanahan et al., 1994; Wilkinson et al., 1995), ed il ripening inhibitor (*rin*), non in grado di attivare l'autocatalisi dell'etilene (Herner and Sin, 1973), sono stati utilizzati a questo scopo. I risultati ottenuti evidenziano che LeACS1A e LeACS6 appartengono al sistema 1 di biosintesi e cioè sono costitutivamente espressi in questi tessuti cellulari, mentre LeACS2 e LeACS4 sono sensibili all'autocatalisi dell'etilene. Inoltre se LeACS2 viene silenziata, utilizzando trascritti antisenso, si osserva una drastica riduzione della respirazione e di tutti i fenomeni correlati con la maturazione. (Oeller et al., 1991). Non è ancora chiaro se la regolazione genica dell'ACS sia trascrizionale, post-trascrizionale o post-traduzionale. Un'ipotesi, che è stata confermata solo in vitro, vede nella fosforilazione (post-traduzionale) un meccanismo di protezione dell'enzima dalla degradazione (Tatsuki and Mori, 2001).

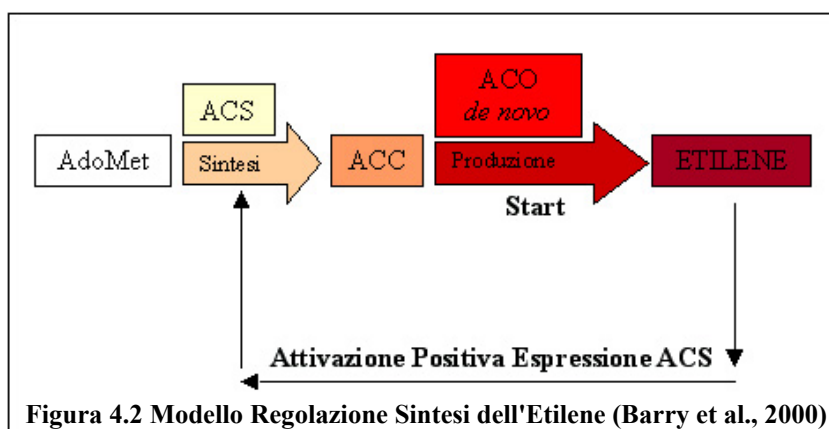


Figura 4.2 Modello Regolazione Sintesi dell'Etilene (Barry et al., 2000)

Inizialmente era stato ipotizzato che fosse l'ACS l'enzima chiave della via biosintetica dell'etilene e che l'ACO fosse costitutivamente espresso (Yang and Hoffman, 1984), ma recentemente il ruolo dell'ACC Ossidasi è stato rivisto, infatti l'aumento dell'attività dell'ACO

precede quello dell'ACS (Lui et al., 1985). Anche la funzione di questo enzima è stata verificata in vitro attraverso l'espressione eterologa del gene in oociti di *Xenopus* ed in lievito (Hamilton et al., 1991; Spanu et al., 1991). Anche l'ACC Ossidasi appartiene ad una famiglia multigenica e nel corso degli anni sono stati isolati geni ACO in diverse specie fra cui anche il pesco (Ruperti et al., 2001). Anche in questo caso sono stati individuati due geni PP-ACO1, fortemente espresso durante la maturazione, e PP-ACO2, espresso in diversi tessuti vegetali e non-autocatalitico. In pomodoro sono stati individuati 4 geni codificanti per ACO, ancora una volta due di questi rispondono alla presenza dell'etilene e due no, a dimostrare l'effettiva presenza di vie biosintetiche regolate in modo differenziale a seconda del tessuto (Barry et al., 1996; Nakatstuka et al., 1998).

Analisi dell'espressione di ACO durante la maturazione, indicano che ogni gene della famiglia è altamente regolato e trascritto a diversi livelli e a differenti stadi di sviluppo (Barry et al., 1996). Sembra verosimile che il primo passaggio nella via biosintetica dell'etilene sia la sintesi *de novo* di ACO, l'etilene così prodotto induce l'ACC Sintasi che a sua volta inizia a processare AdoMet per ottenere ACC e che causa una nuova produzione di etilene ricominciando il ciclo (Figura 4.2).

4.1.1 Recettori per l'Etilene e traduzione del segnale durante la maturazione

Come è stato delucidato nel paragrafo precedente, durante la maturazione dei frutti climaterici la sintesi autocatalitica dell'etilene coordina ed accelera i processi della maturazione. Il meccanismo della trasduzione del segnale non è ancora del tutto chiaro; è noto che i recettori per l'etilene sono omologhi ai recettori batterici a due componenti, questi ultimi sono formati da un dominio che recepisce il segnale e da un dominio recettore che risponde ad esso, questo tipo di meccanismo consente ai batteri di rispondere a differenti stimoli ambientali (Chang and Stewart, 1998). Tutti i recettori per l'etilene hanno il dominio recettore in cui si riconoscono una porzione transmembrana, un dominio GAF, ed un residuo istidina chinasi. La ligazione dell'etilene sul recettore è mediata dal rame, che suggerisce la fosforilazione come segnale a valle; la presenza dell'etilene causa un aumento transiente della fosforilazione di polipeptidi (Raz and Flhur, 1993). Una proteina (CTR) simile ad una MAP3K (Proteina mitogen attivata chinasi chinasi chinasi) è stata infatti individuata come altro componente della via di trasduzione del segnale (Kieber et al., 1993). CTR è un regolatore negativo della via, infatti mutanti CTR1, in cui la proteina non è funzionante, rispondono costitutivamente all'etilene anche in sua assenza (Figura 4.3).

In pomodoro sono state individuate almeno 6 famiglie di recettori per l'etilene, anche in questo caso, come avviene per ACS e ACO, i diversi geni sono espressi in tessuti ed in momenti dello sviluppo differenti. Alcuni recettori sono a loro volta etilene indotti, l'aumento dei recettori causa una diminuzione della sensibilità all'etilene supportando anche la teoria di un modello di regolazione negativa dell'etilene mediato dai recettori stessi (Ciardi et al. 2000). Inoltre nonostante le loro strutture differenti i recettori per l'etilene si sono dimostrati funzionalmente ridondanti. I recettori finora citati sono collocati sulla membrana citoplasmatica, la CTR nel citosol attiva una cascata di fosforilazione che va a colpire, in modo ancora sconosciuto, un recettore di membrana nucleare (EIN). EIN a sua volta attiva la trascrizione del ERF (Ethylene Responsive Factor) che funge, a sua volta, da fattore di trascrizione dei geni etilene-dipendenti (Figura 4.4) ed il processo della maturazione dei frutti. (Alexander and Grierson, 2002)

Figura 4.3 Modello di trasmissione del segnale dal recettore al CTR in presenza di etilene, in assenza di etilene ed in presenza di etilene in un mutante con recettore non funzionante

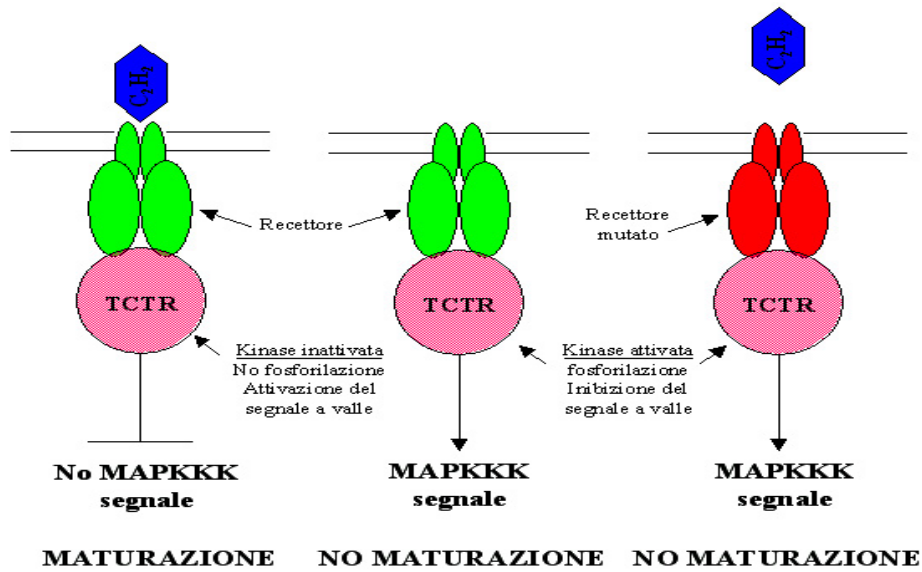
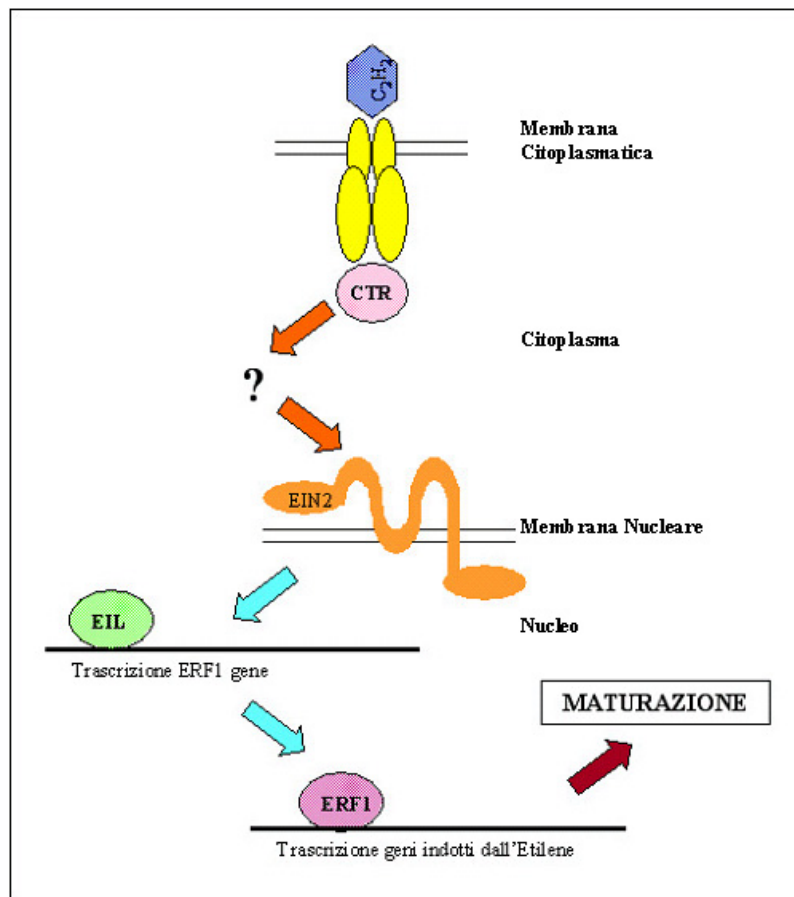


Figura 4.4 Diagramma della possibile via di trasduzione del segnale in frutti durante la maturazione



4.2 Geni coinvolti nell'ammorbidimento della polpa

A causa della loro importanza per la trasportabilità e la conservazione post-raccolta del frutto, gli enzimi che sono implicati nell'ammorbidimento della parete cellulare sono stati ben studiati, sia *in*

vivo sia *in vitro* (Harker et al., 1997). Durante la maturazione i cambiamenti nella consistenza della polpa e di conseguenza, il suo ammorbidimento sono dovuti in gran parte al parziale smantellamento della parete cellulare. Con il procedere della maturazione, la parete delle cellule diventa sempre più idratata attraverso l'idrolisi parziale delle pectine. I cambiamenti nella coesione delle pectine sono responsabili delle modificazioni nella consistenza della polpa caratteristica di un frutto maturo.

I principali geni responsabili dell'ammorbidimento della polpa in pesco, sotto il controllo positivo dell'etilene, sono: le Endopoligalatturonasi (*endo*-PG), le β -galattosidasi (β -gal), la Pectinmetilesterasi (PME), le Pectato liasi (PL) e le Espansine (Exp).

E' stato evidenziato come le poligalatturonasi (PG) siano le maggiori responsabili della degradazione dei poliuronidi della parete e che la loro trascrizione avvenga durante la maturazione (DellaPenna et al, 1989; Montgomery et al, 1993; Orr and Brady, 1993). Il promotore genico delle PG contiene un elemento etilene dipendente (Nicholass et al, 1995) e l'analisi di pomodori transgenici sensibili a bassi livelli di etilene dimostrano l'attivazione delle PG anche a queste condizioni (Sitrit and Bennett, 1998).

Durante la maturazione le PME sono responsabili della de-esterificazione dei poligalatturonidi della parete cellulare, che sono altamente metil-esterificati; questa modificazione rende i poliuronidi suscettibili alle PG (Koch and Nevins, 1989; Carpita and Gibeaut, 1993).

La degradazione della parete cellulare avviene anche attraverso l'azione delle β -galattosidasi che smantellano i polimeri di galattosio facendo incrementare, durante tutta la maturazione, il galattosio monomero libero all'interno del citoplasma.

Una classe importante di enzimi, che ha il ruolo di rendere disponibili le pectine per le successive degradazioni, è rappresentato dalle Pectato liasi; è stato osservato come, in pesco, ne agiscano almeno due e che la loro azione sia temporalmente anticipata rispetto a quella dei tre enzimi finora citati (Trainotti et al., 2003), che quindi probabilmente necessitano dell'azione delle PL per poter avviare la loro azione di smantellamento.

Dei differenti enzimi che presiedono alla degradazione delle emicellulose, solo le espansine vengono fortemente espresse durante la maturazione (Hyama et al., 2000); questo conferma l'ipotesi che la degradazione delle pectine prevalga sulla degradazione delle emicellulose.

4.3 Pesche Melting, Non-Melting e Stonyhard

L'ammorbidimento della polpa in pesco avviene in due fasi: una, all'inizio del processo della maturazione, in cui si osserva un lento declino della consistenza del frutto e che avviene, e, l'altra che coincide con la fine del processo in cui si ha un rapido ammorbidimento (fase di melting)

(Pressey et al., 1971). Il fenotipo fondente (melting flesh) è associato ad un elevato incremento delle pectine solubili e ad una progressiva depolimerizzazione delle pectine nella seconda fase (Pressey et al., 1971; Dawson et al., 1992; Fishman et al., 1993). Nelle pesche con fenotipo croccante (non-melting flesh) la fase finale di melting è assente, le pectine sono sottoposte ad una depolimerizzazione minima e quindi i frutti rimangono relativamente consistenti anche in piena maturazione (Pressy and Avanti, 1978; Fishman et al., 1993). L'accumulo di trascritti relativi alle *endo*-PG viene riscontrato solo nelle pesche con fenotipo melting e coincide con la fase drastica di perdita di consistenza (Pressy and Avant, 1978; Orr and Barry, 1993). L'assenza della fase di melting nelle pesche con fenotipo croccante sembra essere dovuta ad una delezione nel gene delle *endo*-PG che causa la sintesi di una proteina tronca e non funzionante (Lester et al., 1994, 1996; Callahan et al., 2004).

Anche il fenotipo 'stony hard' riguarda la consistenza della polpa ed è stato descritto per la prima volta da Yoshida nel 1976. Questa tipologia di pesche ha la caratteristica di avere una polpa estremamente consistente a causa della mancata sintesi dell'etilene (Haji et al., 2001). Analisi biochimiche hanno evidenziato come le pesche 'stony hard' producono etilene a livelli prossimi allo zero.

Sia il fenotipo croccante che 'stony hard' sono molto importanti economicamente, in quanto presentano un'elevata resistenza alle manipolazioni ed consentono tempi di conservazione più lunghi rispetto a quelle del tipo fondente. Recentemente sono stati pubblicati degli importanti risultati che dimostrano la teoria di Monet (1989), secondo la quale i caratteri consistenza della polpa ed aderenza della polpa al nocciolo sono differenti forme alleliche dello stesso locus. Quindi le endopoligalatturonasi sono responsabili anche dell'aderenza della polpa (Peace et al., 2005). Entrambi i caratteri, consistenza della polpa e stony hard, sono due caratteri sono monogenici recessivi, inoltre il carattere Melting è allelico del carattere aderenza della polpa al nocciolo (Peace et al., 2005). Ad oggi il carattere stony hard non è stato localizzato in nessuna delle numerose mappe del genere *Prunus*, mentre il carattere consistenza della polpa è stato mappato come aderenza della polpa al nocciolo (Dettori et al., 2001)

5 SCOPO DEL LAVORO

Il pesco [*Prunus Persica* (L.) Batsch] è una delle colture più importanti tra le specie arboree da frutto. L'attività di miglioramento genetico del pesco è stata da sempre molto intensa ed ha prodotto, negli anni, un notevole numero di nuove cultivar, basti pensare che solo fra il 1990 e il 1996 sono state introdotte sul mercato oltre 500 nuove varietà.

Il miglioramento genetico effettuato fino ad ora ha certamente assicurato un innalzamento degli standard agronomici e pomologici delle cultivar, ma ha anche messo in evidenza una drastica riduzione della possibilità di migliorare ulteriormente le caratteristiche già presenti nelle attuali cultivar.

La variabilità genetica presente nel germoplasma di pesco è stata ormai ampiamente sfruttata e le nuove cultivar hanno caratteristiche sempre più simili tra di loro.

Pur esistendo delle cultivar di pesco che possiedono delle ottime caratteristiche qualitative, per esigenze di stoccaggio e commercializzazione, la raccolta dei frutti non viene effettuata allo stadio di maturazione fisiologica, ma anticipata notevolmente. Questa pratica causa la perdita di alcune caratteristiche organolettiche del frutto impedendone la completa espressione.

La nascita di una nuova consapevolezza del consumatore verso tutti i prodotti alimentari, e della frutta in particolare, impone lo sviluppo di nuove varietà che abbiano delle buone caratteristiche organolettiche, una polpa consistente, croccante e succosa, livelli elevati di vitamine, di composti antiossidanti e di altri nutraceutici.

In questo quadro generale bisogna considerare il notevole apporto che hanno dato e che possono dare le biotecnologie al miglioramento genetico tradizionale; infatti le tecniche di genetica molecolare permettono di ottenere informazioni importanti sulla modalità di trasmissione dei caratteri, di investigare e conoscere la sequenza e la funzione dei singoli geni e di fornire una visione d'insieme di ciò che accade a livello genico in un determinato momento.

Una delle possibili applicazioni delle informazioni ottenute dalle mappe di associazione, è offerta dalla MAS (Marker Assisted Selection); infatti individuando dei marcatori molecolari strettamente associati a caratteri di importanza agronomica si può procedere ad una selezione precoce dei sementali con un notevole risparmio di tempo, di spazio e quindi indirettamente di costi.

L'obiettivo principale è quello di individuare ed analizzare i caratteri monogenici e poligenici che concorrono a determinare la qualità del frutto. In particolare, verrà preso in considerazione il carattere "stony hard". Questo carattere, di tipo monogenico recessivo (hd/hd) (Yoshida M. 1976, Scorza R. & Sherman W.B. 1996), conferisce una croccantezza alla polpa che viene oggi particolarmente perseguita nei programmi di Miglioramento genetico, non solo per i vantaggi riscontrati per la conservabilità e la trasportabilità dei frutti (Liverani et. al. 2002), ma anche perché

consente di effettuare la raccolta allo stadio di maturazione fisiologica con conseguente esaltazione delle qualità organolettiche del frutto.

A questo scopo verrà utilizzata una popolazione F2 ottenuta presso l'Istituto Sperimentale per la Frutticoltura di Forlì, dall'autofecondazione di un semenzale F1 derivato dall'incrocio Yumyeong x O'Henry. La cultivar Yumyeong è portatrice del carattere "stony hard".

Per raggiungere tale scopo è necessario procedere a:

- 1) Costruzione della mappa di associazione della popolazione F2 Yumyeong X O'Henry (fornita dalla sezione di Folri dell'Istituto Sperimentale per la Frutticoltura) utilizzando le seguenti tipologie di marcatori molecolari: SSRs, AFLPs e SNPs.
- 2) Individuazione di EST correlate alla qualità, analizzando librerie di cDNA ottenute da frutti a differenti stadi di maturazione
- 3) Isolamento e caratterizzazione del gene responsabile del carattere "stony hard" tramite la tecnica del cDNA-AFLP Differential Display

I risultati ottenuti nell'ambito della ricerca proposta consentiranno di rendere disponibili, a quanti operano nel campo del miglioramento genetico, nuovi marcatori molecolari (SNP, AFLP, SSR) che siano di facile utilizzo per la predizione e la selezione precoce dei caratteri relativi alla qualità del frutto.

Inoltre la comunità scientifica potrà utilizzare gli EST prodotti ed il gene individuato per studi di funzionalità ed espressione (microarray) o per analisi filogenetiche.

6 MATERIALI e METODI

6.1 COSTRUZIONE MAPPA D'ASSOCIAZIONE

6.1.1 Materiali Vegetali

Una popolazione F_2 (O'H x Y) di 57 individui è stata utilizzata per la costruzione della mappa d'associazione. La popolazione deriva dall'autofecondazione del semenzale F_1 H51 ottenuto dall'incrocio O'Henry x Yumyeong. Yumyeong è il parentale portatore del carattere "stony hard", è una pesca a polpa bianca ed aderente. O'Henry è una pesca a polpa gialla, spicca e fondente, eterozigote per il carattere pesca/nettarina (Okie, 1998).

Gli apici fogliari e le foglie giovani, non ancora espanse, sono stati prelevati da tutti gli individui della F_2 , dei parentali, dell'ibrido F_1 e conservate a -80 °C.

Il DNA è stato estratto da 1gr di tessuto fogliare utilizzando il Nucleon™ PhytoPure™ Genomic DNA Extraction Kits (Amersham Biosciences), i campioni sono stati quantificati (utilizzando un fluorimetro DyNA quantum™ 200, Hoefer) e diluiti fino ad una concentrazione finale di 10 ng/μl.

6.1.2 Caratteri Morfologici

Nella popolazione O'H x Y sono stati valutati 4 caratteri morfologici:

1. Colore della polpa: Polpa Bianca/Polpa Gialla Y/y, (Connors 1920). Studi sulla segregazione del carattere "colore della polpa" indicano che si tratta di un carattere mendeliano semplice, e che il colore bianco è dominante sul giallo (Hesse 1975). Questa evidenza suggerisce l'ipotesi che il gene *Y* codifichi per un enzima coinvolto nella degradazione dei carotenoidi, come la famiglia delle carotenoide ossigenasi (Giuliano et al, 2003).
2. Aderenza della polpa: Spicca/Aderente F/f, (Bailey & French, 1941-1949). Il carattere adesione della polpa è descritto come monogenico; la polpa spicca si stacca agilmente dal nocciolo è dominante rispetto a quella aderente. Già Monet nel 1989 e recentemente Peace (Peace et al., 2005) descrivono questo carattere come strettamente correlato con la consistenza della polpa e con l'attività delle endopoligalatturonasi.
3. Consistenza della polpa: Polpa Fondente/ Polpa "stony hard" Hd/hd, (Yoshida, 1976) Il carattere "stonyhard" è un carattere monogenico recessivo, che conferisce croccantezza e conservabilità alla polpa. Non sono del tutto note le sue relazioni con il carattere "Melting".

4. Morfologia della buccia: Pesca/Nettarina G/g, (Blake, 1932)

6.1.3 Marcatori Microsatelliti (Simple Sequence Repeats - SSRs)

I Microsatelliti sono fra i marcatori molecolari quelli più utilizzati per l'analisi genetica poiché sono codominanti, altamente ripetibili e di facile valutazione anche se richiedono una complicata fase per l'isolamento.

Per la costruzione della mappa d'associazione sono state utilizzate diverse coppie di primer microsatelliti, reperite in bibliografia, provenienti da specie appartenenti al genere *Prunus* (Tab. 5.1) quali: Mandorlo (*Prunus dulcis*), Pesca (*Prunus persica*), Albicocco (*Prunus armeniaca*), Ciliegio Selvatico (*Prunus avium*), Susino Giapponese (*Prunus salicina*).

Tabella 6.1 Sigla, specie e quantità delle coppie di primer SSR testate sui parentali della popolazione F2 Y x O'H; quantità di DNA e Magnesio utilizzati per le reazioni di PCR

SSR sigla	Specie	N° SSR testati	ng DNA per reaz.	mM MgCl ₂ per reaz.	Riferimento
BPPCT	Pesco	42	40	3	Dirlewanger et al. (2002)
CPDCT	Mandorlo	33	20	1.5	Mnejia et al. (2005)
CPPCT	Pesco	35	40	3	Aranzana et al. (2002)
CPSCT	Susino Giapponese	34	40	3	Mnejia et al. (2004)
EMPaS	Ciliegio Selvatico	11	20	3	Vaughan et al. (2004)
MA/M	Pesco	34	20	1.5	Yamamoto et al. (2002)
PaC	Albicocco	10	20	2	Decroocq et al. (2003)
PaCITA	Albicocco	12	40	1.5	Lopes et al. (2002)
pchgms pchcms	Pesco	24	40	3	Sosinski et al. (2000) Georgi et al. (2003) Wang et al. (2002b) Verde et al. (2004)
PS		6	40	3	
UDA	Mandorlo	43	40	3	Testolin et al. (2004)
UDAp	Albicocco	26	20		Messina et al. (2004)
UDP	Pesco	15	40	3	Testolin et al. (1999)
Totale		325			

Le amplificazioni sono state eseguite in un volume finale di 10 µl utilizzando, per ogni reazione, dai 20 ai 40 ng di DNA genomico (Tab. 5.1), 20 mM Tris-HCl (pH 8,4), 50 mM KCl, 1.5-3 mM MgCl₂ (Tab. 5.1), 0.2 mM dNTPs, 0.2 µM per primer e 0.5 U di Platinum[®] GenoType *Tsp* DNA Polymerase (Invitrogen). Il tipo di polimerasi scelto è stato appositamente studiato per l'amplificazione di sequenze microsatelliti. Le amplificazioni sono state eseguite utilizzando un termociclatore GENEAMP[®] PCR System 9600 (Perkin-Elmer) seguendo i cicli di reazione e le temperature di annealing pubblicati nei lavori di riferimento citati. I prodotti di PCR ottenuti sono stati separati con elettroforesi orizzontale su gel d'agarosio al 3% in una soluzione di TBE 1X applicando un voltaggio pari a 5 V/cm, quindi i frammenti sono stati visualizzati tramite

colorazione con bromuro d'etidio (BrEt). I gel sono stati fotografati con una macchina analogica utilizzando pellicole Kodak (F9020 Sigma Kodak black and white film) ed in seguito utilizzando un acquirente d'immagini digitale (GEL Logic 100 Imaging system, Kodak). L'agarosio utilizzato è il MetaPhor® (Cambrex), questo tipo di supporto, ad alta definizione, consente di separare frammenti in un intervallo compreso fra le 200 e le 800 paia di basi che differiscono fra loro anche di solo il 2%.

I primer sono stati inizialmente saggiati solo su Yumyeong, O'Henry e F₁ H51, quelli risultati polimorfici e segreganti sono stati testati su tutta la popolazione.

6.1.4 Amplified fragment length polymorphisms (AFLPs)

Gli Amplified Fragment Length Polymorphisms (AFLPs) sono dei marcatori estremamente polimorfici anche se dominanti e quindi meno informativi rispetto ai microsatelliti, sono stati utilizzati a causa del basso polimorfismo evidenziato con gli SSR testati in precedenza.

Gli AFLP sono stati visualizzati con due diversi apparati elettroforetici: un Sequenziatore capillare CEQ™ 8000 Genetic Analysis System (Beckman Coulter) e un apparato per elettroforesi verticale Sequi-Gen GT Sequencing Cell (Biorad).

In entrambi i casi è stato applicato il protocollo di Vos (Vos et al., 1995) con delle opportune modifiche, che riguardano, in particolare, la quantità di DNA (100 ng invece di 500) e la concentrazione dei primer che è stata modificata per le reazioni visualizzate tramite il sequenziatore. Il CEQ 8000 ha un'elevata sensibilità, la tecnologia con cui è stato costruito prevede l'utilizzo di primer marcati con 4 fotocromi speciali i Well Dyes (Beckman Coulter) che consentono di diminuire la quantità di primer marcato di un fattore 6.

Il DNA genomico è stato digerito contemporaneamente con due enzimi di restrizione (EcoR1 e MseI). Alle estremità dei frammenti sono stati ligati degli adattatori opportunamente disegnati; gli adattatori hanno la funzione di distruggere i siti di restrizione presenti e di rendere possibile l'amplificazione con coppie di primer con sequenza nota.

Dopo la restrizione e la ligazione è stata effettuata la reazione di preamplificazione che consiste in una reazione di PCR in cui i prodotti della ligazione vengono usati come templati ed i primer (denominati Eco + N e Mse + N) presentano la sequenza complementare a quella degli adattatori ed una base arbitrariamente scelta. In questo caso le preamplificazioni sono state effettuate utilizzando Eco + A e Mse + C.

Successivamente le preamplificazioni sono state, a loro volta, utilizzate come template per delle amplificazioni selettive, sono stati scelti primer in cui vengono aggiunte due o tre basi arbitrarie, le differenti coppie di primer testate sono riportate nella Tabella 6.2

Se le amplificazioni dovevano essere visualizzate tramite il sequenziatore capillare la reazione selettiva detta anche fluorescente (Fluor) è stata ottenuta seguendo il ciclo descritto da Vos, ma con una miscela di reazione diversa mantenendo un volume finale di 20 µl: 5µl PreAmp (diluite 1:15), 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM MgCl₂, 0.25 ng/µl WellRed Eco Primer, 1.5 ng/µl Mse Primer, 0.4 U *Taq* DNA Polymerase Recombinant (Invitrogen).

Tabella 6.2 Combinazioni AFLP testate sui parentali della popolazione Y x O'H

Combinazioni Primer AFLP Ceq8000	Combinazioni Primer AFLP Gel Poliacrilamide
Eco CAC/Mse AAC	Eco CA/Mse AAG
Eco CAC/Mse ACA	Eco CA/Mse ACA
Eco CAC/Mse ACT	Eco CA/Mse ACC
Eco CAC/Mse ACC	Eco CA/Mse AGA
Eco CAC/Mse AGG	Eco CA/Mse AGG
Eco CGA/Mse ACC	Eco CA/Mse ATA
Eco CGA/Mse ACA	Eco CA/Mse AAC
Eco CGA/Mse ACT	Eco CG/Mse AA
Eco CGA/Mse AAC	Eco CG/Mse AT
Eco CGA/Mse AGG	Eco CG/Mse AG
Eco CA/Mse AA	Eco CG/Mse AC
Eco CA/Mse AAA	Eco CA/Mse ATG
Eco CA/Mse AAT	Eco CA/Mse AGC
Eco CA/Mse AAC	Eco CA/Mse ACT
Eco CA/Mse AAG	Eco CA/Mse AAA

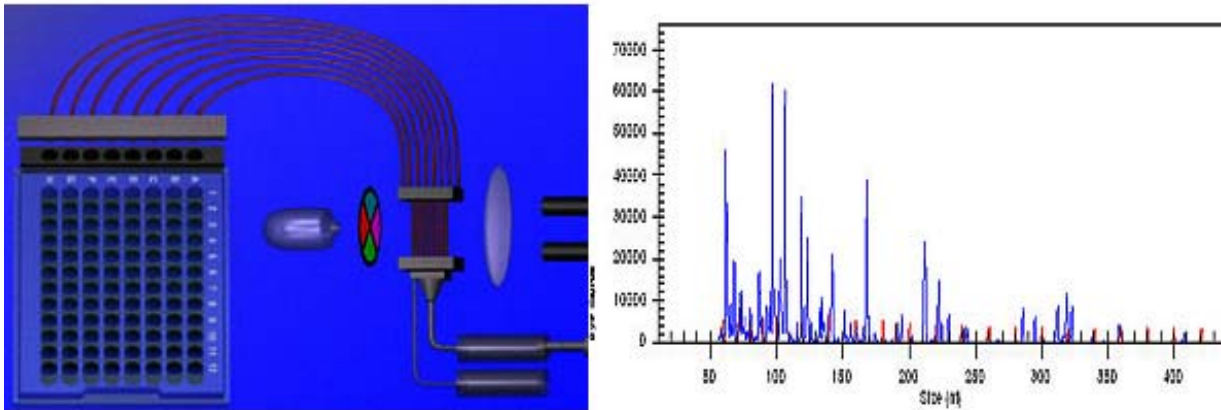
Dopo la reazione di PCR i campioni sono stati caricati in piastre da 96 pozzetti (Beckman Coulter), sul sequenziatore capillare. In ogni pozzetto è stata dispensata una miscela di corsa così composta: 1 µl di Fluor, 0.5 µl di CEQ DNA Size Standard Kit – 400 (Beckman Coulter), 8µl Sample Loading Solution (Beckman Coulter).

La piastra viene posizionata all'interno del sequenziatore che sottopone i campioni a denaturazione, iniezione all'interno del capillare e corsa (voltage 4.2 kV, durata 85 minuti). Contemporaneamente viene posizionata all'interno del sequenziatore una piastra contenente il tampone di corsa CEQ Separation Buffer (Beckman Coulter). All'estremità opposta del capillare rispetto a quella del caricamento, sono posizionati dei laser che sollecitano debitamente i fotocromi; i segnali emessi sono inviati al computer che li visualizza come picchi blu per i frammenti AFLP e rossi per i frammenti dello standard interno (Figura 6.1).

Per quello che concerne la risoluzione delle reazioni selettive su gel di poliacrilamide le combinazioni di primer sono state saggiate su Yumyeong, O'Henry, F₁ H51 e 4 individui arbitrariamente scelti al fine di individuare quelle polimorfiche e segreganti che sono state poi testate su tutta la popolazione.

In questo caso è stato necessario marcare il primer Eco con fosforo 33 (P^{33}). La marcatura dei primer, le reazioni selettive e la corsa elettroforetica sono avvenute utilizzando quanto riportato da Lauria et al. nel 2004.

Figura 6.1 Schema che illustra il funzionamento del sequenziatore capillare ed esempio di elettroferogramma



6.1.5 Analisi dei Dati

I dati ottenuti sono stati analizzati con il programma JoinMap ® 3.0 (Van Ooijen and Voorrips, 2001). A causa dell'elevato numero di marcatori molecolari con fase ignota (AFLP), la popolazione F_2 è stata considerata come derivante da un incrocio tra due individui eterozigoti, utilizzando l'opzione CP. I gruppi d'associazione sono stati determinati ad un valore di $LOD \geq 2$ ed una Frequenza di Ricombinazione ≤ 0.4 . Per la procedura di mappatura è stati presi in considerazione i loci con valori di Rec minori di 0.4 e di LOD maggiore di 0,5. I marcatori che aumentavano i valori di chi-quadro della mappa di 3 (Jump threshold ≥ 3) venivano inizialmente rimossi e poi aggiunti uno alla volta; questi loci venivano forzati all'interno della mappa ed accettati solo se non comportavano riarrangiamenti importanti nell'ordine dei loci.

Tutti i marcatori analizzati sono stati sottoposti al test del chi-quadro per verificare se le segregazioni osservate erano in accordo con quelle mendeliane attese (1:2:1; 3:1).

Per convertire le frequenze di ricombinazione in distanze di mappa (centiMorgan) è stata utilizzata la funzione di Kosambi (1944).

6.2 BULKED SEGREGANT ANALYSIS (BSA)

6.2.1 Materiali Vegetali e Costituzione dei Bulk

Al fine di individuare il gene responsabile del carattere “stony hard” ed un marcatore ad esso correlato, è stato scelto l’approccio della bulked segregant analysis (BSA).

Presso la sezione di Forlì dell’Istituto Sperimentale per la Frutticoltura è presente una popolazione F₂ ottenuta dall’autofecondazione di un semenzale F₁ derivato dall’incrocio O’Henry x Yumyeong (Yumyenog portatore del carattere ‘stony hard’). Gli individui di questa popolazione erano stati valutati per il carattere ‘stony hard’, essendo entrati nella fase produttiva già durante il primo anno della ricerca. La popolazione essendo già stata sottoposta ad una forte selezione non era però adatta a studi di segregazione. All’interno di questa popolazione sono state individuate due classi fenotipiche quella degli “stony hard” e quella dei “non stony hard”. Quindi sono stati costituiti 4 bulk ognuno composto da 8 individui:

- ✓ 2 contengono individui classificati come “stony hard” (SH1, SH2).
- ✓ 2 contengono individui classificati come “non stony hard” (NSH1, NSH2).

Il DNA è stato estratto singolarmente da ciascun individuo componente i bulk, da Yumyeong e da O’Henry utilizzando il metodo Doyle and Doyle (1990) con alcune modifiche descritte in Dettori et al. (2001). I campioni sono stati quantificati utilizzando il fluorimetro DyNA quantum™ 200 (Hofer), quindi per ciascun bulk sono stati miscelati i DNA in ugual misura per ottenere una concentrazione finale di 10 ng/μl (pari ad 1.25 per ogni campione).

6.2.2 Random Amplified Polymorphic DNA (RAPDs)

I 4 bulk SH1, SH2, NSH1, NSH2, ed i due parentali (Yumyeong, O’Henry) sono stati analizzati via PCR utilizzando 296 primer della Operon Technologies (Alameda, Calif.) tabella 5.3. Le amplificazioni sono state eseguite in un volume finale di 10 μl, utilizzando 10 ng di DNA, 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM MgCl₂, 0.1 mM dNTPs, 0.25 μM Primer, 1 U *Taq* DNA Polymerase Recombinant (Invitrogen). Le Amplificazioni sono state ottenute utilizzando il termociclatore GeneAmp PCR System 9600 (Perkin-Elmer) con il seguente protocollo: 45 cicli, 94°C per 1 minuto, 36 °C per 1 minuto, 72 °C per 2 minuti ed un’estensione finale a 72 °C per 8 minuti.

I prodotti di PCR ottenuti sono stati separati con elettroforesi orizzontale su gel d’agarosio al 1.5% in una soluzione di TAE 1X applicando un voltaggio di 3 V/cm.

I RAPD che evidenziavano il profilo ricercato venivano testati nuovamente nei bulk.

6.2.3 Amplified Fragment Length Polymorphisms (AFLPs)

I 4 bulk SH1, SH2, NSH1, NSH2, ed i due parentali sono stati analizzati via PCR utilizzando i marcatori molecolari AFLP e visualizzando i frammenti con il Sequenziatore capillare CEQ™ 8000 Genetic Analysis System (Beckman Coulter). Per la metodica si fa riferimento al paragrafo dedicato agli AFLP nel capitolo dedicato alla costruzione della Mappa d'Associazione. Le 8 combinazioni testate sono: Eco ACA/Mse CAT, Eco ACA/Mse CAA, Eco AC/Mse CAT, Eco AG/Mse CA, Eco AG/Mse CAA, Eco AG/Mse CAC, Eco AG/Mse CAT, Eco AC/Mse CAT.

Tabella 6.3 Sigla primer RAPD saggiati nella BSA

	OPA2	OPA3	OPA4	OPA5	OPA6	OPA7			
OPG1	OPG2	OPG3	OPG4	OPG5	OPG6	OPG7	OPG8	OPG9	OPG10
OPG11	OPG12	OPG13	OPG14	OPG15	OPG16	OPG17	OPG18	OPG19	OPG20
OPH1	OPH2	OPH3	OPH4	OPH5	OPH6	OPH7	OPH8	OPH9	OPH10
OPH11	OPH12	OPH13	OPH14	OPH15	OPH16	OPH17	OPH18	OPH19	OPH20
OPI1	OPI2	OPI3	OPI4	OPI5	OPI6	OPI7	OPI8	OPI9	OPI10
OPI11	OPI12	OPI13	OPI14	OPI15			OPI18		
POJ1									
		OPJ13		OPJ15		OPJ17	OPJ18		
OPK1	OPK2	OPK3	OPK4	OPK5	OPK6	OPK7	OPK8	OPK9	OPK10
OPK11	OPK12	OPK13	OPK14	OPK15	OPK16	OPK17	OPK18	OPK19	OPK20
OPL1	OPL2	OPL3	OPL4	OPL5	OPL6	OPL7	OPL8	OPL9	OPL10
OPL11	OPL12	OPL13	OPL14	OPL15	OPL16	OPL17	OPL18	OPL19	OPL20
OPM1	OPM2	OPM3	OPM4	OPM5	OPM6	OPM7	OPM8	OPM9	OPM10
OPM11	OPM12	OPM13	OPM14	OPM15	OPM16	OPM17	OPM18	OPM19	OPM20
OPN1	OPN2	OPN3	OPN4	OPN5	OPN6	OPN7	OPN8	OPN9	OPN10
OPN11	OPN12	OPN13	OPN14	OPN15	OPN16	OPN17	OPN18	OPN19	OPN20
OPO1	OPO2	OPO3	OPO4	OPO5	OPO6	OPO7	OPO8	OPO9	OPO10
OPO11	OPO12	OPO13	OPO14	OPO15	OPO16	OPO17	OPO18	OPO19	OPO20
OPP1	OPP2	OPP3	OPP4	OPP5	OPP6	OPP7	OPP8	OPP9	OPP10
OPP11	OPP12	OPP13	OPP14	OPP15	OPP16	OPP17	OPP18	OPP19	OPP20
OPQ1	OPQ2	OPQ3	OPQ4	OPQ5	OPQ6				
OPR1	OPR2	OPR3	OPR4	OPR5	OPR6	OPR7	OPR8	OPR9	OPR10
OPR11	OPR12	OPR13	OPR14	OPR15	OPR16	OPR17	OPR18	OPR19	OPR20
OPS1	OPS2	OPS3	OPS4	OPS5	OPS6	OPS7	OPS8	OPS9	OPS10
OPS11	OPS12	OPS13	OPS14	OPS15	OPS16	OPS17	OPS18	OPS19	OPS20
OPT1	OPT2	OPT3	OPT4	OPT5	OPT6	OPT7	OPT8	OPT9	OPT10
OPT11	OPT12	OPT13	OPT14	OPT15	OPT16	OPT17	OPT18	OPT19	OPT20
OPU1	OPU2	OPU3	OPU4	OPU5	OPU6	OPU7	OPU8	OPU9	OPU10
OPU11	OPU12	OPU13	OPU14	OPU15	OPU16	OPU17	OPU18	OPU19	OPU20
OPAD1	OPAD2	OPAD3	OPAD4	OPAD5	OPAD6	OPAD7	OPAD8	OPAD9	OPAD10
OPAD11	OPAD12	OPAD13	OPAD14	OPAD15	OPAD16	OPAD17	OPAD18	OPAD19	OPAD20

Il polimorfismo ricercato sia nel caso dei marcatori RAPD che in quello degli AFLP era del tipo presenza della banda nei bulk 'stony hard' ed in Yumyeong ed assenza nei bulk 'non stony hard' ed in O'Henry; quando veniva riscontrato questo tipo di polimorfismo la PCR veniva ripetuta per confermare il risultato. Quindi si ripeteva l'analisi sui singoli individui che compongono i Bulk per

verificare se il polimorfismo era effettivamente presente in ogni individuo 'stony hard' ed assente nei 'non stony hard'.

6.3 EXPRESSED SEQUENCE TAGS (ESTs) DA FRUTTO DELLA CV YUMYEONG

Per individuare i geni espressi all'interno di un determinato tessuto o organo è necessario analizzare l'insieme dei trascritti in quel tessuto o organo in un preciso momento, per questo scopo si è deciso di produrre delle Expressed Sequence Tags (ESTs).

6.3.1 Estrazione RNA

Per costruire la libreria cDNA è stato estratto RNA dal mesocarpo di un frutto della cv. Yumyeong allo stadio S4 (maturazione fisiologica). L'estrazione dell'RNA totale è avvenuta a partire da 36 grammi di polpa, conservati a -80°C , utilizzando il procedimento descritto da Manning (1990) con alcune modificazioni minori. La particolarità di questo procedimento sta nell'uso del 2-Butossietanolo un solvente organico che a differenti concentrazioni riesce a promuovere una precipitazione differenziale. Aggiunto in un rapporto 2:5 porta alla precipitazione dei polisaccaridi (presenti in grande quantità nel frutto e che possono inibire l'estrazione degli acidi nucleici) mentre in un rapporto 1:1 porta alla precipitazione degli acidi nucleici.

Per quantificare l'RNA totale e per valutarne la qualità in termini di purezza è stato utilizzato uno spettrofotometro a due lunghezze d'onda 260nm e 280nm. L'RNA viene utilizzato se il rapporto fra la lettura a 260 e la lettura 280 è compreso fra 1.9 e 2.1.

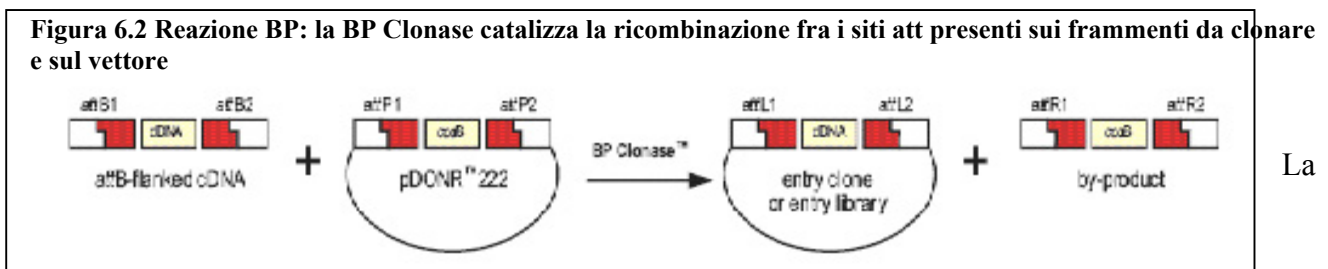
Per verificare l'integrità dell'RNA 1 μg è stato sottoposto ad una corsa elettroforetica su gel d'agarosio in tampone TAE 1X (5 V/cm). L'RNA è stato denaturato con uno shock termico e miscelato con un tampone a base di formaldeide che mantiene la denaturazione durante la corsa. L'RNA è stato visualizzato tramite colorazione con BrEt, e l'immagine è stata ottenuta utilizzando l'acquisitore d'immagine digitale (GEL Logic 100 Imaging system, Kodak).

Per costruire una libreria cDNA è necessario utilizzare solo il messaggero (poly A⁺ RNA), quindi l'RNA totale va purificato. Per la purificazione è stato utilizzato il Kit NucleoTrap® mRNA (Clontech), che posiziona su delle biglie di silicone delle code di poly T le quali legano le code poly A, caratteristiche degli RNA messaggeri. Per verificare l'avvenuta purificazione l'RNA viene nuovamente sottoposto ad una corsa elettroforetica su gel d'agarosio

6.3.2 Costruzione della cDNA Library

La libreria cDNA è stata ottenuta utilizzando il CloneMiner cDNA Library construction Kit (Invitrogen); questo kit è stato scelto perché presenta numerosi vantaggi: è rapido, non necessita di una grande quantità di RNA iniziale, utilizza come ospiti per la trasformazione batteri e non fagi come, invece, la maggior parte dei kit presenti sul mercato, si avvale della tecnologia Gateway® e non è obbligatorio l'uso di isotopi marcati radioattivamente.

La tecnologia Gateway® è un sistema universale di clonaggio che si basa sulla ricombinazione sito specifica propria del batteriofago Lambda (Landy, 1989), questo tipo di ricombinazione consente al batteriofago di inserirsi nel cromosoma di *E.coli* e di passare così dal ciclo litico al ciclo lisogenico (Ptashne, 1992). Nella tecnologia Gateway®, i componenti della ricombinazione di Lambda sono stati migliorati al fine di implementare la specificità e l'efficienza del sistema (Bushman et al., 1985). La ricombinazione di Lambda avviene attraverso siti specifici (*att*): *attB* sul cromosoma di *E.coli* e *attP* sul cromosoma di Lambda. Questi siti fungono da siti di riconoscimento per le proteine di ricombinazione (Weisemberg and Landy, 1983); durante l'integrazione di Lambda avvengono fra questi siti di riconoscimento delle ricombinazioni che alla formazione delle regioni *attL* e *attR*(Figura 6.2). L'enzima BP Clonasi facilita la ricombinazione fra le sequenze fiancheggianti. All'interno del vettore di clonaggio pDONR™222 è stato inserito il gene *ccdB* che consente di effettuare una selezione negativa dopo le reazioni di ricombinazione e trasformazione. Questo gene interferisce con la crescita della cellula portando alla morte quelle colonie di *E.coli* che posseggano un vettore non ricombinante evitando quindi lo screening blue/bianco delle colonie.



costruzione della libreria ha comportato diverse fasi:

1. Ligazione del primo adattatore *attB2*-Biotina ai filamenti di RNA (Figura 6.3)
2. Retrotrascrizione dell'RNA: conversione dell'RNA a singolo filamento in DNA a singolo filamento (Figura 6.3).
3. Conversione da cDNA a singolo filamento a doppio filamento (Figura 6.3)
4. Ligazione del secondo adattatore *attB1* all'estremità 5' dei frammenti (Figura 6.3)

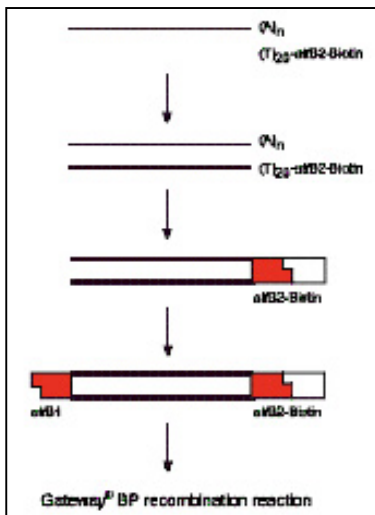


Figura 6.3 Retrotrascrizione dell'RNA

5. I frammenti ottenuti vengono inseriti nel vettore pDONRTM222 con la tecnologia Gateway®
6. Trasformazione del ceppo cellulare di *E.coli* ElectroMAXTM DH10BTM T1 Phage Resistent Cell

Per la costruzione della libreria sono stati utilizzati 2 µg di mRNA ed è stato seguito in modo assolutamente fedele il protocollo fornito dal kit. Non essendo il nostro laboratorio attrezzato per l'uso di materiale radioattivo per il frazionamento del cDNA prima della legazione nel vettore di clonaggio è stata utilizzata la metodica non radioattiva riportata nell'appendice web del kit stesso.

Durante tutti i passaggi necessari per la costruzione sono stati inseriti diversi controlli, come suggerito dal manuale.

La libreria ottenuta è stata sottoposta a diverse analisi per verificarne il titolo (colonie formanti unità) e determinare la dimensione media dei frammenti inseriti. A questo scopo sono state piastrate, su piastre contenenti terreno LB (10 g NaCl, 10 g Triptone, 5 g Estratto di lievito, 20 g Agar) e kanamicina (50 µg/ml), diverse diluizioni della libreria. La titolazione della libreria è stata ottenuta applicando la formula:

$$\text{cfu/ml} = \frac{\text{colonie in piastra} \times \text{fattore di diluizione}}{\text{volume piastrate}}$$

Il numero totale delle colonie formanti unità è stato dedotto dalla formula:

$$\text{Total CFU} = \text{cfu/ml volume} \times \text{volume totale della libreria cDNA}$$

Da una delle piastre ottenute sono state prelevate arbitrariamente 24 colonie che sono state messe in terreno di coltura in LB (10 g NaCl, 10 g Triptone, 5 g Estratto di lievito) e Kanamicina (50 µg/ml) a 37°C per tutta la notte (o/n), quindi sono stati estratti i plasmidi con il metodo della lisi alcalina secondo Maniatis.

I plasmidi sono stati visualizzati tramite colorazione con SYBR Green I Nucleic Acid Gel Stain (Invitrogen) dopo una corsa elettroforetica (3 v/cm) su gel d'agarosio (1%) in tampone TAE 1X.

Per stimare la dimensione degli inserti contenuti all'interno dei plasmidi, questi sono stati sottoposti a digestione enzimatica con l'enzima *Bsr*GI. Il vettore pDONRTM222 possiede per questo enzima 3 siti restrizione al suo interno, se sottoposto a digestione enzimatica il suo pattern risulta essere composto da 3 bande di 2.5 1.4 e 0.79 Kb. I plasmidi che contengono i frammenti di cDNA target presentano solo 2 siti di restrizione per *Bsr*GI e se digeriti, il profilo che si ottiene presenta una

banda a 2.5 Kb corrispondente al vettore e delle bande addizionali relative al frammento di cDNA inserito.

Si considera la libreria di buona qualità quando il numero dei cloni che la formano è nell'ordine di 10⁶-10⁷, quando il numero dei ricombinanti è maggiore del 95% e il valore medio dell'inserito è maggiore o uguale a 1000bp.

6.3.3 Sequenziamento Cloni cDNA Library

Questa parte del lavoro è stata svolta presso il Genome Research Laboratory (GRL), della North Carolina State University. Sono stati sequenziati 1907 cloni della libreria di cDNA prodotta.

Circa 10.000 cloni della Libreria sono stati piastrati sulle Qplate (piastre 25cmX25cm) contenenti terreno LB (10 g NaCl, 10 g Triptone, 5 g Estratto di lievito, 20 g Agar) e kanamicina (50 µg/ml), 3000 di questi sono stati isolati, utilizzando il Qbot 110V (Genetix) e trasferiti in piastre da 384 pozzetti contenenti LB (10 g NaCl, 10 g Triptone, 5 g Estratto di lievito) e kanamicina (50 µg/ml) e messi in coltura a 37°C tutta la notte. Dalla coltura liquida è stata prelevata una aliquota e questa è stata utilizzata come template per la reazione di PCR al fine di amplificare i frammenti contenuti all'interno dei plasmidi.

Le amplificazioni sono state eseguite in un volume finale di 20 µl, utilizzando 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM MgCl₂, 0.3 mM dNTPs, 0.25 µM Primer M13 forward, 0.25 µM Primer M13 reverse, 1 U *Taq* DNA Polymerase Recombinant (Invitrogen). Le Amplificazioni sono state fatte utilizzando il termociclatore GeneAmp[®] PCR System 9700 (Perkin-Elmer) con il seguente protocollo: una denaturazione iniziale a 94°C per 3 minuti quindi 25 cicli, 94°C per 1 minuto, 47 °C per 30 secondi, 72 °C per 2 minuti ed un'estensione finale a 72 °C per 8 minuti.

I prodotti di PCR ottenuti sono stati separati con elettroforesi orizzontale su gel d'agarosio al 1.5% in una soluzione di TBE 1X applicando un voltaggio di 3 V/cm, quindi i frammenti sono stati visualizzati tramite colorazione con BrEt. L'immagine è stata acquisita utilizzando una fotocamera digitale (GEL Logic 100 Imaging system, Kodak).

Dopo aver verificato l'amplificazione degli inserti, i prodotti dell'amplificazione sono stati purificati aggiungendo un volume di 3M NaOAc pH 5.2 in un rapporto 1/10 e Isopropanolo in un rapporto 3/4, quindi i campioni sono stati riposti a -80°C o/n, poi centrifugati per 30 minuti a 3000 rpm, è stato eliminato il sovrantante ed il pellet è stato lavato con EtOH 70% (freddo). I campioni sono stati centrifugati per 30 minuti a 3000 rpm, è stato eliminato il sovrantante il pellet è stato seccato e risospeso in 8µl di ddH₂O sterile.

I frammenti purificati sono stati quantificati con elettroforesi orizzontale su gel d'agarosio al 1.5% in una soluzione di TBE 1X applicando un voltaggio di 3 V/cm, quindi i frammenti sono stati visualizzati tramite colorazione con *BrEt*. L'immagine è stata acquisita utilizzando una fotocamera digitale (GEL Logic 100 Imaging system, Kodak).

Fra i 5 ed i 20 ng di DNA sono stati utilizzati per la reazione di sequenza che è stata effettuata utilizzando il sequenziatore capillare ABI®Prism 3700 DNA Analyser (Applied Biosystems) ed il kit di sequenza BigDye®Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems).

Il kit utilizzato è un'evoluzione del metodo dei dideossi ideato da Sanger. Ad ognuno dei quattro dideossinucleotidi sono associati quattro diversi fluorocromi che emettono, se opportunamente eccitati, radiazioni luminose di colore diverso. Questi Fluorocromi sono costituiti da due componenti: un donatore 6-carboxilfluoresceina (6-FAM) ed un accettore differente per ogni singolo nucleotide derivante dalla diclorodamina (dRhodamine). Oltre ai nucleotidi Terminatori il kit contiene: i dideossinucleotidi, il tampone di reazione, il MgCl₂ e la *taq* polimerasi. La miscela dei desossinucleotidi ne contiene due modificati: il dITP al posto del dGTP, per minimizzare la compressione delle bande ed il UTP al posto del dTTP, per migliorare il segnale dato dalle bande T. La *taq* polimerasi utilizzata dal kit è la AmpliTaq® DNA Polymerase (Applied Biosystems). Questa *Taq* presenta una doppia particolarità: possiede una mutazione puntiforme nel sito attivo, ciò produce una minore discriminazione nei confronti dei dideossinucleotidi ed ha anche una seconda mutazione nel dominio amino-terminale che elimina l'attività esonucleasica 5'→3'.

Le reazioni di sequenza sono state eseguite in un volume finale di 10 µl, utilizzando 2µl templatato , 2µl BigDye, 1µl (3.2 pmol) Primer M13 forward. Le Amplificazioni sono state fatte utilizzando il termociclatore GeneAmp® PCR System 9700 (Perkin-Elmer) con il seguente protocollo: una denaturazione iniziale a 96°C per 1 minuto quindi 25 cicli, 96°C per 10 secondi, 50 °C per 5 secondi, 60 °C per 4 minuti.

Le sequenze sono state purificate per eliminare tutti i terminatori non incorporati utilizzando un kit
Le sequenze purificate sono state caricate sul sequenziatore per la separazione che è avvenuta utilizzando il programma standard per le reazioni di sequenza.

6.3.4 Annotazione delle sequenze

Le sequenze sono state annotate presso il Clemson University Genomic Institut utilizzando un software pubblico ed una pipeline da loro sviluppata (ProcEST.pl). Il procedimento consta di tre fasi principali:

Fase I: Processamento dei Cromatogrammi

I cromatogrammi sono stati convertiti in formato FASTA, l'assegnazione delle basi e gli indici di qualità corrispondenti sono stati effettuati utilizzando il programma Phred (Ewing et al., 1998). Le contaminazioni del vettore e delle sequenze dell'ospite sono state mascherate utilizzando il programma di comparazione delle sequenze cross_match (Gordon et al., 1998). Dopo che le sequenze sono state decontaminate dal vettore da ognuna sono state eliminate le basi di bassa qualità che si trovano all'estremità dello script (basi a cui Phred ha assegnato un valore inferiore a 20). Le sequenze sono state eliminate se possedevano più del 5% di basi ambigue, dei tratti composti da ripetizioni di A o T maggiori di 40, o meno di 100 basi di alta qualità (valore di Phred pari a 20). A questo stadio il processo ha fornito diversi file: overall summary report file (statistica generata dalla pipeline basata sui cromatogrammi), clone report file (statistica generata dalla pipeline che considera il successo delle singole sequenze), GeneBank submission file (sottomissione automatica all'NCBI utilizzando le informazioni sulla libreria fornite inizialmente) ed un file con tutte le sequenze in formato FASTA con, per ognuna, i valori di Phred. A questo livello dalle sequenze vengono anche eliminate tutte quelle omologhe con le sequenze ribosomiali, cloroplastiche e mitocondriali (se note).

Fase II: Costituzione dei Contig utilizzando le sequenze di alta qualità

In questa fase le sequenze di elevata qualità sono state assemblate per la costituzione dei contig utilizzando il programma CAP3 (Huan and Madan, 1999). Per prevenire errati assemblaggi sono stati utilizzati dei parametri molto stringenti (p 90, d 60), questo tipo di stringenza può anche rivelarsi utile per l'individuazione di geni paraloghi.

Fase III: Annotazione

L'annotazione del gruppo degli unigene consiste nel confrontare il set dei contig e delle singole sequenze della libreria con tutte le proteine presenti nel database nr di GeneBank utilizzando l'algoritmo fastx3.4 (Pearson and Lipman, 1988). Le sequenze sono state anche caratterizzate tramite il confronto con il Genbank Rosaceae EST dataset (<http://www.ncbi.nlm.nih.gov>) e con il Genome Database for Rosaceae (GDR) relativo alle EST mappate sulla mappa fisica di pesco in costruzione presso la Clemson University utilizzando il pacchetto dei programmi BLAST (Altschul et al., 1997).

Gli appaiamenti più significativi per ogni contig e per ogni singolo clone della libreria vengono registrati. Le sequenze Microsatelliti (SSR) sono state individuate all'interno del set degli unigene utilizzando il programma CUGISSR.pl.

6.3.5 Caratterizzazione Marcatori Molecolari Microsatelliti derivati dall' EST

All'interno delle EST sono state individuate 698 sequenze microsatelliti. Di queste 32 sono state utilizzate per disegnare dei primer corrispondenti alle regioni fiancheggianti il locus microsatelliti. A tal fine è stato utilizzato il programma Primer3 (Rozen & Skaletsky 1998), pubblico e disponibile in rete (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Questo software permette di disegnare dei primer tenendo conto di diversi parametri che influiscono sulla funzionalità della coppia:

1. Lunghezza dei primer. Vengono richiesti 3 valori, espressi in basi, che corrispondono alla lunghezza minima, massima e a quella ottimale che i primer devono avere;
2. Temperature di Melting. Questo è uno dei parametri fondamentali, la T_m è quel valore per cui un frammento a doppio filamento rompe i suoi legami covalenti e diventa a singolo filamento. Anche in questo caso, sono richiesti un valore minimo, uno ottimale ed uno massimo;
3. Differenza di Temperatura di Melting tra i due primer. Questo parametro è importante per evitare le associazioni semispecifiche dei primer a sequenze non del tutto omologhe;
4. Percentuale in GC per ognuno dei due primer. Questo valore influenza la temperatura di melting dei primer e la stringenza (temperatura di annealing alta) della reazione di PCR, anche in questo caso bisogna trovare un giusto compromesso, di solito un buon valore è intorno al 40-50%;
5. Grado di Autocomplementarietà di ogni singolo primer. Questo parametro serve per evitare che il primer formi dei loop per la presenza di estremità palindromiche;
6. Grado di Complementarietà fra i due primer. Minore è questo valore minore sarà la possibilità che l'estremità 3' dei due primer si appaino tra loro.

Il programma propone una coppia di primer che rappresenta il miglior compromesso fra le caratteristiche proprie della sequenza ed i parametri sopracitati tenendo conto dei principali fattori chimico fisici e termodinamici che entrano in gioco durante una reazione di amplificazione.

Le 32 coppie di primer sono state testate sui parentali delle 3 popolazioni segreganti disponibili presso il nostro laboratorio una progenie F_2 derivante dall'autofecondazione di una F_1 di un ibrido pesco-mandorlo, una popolazione Back-Cross derivante dall'incrocio di un'accessione di *P.ferganensis*, maschio donatore, con un'accessione di Pesco IF7310828, come femmina ricorrente; e con 8 genotipi appartenenti al genere *Prunus*: 1mandorlo, 2 albicocchi, 3 susini e 2 ciliegi (Tabella 6.4). Le amplificazioni sono state eseguite in un volume finale di 10 μ l utilizzando, per ogni reazione, 20 ng di DNA genomico, 20 mM Tris-HCl (pH 8,4), 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.2 μ M per primer e 0.5 U di Platinum[®] GenoType *Tsp* DNA Polymerase (Invitrogen).

Le amplificazioni sono state eseguite utilizzando un termociclatore GENEAMP® PCR System 9600 (Perkin-Elmer) seguendo i cicli di reazione e le temperature di annealing pubblicati in Testolin et al. (1999). I prodotti di PCR ottenuti sono stati separati con elettroforesi orizzontale su gel d'agarosio (MetaPhor® (Cambrex)) al 3% in una soluzione di TBE 1X applicando un voltaggio di 5 V/cm, quindi i frammenti sono stati visualizzati tramite colorazione con *BrEt*. I gel sono stati fotografati con un acquirente d'immagini digitale (GEL Logic 100 Imaging system, Kodak). All'interno dei Peschi è stata calcolata l'eterozigosità osservata e attesa.

Tabella 6.4 Lista varietà utilizzate per testare i primer microsatelliti EPPISF

Yumyeong	pesco	Ferragnes	mandorlo
O'Henry	pesco	Ninfa	albicocco
F1 (O'HxY)	pesco	Laycot	albicocco
Prunus Ferganensis	pesco	Firenze 90	susino europeo
IF7310828	pesco	Angeleno	susino giapponese
F1 (PxY)	pesco	Green Sun	susino giapponese
Early Gold	pesco	Burlat C1	ciliegio dolce
Texas	mandorlo	Santa Rosa	ciliegio dolce
F1 (TxY)	ibrido pesco-mand		

6.4 ANALISI DELL'ESPRESSIONE DELLE ENDOPOLIGALACTURONASI (*ENDOPG*) E DELL' 1-AMINOCICLOPROPANO-1-CARBOSSILATO OSSIDASI (*ACO*)

6.4.1 Estrazione RNA Mesocarpo frutto Stadio S4

Al fine di analizzare l'espressione di due degli enzimi coinvolti nel processo di ammorbidimento della polpa è stato estratto l'RNA totale dal mesocarpo maturo allo stadio S4 della cultivar Yumyeong, Françoise e dai singoli individui che compongono la popolazione F₂ O'H x Y, utilizzata per la costruzione della mappa d'associazione.

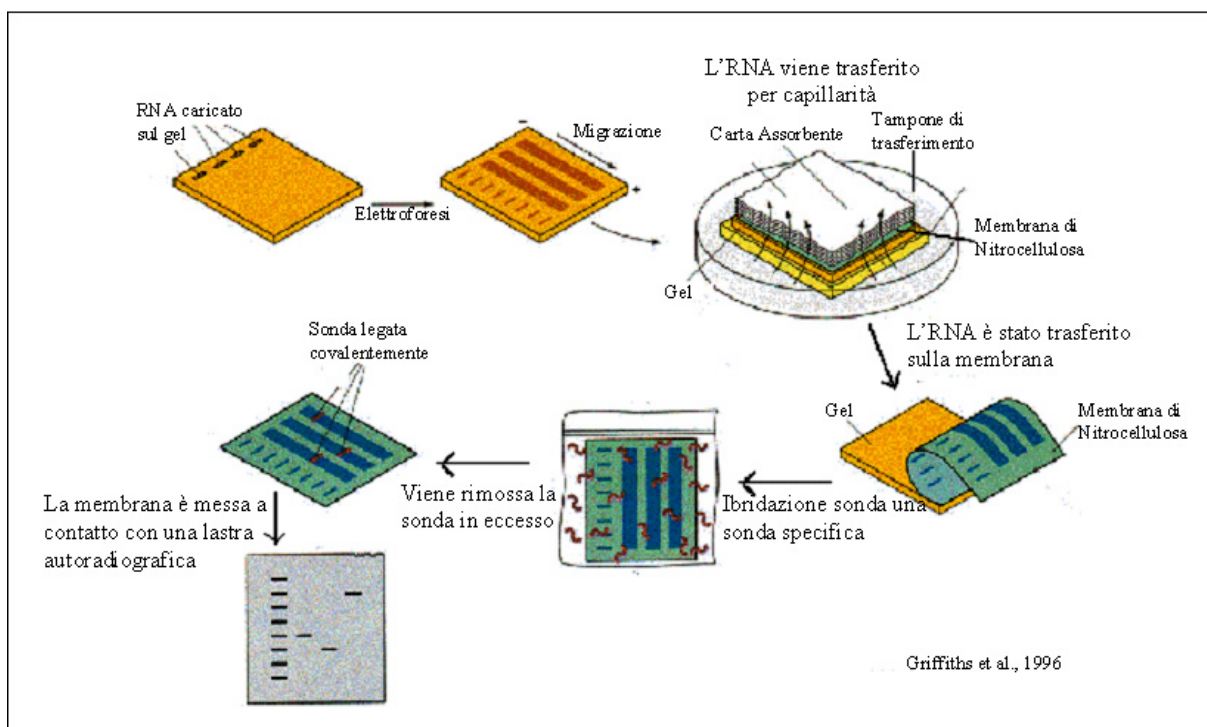
L'estrazione dell'RNA totale è avvenuta a partire da 6 grammi di polpa, che era stata conservata a -80 °C, utilizzando il procedimento descritto da Manning (1990) con alcune modificazioni minori. Per quantificare l'RNA e per valutarne la qualità in termini di purezza è stato utilizzato uno spettrofotometro a due lunghezze d'onda 260nm e 280nm. L'RNA viene utilizzato se il rapporto fra la lettura a 260 e la lettura 280 è compreso fra 1.9 e 2.1.

Per verificare l'integrità dell'RNA, 1 µg viene sottoposto ad una corsa elettroforetica su gel d'agarosio in tampone TAE 1X (5 V/cm).

6.4.2 Analisi Northern

Cinque microgrammi dell'RNA totale delle pesche analizzate è stato utilizzato per effettuare una corsa elettroforetica seguendo gli stessi parametri citati in precedenza. Il gel, un ponte di carta 3MM spessa, 4 fogli di carta 3MM sottili (della dimensione del gel) sono stati immersi per 5 minuti nel Transfer Buffer (3M NaCl, 8mM NaOH, 2mM sarkosyl). Anche la membrana (Nylon Membranes, positively charged; Roche) è stata immersa nel buffer per 5 minuti. Quindi è stato montato l'apparato ed il trasferimento è avvenuto o/n (Figura 6.4). A questo punto l'RNA è stato fissato covalentemente alla membrana, tramite trattamento termico (60°C per 15 minuti) ed ai raggi UV. La membrana ottenuta è stata utilizzata per l'ibridazione con sonde specifiche.

Figura 6.4 Northern Blot ed Ibridazione della Membrana



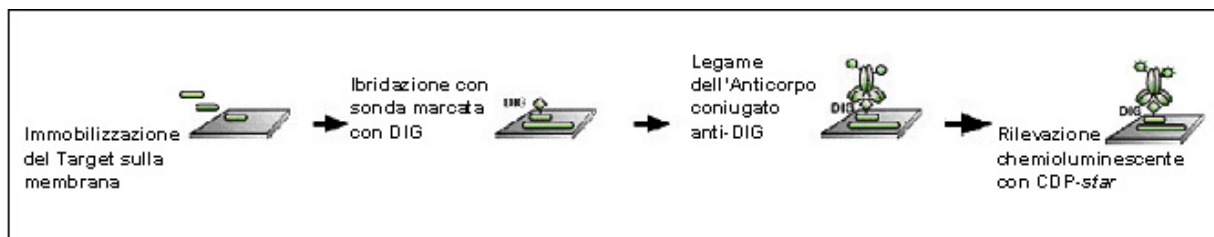
6.4.3 Ibridazione

Le due sonde scelte per l'analisi corrispondono alla porzioni 3'UTR dei due geni presi in esame: endopoligalatturonasi (*endoPG*), 1-aminociclopropano-1-carbossilato ossidasi (ACO) (Tonutti et al., 2000).

I due frammenti sono stati marcati utilizzando una metodica non radioattiva basata sull'uso della digossigenina DIG-dUTP. La digossigenina è un dinucleotide modificato omologo alla digossitimina che via PCR viene inglobato all'interno della sonda; la DIG è stericamente più ingombrante rispetto al dTTP e quindi bisogna trovare il giusto rapporto fra questi due nucleotidi per ottenere un risultato ottimale. La verifica della marcatura avviene tramite elettroforesi mettendo

a confronto il frammento marcato con quello nativo, se la DIG è stata inglobata efficientemente il frammento marcato è più “pesante” e quindi corre più lentamente. La rilevazione del segnale dell’ibridazione avviene tramite una reazione immunochimica. La fosfatasi alcalina coniugata ad un anticorpo anti-DIG si lega alla sonda marcata; quindi l’ibrido anticorpo-sonda viene visualizzato da un substrato chemiluminescente attraverso la fosfatasi alcalina (Figura 6.5). Il substrato produce una radiazione luminosa che può impressionare una lastra autoradiografica.

Figura 6.5 Schema procedura d'ibridazione con Digossigenina



La marcatura della sonda, l’ibridazione della membrana e la rilevazione sono state effettuate seguendo il protocollo fornito dalla Roche. La membrana viene inserita in una bottiglia di ibridazione e messa a preibridare all’interno di un fornello rotante per 1 ora a 50°C con 15 ml di DIG Easy Hyb Buffer (Roche); la sonda viene denaturata con uno shock termico e inserita all’interno della bottiglia, l’ibridazione avviene a 50°C ed ha una durata di 12-16 h. Quindi la membrana viene sottoposta a dei lavaggi, per eliminare la sonda in eccesso che non si è legata alla membrana. I lavaggi vengono ripetuti due volte: 5 minuti in LOW STRINGENCY BUFFER (2x SSC, 0,1% SDS) a temperatura ambiente; 15 minuti in HIGH STRINGENCY BUFFER (0,1x SSC, 0,1% SDS) a 50°C. Dopo questa fase inizia la fase di rilevamento che avviene a temperatura ambiente: 5 minuti in WASHING BUFFER (1x Tampone Maleico, 0,3% v/v Tween 20); 30 minuti in BLOCKING SOLUTION (1:10 10 x Blocking Solution: 1 x Tampone Maleico); 30 minuti in Soluzione con ANTICORPO (0.5 µl in 10 ml 1x Blocking Solution); lavare la membrana per 15 minuti con 10 ml di WASHING BUFFER; ripetere l’operazione. Equilibrare la membrana per 5 minuti con 10 ml di DETECTION BUFFER (0.1M TrisHCl; 0.1M NaCl). Diluire 50 µl di CDP-Star in 5 ml DETECTION BUFFER e assicurarsi di coprire perfettamente la membrana.

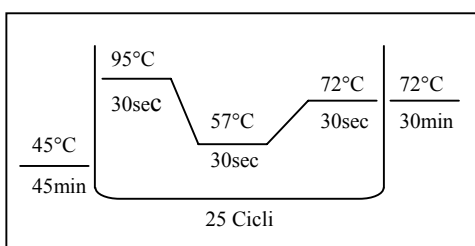
Quindi la membrana viene sigillata con pellicola trasparente e messa a contatto con la lastra autoradiografica. Il tempo d’esposizione della lastra è variabile e dipende dall’intensità del segnale di ogni sonda, in genere si utilizza un tempo d’esposizione di 30 minuti. I blot sono stati conservati a 4°C protetti da fogli di carta assorbente e avvolti in pellicola trasparente.

6.4.4 ReverseTrascriptase-PCR

Al fine di verificare l'espressione dell'ACC ossidasi in ognuno dei singoli individui utilizzati per la costruzione della mappa è stata effettuata una retrotrascrizione del gene. Questa metodica utilizza una retrotrascrittasi virale che trasforma l'RNA in DNA a singolo filamento in unione con una polimerasi standard che è in grado di utilizzare il filamento di DNA come stampo. A tal fine vengono impiegate una o più coppie di primer specifici, in questo caso particolare è stata utilizzata la stessa coppia di primer disegnata sulla sequenza della 3'UTR dell'ACO, già usata per la marcatura della sonda (Ruperti et al.,2001). Nella pratica è stato utilizzato il kit AccessQuick RT-PCR System (Promega) alle condizioni riportate in Tabella 5.6. I frammenti ottenuti sono stati separati tramite elettroforesi orizzontale su gel d'agarosio al 1.5% in una soluzione di TAE 1X applicando un voltaggio di 5 V/cm.

Tabella 6.5 Miscela di reazione e ciclo utilizzati per l'RT-PCR

Reagenti	Quantità
RNA	1.5
AccessQuick MasterMix (2x)	13.25
Aco F (10 μ M)	0.5
Aco R (10 μ M)	0.5
AMV Reverse Trascriptase (5U/ μ l)	0.5
Nulease Free H ₂ O	11



7 RISULTATI

7.1 COSTRUZIONE DELLA MAPPA D'ASSOCIAZIONE O'H x Y

7.1.1 *Marcatori Microsatelliti*

Utilizzando i parentali della popolazione (O'Henry e Yumyeong) e l'ibrido F1 sono state saggiate 357 coppie di primer microsatelliti, 325 reperite in bibliografia e 32 disegnate sulla base delle EST ottenute sequenziando una libreria di cDNA derivante de mesocarpo allo stadio S4 della cultivar Yumyeong. L'84% dei marcatori saggiati ha dato un prodotto di amplificazione; 122 di questi, pari al 40,7%, erano polimorfici nei due parentali (Tabella 7.1), questo valore percentuale è da considerarsi medio, infatti sono evidenti le notevoli differenze fra i tassi di polimorfismo riscontrato nei diversi gruppi di SSR analizzati. Dei 122 primer polimorfici solo 40 (13,3%) erano eterozigoti nell'ibrido F₁ (dal quale discende la popolazione in esame), e quindi si aspettava segregassero nella popolazione F₂ (Tabella 7.1). Trentasei di questi sono stati saggiati su tutta la popolazione per costruire la mappa d'associazione. Quattro marcatori segreganti sono stati scartati a causa di un profilo di difficile lettura.

Tabella 7.1 Sigla, origine, numero e tasso di polimorfismo dei marcatori SSR saggiati sulla popolazione O'H x Y (O'H O'Henry, Y Yumyeong)

SSR sigla	Specie	N° SSR saggiati	N° SSR amplificati	N° SSR polimorfici	% Polimorfismo O'H, Y	N° SSR eterozigoti in F ₁	N° SSR saggiati in F ₂
EPPISF	Pesco	32	29	6	20,7	2	2
BPPCT	Pesco	42	41	15	36,6	7	6
CPDCT	Mandorlo	33	21	4	19,0	1	1
CPPCT	Pesco	35	33	7	21,2	2	2
CPSCT	Susino Giapponese	34	33	8	24,2	4	3
EmpaS	Ciliegio Selvatico	11	11	2	18,2	2	2
MA/M	Pesco	34	32	13	40,6	5	5
PaC	Albicocco	10	0	0	0,0	0	0
PaCITA	Albicocco	12	7	6	85,7	2	2
Pchgms/ Pchcms	Pesco	24	24	12	50,0	5	5
PS	<i>Prunus salicina</i>	6	6	3	50,0	2	1
UDA	Mandorlo	43	39	6	15,4	4	4
UDAp	Albicocco	26	9	3	33,3	1	0
UDP	Pesco	15	15	11	73,3	3	3
Totale		357	300	122	40,7	40	36

Tutte le coppie di primer hanno messo in evidenza un singolo locus, tranne UDA-005 e CPDCT 029 che ne hanno evidenziati 2.

Utilizzando una popolazione segregante F₂, ed essendo i microsatelliti dei marcatori codominanti, la segregazione attesa era 1:2:1. Le analisi hanno invece evidenziato il 45% di segregazione del tipo 3:1, tipico dei marcatori dominanti.

Tabella 7.2 Rapporti di segregazione e posizione sulla mappa T x E, dei loci SSR analizzati

Sigla Primer	Genotipo Parentali	Rapporto Segregazione	Significatività (P < 0,05)			Gruppo Associazione (T x E)
			Pop Unica	Pop 1	Pop2	
BPPCT001	aa :bb	1:2:1	ns	ns	ns	G2
BPPCT016	aa :bb	1:2:1	ns	0,05	ns	G1
BPPCT024	ab :bb	3:1	0,0001	Non Segregante	0,0001	G2
BPPCT025	a- :--	3:1	ns	ns	ns	G6
BPPCT037	aa :ab	1:2:1	ns	ns	ns	G5
BPPCT038	aa :ab	1:2:1	ns	ns	ns	G5
CPDCT029 a	a- :--	3:1	0,0001	Non Segregante	0,0001	--
CPDCT029 b	a- :--	3:1	0,0001	Non Segregante	0,0001	--
CPPCT026	aa :bb	1:2:1	ns	ns	ns	G1
CPPCT030	ab :bb	1:2:1	ns	ns	ns	G6
CPSCT009	aa :ab	1:2:1	ns	ns	0,05	--
CPSCT034	a- :--	3:1	0,0001	Non Segregante	0,0001	G2
EmpaS001	a- :--	3:1	0,0001	Non Segregante	0,0001	--
EmpaS002	ab :bb	1:2:1	ns	ns	ns	--
EPPISF002	aa :bb	1:2:1	ns	ns	ns	--
EPPISF032	a- :--	3:1	ns	ns	0,0001	--
M1	ab :bb	1:2:1	ns	ns	ns	G2
M4	aa :bb	1:2:1	ns	ns	0,05	G1
MA26	-- :-b	3:1	ns	ns	0,001	--
MA27	aa :bb	1:2:1	ns	ns	ns	--
MA7	aa :bb	1:2:1	ns	ns	ns	G1
PaCITA006	a- :--	3:1	0,0001	Non Segregante	0,0001	G4
pchgms26	aa :bb	1:2:1	ns	ns	ns	G2
pchgms27	a- :--	3:1	ns	ns	0,05	G7
pchgms28	aa :bb	1:2:1	ns	ns	ns	G1
PS05C03	a- :--	3:1	0,0001	Non Segregante	ns	G5
UDA-005	aa :bb	1:2:1	ns	ns	ns	G2
UDA-005b	a- :--	3:1	ns	ns	ns	G2
UDA023	a- :--	3:1	0,0001	Non Segregante	0,0001	G2
UDA-025	aa :bb	1:2:1	0,05	ns	ns	G2
UDA-029	a- :--	3:1	0,0001	0,0001	ns	G2
UDP-001	a- :--	3:1	0,0001	Non Segregante	0,0001	G6
UDP-003	a- :--	3:1	0,0001	Non Segregante	0,001	G4
UDP-013	aa :bb	1:2:1	ns	ns	Ns	G2

Trentaquattro loci SSR analizzati nella popolazione O'H x Y il 36,4% ha evidenziato un'elevata distorsione ($P < 0,05$). Dieci marcatori, che ci si aspettava essere segreganti, evidenziavano una fortissima selezione contro uno degli alleli attesi. Solo un ristretto pool di campioni presentava l'allele selezionato. Dal confronto con la mappa T x E, inoltre, si è evinto che i loci distorti erano distribuiti in 6 regioni del genoma in 4 gruppi d'associazione (Dirlewanger et al., 2004). Sulla base delle anomalie riscontrate si è ipotizzato che la popolazione analizzata potesse non derivare dall'autofecondazione di un unico ibrido F1. Quindi sono stati analizzati separatamente i dati provenienti dalle due popolazioni: la prima (Pop1) composta da 40 individui e la seconda (Pop2) composta dai 20 individui che rappresentavano, per i marcatori fortemente distorti, l'allele in difetto (Tabella 7.2).

Nella Pop2 i marcatori sbilanciati restavano tali, di conseguenza si è deciso di considerare gli individui che la costituivano come dei 'fuori popolazione' e costruire la mappa solo con gli individui che definivano la Pop1.

Il numero dei loci SSR analizzati, è stato così ridotto a 24, 17 erano già stati localizzati in mappe d'associazione del genere *Prunus* (T x E) (Dirlewanger et al., 2004; Verde et al., 2005), mentre 6 (CPSCT009, EPPISF002, EPPISF032, MA26, EmpaS001, pchgms26, UDA005b) non erano ancora stati localizzati su nessuna mappa.

7.1.2 Marcatori AFLP (Amplified Fragment Length Polymorphisms)

Sono state saggiati, sui parentali e su 5 individui della popolazione, 30 combinazioni di primer AFLP, per 15 di questi è stato utilizzato il Sequenziatore Capillare Ceq8000 e per le altre 15 è stato utilizzato un apparato per elettroforesi verticale su gel di poliacrilammide; il numero delle bande totali ed il numero di quelle polimorfiche è risultato notevolmente inferiore per le combinazioni corse sul sequenziatore rispetto a quelle corse sull'acrilammide. In totale, in questi saggi preliminari, sono stati valutati 2334 AFLP, con una media di 27,1 frammenti per combinazione, se analizzati con il sequenziatore, e di 128,5 frammenti, se su gel di acrilammide (Tabella 7.3). Le coppie di primer con 3 nucleotidi selettivi (*EcoRI*+NNN/*MseI*+NNN) hanno evidenziato una drastica riduzione del numero delle bande totali, con una media di 13 frammenti per combinazione contro i 55,4 frammenti ottenuti utilizzando coppie nelle quali il primer *EcoRI* ha due sole basi selettive (*EcoRI*+NN/*MseI*+NNN).

Tabella 7.3 Combinazioni di primer AFLP saggate, numero dei frammenti totali e di quelli polimorfici ottenuti

Combinazioni Primer AFLP Ceq8000	N° frammenti totali	N° frammenti polimorfici	Combinazioni Primer AFLP Gel Poliacrilammide	N° frammenti totali	N° frammenti polimorfici
EcoCAC/MseAAC	14	0	EcoCA/MseAAG	142	11
EcoCAC/MseACA	18	1	EcoCA/MseACA	120	16
EcoCAC/MseACT	13	1	EcoCA/MseACC	150	13
EcoCAC/MseACC	21	1	EcoCA/MseAGA	130	11
EcoCAC/MseAGG	6	1	EcoCA/MseAGG	103	10
EcoCGA/MseACC	6	2	EcoCA/MseATA	180	7
EcoCGA/MseACA	10	0	EcoCA/MseAAC	149	13
EcoCGA/MseACT	11	1	EcoCG/MseAA	147	10
EcoCGA/MseAAC	13	0	EcoCG/MseAT	110	9
EcoCGA/MseAGG	18	0	EcoCG/MseAG	105	17
EcoCA/MseAA	60	--	EcoCG/MseAC	93	17
EcoCA/MseAAA	50	1	EcoCA/MseATG	108	5
EcoCA/MseAAT	72	--	EcoCA/MseAGC	120	5
EcoCA/MseAAC	45	0	EcoCA/MseACT	110	4
EcoCA/MseAAG	50	1	EcoCA/MseAAA	160	4
Totale Frammenti	407	9		1927	152

In totale 13 combinazioni di primer, 2 corse sul sequenziatore e 11 corse su gel di acrilammide, sono state saggate su tutti gli individui della popolazione ottenendo 137 frammenti polimorfici. Nella popolazione solo uno dei frammenti ha deviazioni significative ($P < 0,05$) rispetto alla segregazione mendeliana attesa (3:1).

E' rilevante notare come un 23,3% dei frammenti (32 bande), pur assenti nell'ibrido F1 segregava perfettamente all'interno della popolazione (Tabella 7.3). Anche gli AFLP sono stati analizzati separatamente nelle due popolazioni; in questo caso il numero dei loci distorti è 5 nella Pop1 e 15 nella Pop2. Il numero di loci AFLP totali analizzati è stato ridotto a 128.

Tabella 7.4 Numero dei frammenti polimorfici per combinazione che non si evidenziavano nell'ibrido F1, ma mostravano una segregazione perfetta all'interno della popolazione

Combinazioni Primer AFLP	N. Frammenti polimorfici	N. Frammenti polimorfici F1 con banda	N. Frammenti polimorfici F1 senza banda
EcoCA/MseAAC	13	11	2
EcoCA/MseAAG	11	9	2
EcoCA/MseACA	16	13	3
EcoCA/MseACC	13	8	5
EcoCA/MseAGA	11	8	3
EcoCA/MseAGG	10	6	4
EcoCA/MseATA	7	7	0
EcoCG/MseAA	10	7	3
EcoCG/MseAT	9	8	1
EcoCG/MseAG	17	11	6
EcoCG/MseAC	17	14	3
EcoCGA/MseACC	2	2	0
EcoCGA/MseACT	1	1	0
Totale	137	105	32

7.1.3 *Mappa di Associazione*

Per la costruzione della mappa sono stati considerati i dati ottenuti dall'analisi della sola Pop1. Sono stati utilizzati complessivamente 157 marcatori rappresentati, in particolare, da: 4 Caratteri Morfologici (Colore della Polpa, *Y*; Aderenza al nocciolo, *F*; Tomentosità dell'epidermide, *G*; Stony Hard, *Hd*;), 24 SSR, 128 AFLP, 1 GTM (*endo*-PG1, Peace et al., 2005). Trentanove marcatori sono risultati non associati (3 SSR, 34 AFLP, 2 Caratteri Morfologici). Per 8 AFLP è stata possibile l'assegnazione ai rispettivi gruppi d'associazione, ma non è stato possibile definirne l'ordine. Questi AFLP non vengono riportati nella mappa.

La mappa d'associazione (Figura 7.1) costruita è composta da 110 loci: 2 Caratteri Morfologici (Colore della Polpa, *Y*; Aderenza al nocciolo, *F*), 21 SSR, 86 AFLP 1 GTM (*endo*-PG1). I marcatori si sono distribuiti in 16 gruppi d'associazioni invece degli 8 attesi per il genere *Prunus* ($x=8$). La nomenclatura e l'orientamento, dove presenti loci ancora, sono in accordo con quella già utilizzata nelle mappe di *Prunus* pubblicate in precedenza (Viruel et al., 1995; Joobeur et al., 1998; Dettori et al., 2001); i gruppi in cui non vi sono marcatori di riferimento sono stati classificati arbitrariamente come Gruppo A, B, C, D, E, F, G, H, I, L (dal più grande al più piccolo). Il numero di loci in ogni gruppo è compreso fra i 3 del gruppo L ed i 26 del G2, con una media di 6,88 per gruppo. La mappa copre una distanza di 635 cM, la lunghezza di ogni gruppo è compresa fra i 7 cM per il G7 ed i 102 cM del G2, con una media di 39,7 cM per gruppo.

Il Gruppo 1 (G1) è composto da 18 marcatori (1 carattere morfologico, 5 SSR, 12 AFLP) di cui 6 sono loci ancora (*Y*, M4, UDA-025, CPPCT026, BPPCT016, pchgms28) Copre una distanza di 81 cM e sono presenti due gap maggiori di 10 cM. Il Gruppo 2 (G2) è composto da 26 marcatori (9 SSR, 17 AFLP) copre una distanza 102 cM, e possiede 4 gap maggiori di 10 cM. All'interno del G2 sono stati mappati 4 loci microsatelliti ancora (UDP-013, pchgms26, M1, MA7, UDA-005a) e 5 loci SSR mai localizzati all'interno di altre mappe (pchgms26, CPSCT009, MA26 UDA-005b). Il Gruppo 4 (G4) è costituito da 9 marcatori (6 AFLP, 1 GMT, 1 carattere morfologico, 1 SSR) e copre una distanza di 46 cM. Il Carattere Morfologico (*F*) è un locus ancora ed il GTM (*endo*PG1) cosegrega con quest'ultimo, come ipotizzato da Peace et al. (2005). Abbiamo confermato la posizione del marcatore SSR EPPISF032 nel gruppo analizzando i BIN della mappa di riferimento T x E (Howad et al., 2005). Il Gruppo 5 (G5) è formato da 12 marcatori (2 SSR, 10 AFLP), e copre una distanza di 88 cM, con 5 gap maggiori 10 cM; i due loci SSR (BPPCT037, BPPCT038) sono ancora. Il Gruppo 6 (G6) è costituito da 2 marcatori SSR (CPPCT030 è un locus ancora), e ha una lunghezza di 12 cM. Il microsatellite EPPISF002 è stato mappato per la prima volta. L'ultimo gruppo a cui è stato assegnato un numero in base alle altre mappe del genere è il 7 (G7), che è formato da 2 marcatori (1 SSR, 1 AFLP) e copre una distanza di 7 cM, il locus SSR è un'ancora

(pchgms27). I restanti 10 gruppi d'associazione (A, B, C, D, E, F, G, H, I, L) sono composti da soli marcatori AFLP, ad eccezione del gruppo G al cui interno è localizzato un microsatellite mai mappato prima (EmpaS002). Nel loro insieme coprono una distanza di 299 cM; il gruppo A (61 cM) è composto da 8 marcatori AFLP, il gruppo B (25 cM), e gruppo C (12 cM) da 5, il gruppo D (41 cM) ed il gruppo E (44 cM) da 4 AFLP ed i gruppi F, G, H, I, L ognuno da 3 marcatori

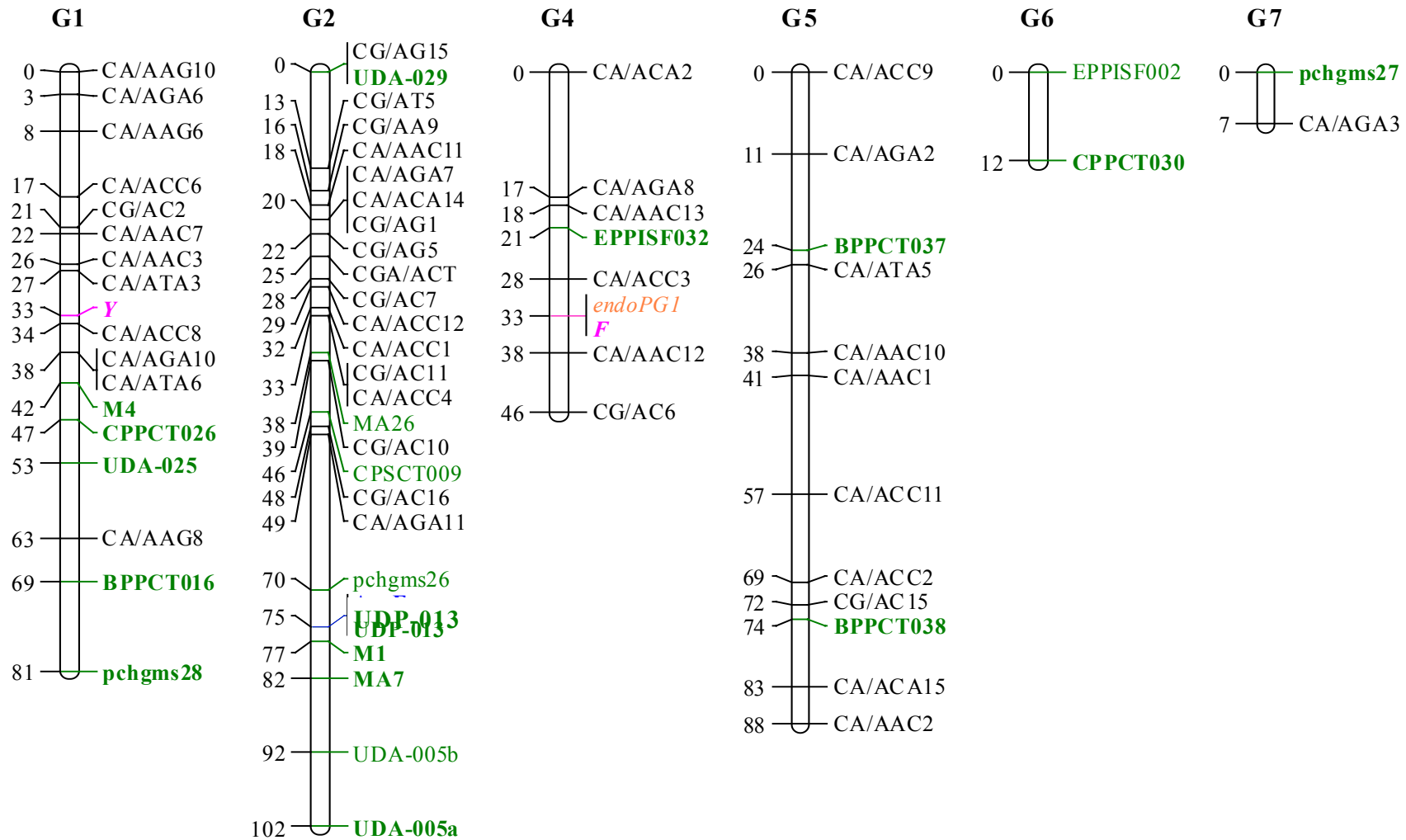


Figura 7.1 Mappa D'Associazione O'H x Y, in grassetto sono evidenziati i loci ancora con altre mappe del genere Prunus, in verde sono i marcatori SSR mai mappati, in fucsia i caratteri morfologici ed in arancio il GTM

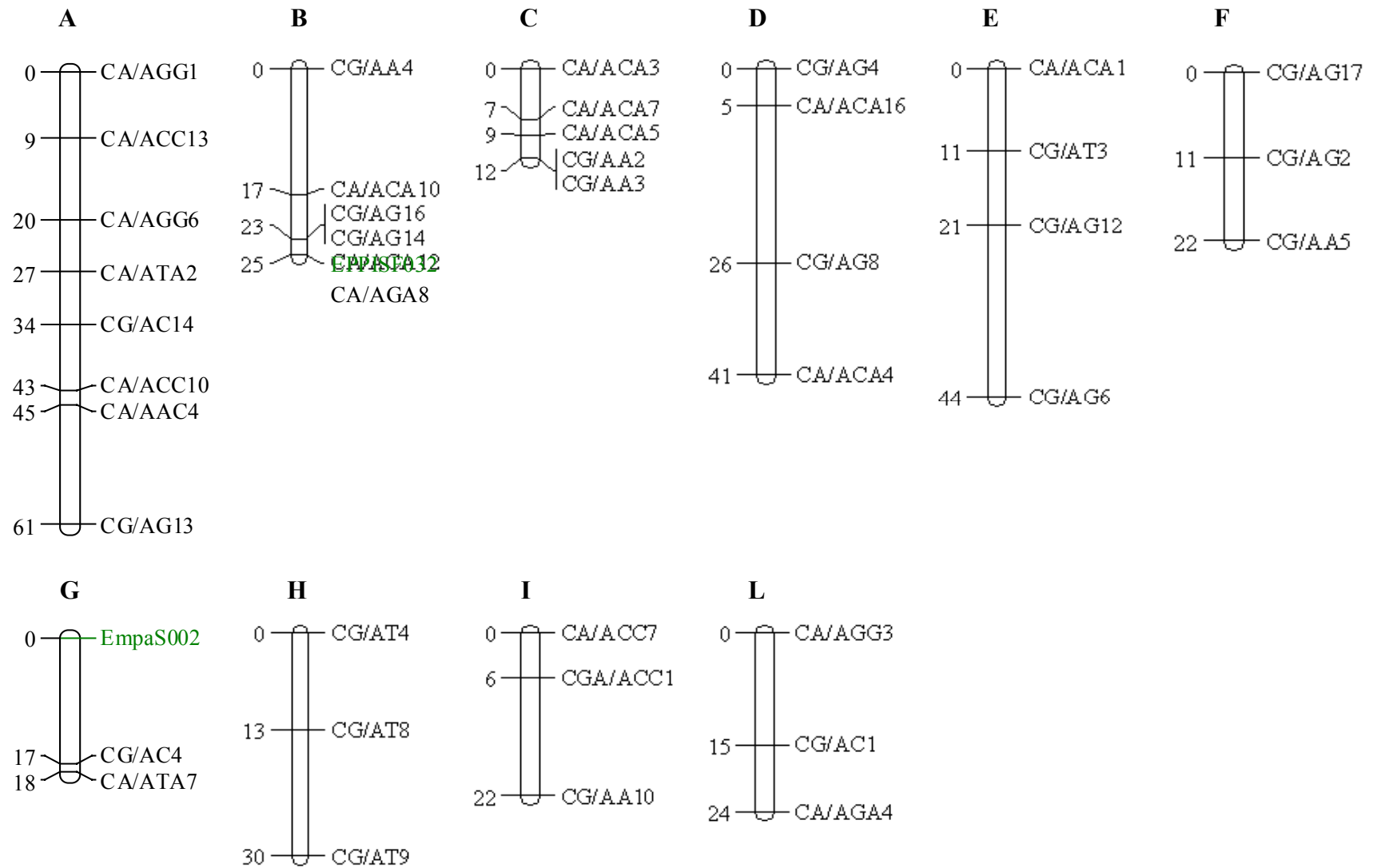


Figura 7.1 continua

7.2 BULKED SEGREGANT ANALYSIS PER L'INDIVIDUAZIONE DI UN MARCATORE ASSOCIATO AL CARATTERE 'STONY HARD'

Al fine di individuare un marcatore strettamente associato al carattere 'stony hard' sono stati costituiti 4 bulk da 8 individui, il cui DNA è stato miscelato in parti uguali. Due bulk sono formati da fenotipi 'stony hard' (Sh1, Sh2) e due da fenotipi 'non stony hard' (NSh1, NSh2). Insieme ai 4 bulk sono stati anche analizzati i due parentali della popolazione. Sono stati saggiati 296 primer RAPD, di questi il 40% (121) ha evidenziato un'amplificazione non accettabile, a causa del basso numero di frammenti prodotti, o non ha evidenziato alcuna amplificazione. In totale sono stati valutati 175 primer RAPD, questi hanno prodotto 1163 frammenti totali di cui 177 evidenziavano nei parentali un polimorfismo del tipo presenza assenza. Le bande che evidenziavano una segregazione fra i bulk ma erano presenti in entrambi i parentali non sono state prese in considerazione. Dei polimorfismi inizialmente individuati (Figura 7.2) solo OPN9 ha confermato il profilo ricercato, presenza del frammento in Yumyeong e nei Bulk 'stony hard' (Sh1, Sh2) ed assenza in O'Henry e nei bulk non 'stony hard' (NSh1, NSh2). Il frammento di interesse è OPN91100.

Il primer OPN9, è stato saggiato sui singoli componenti dei bulk ed ha confermato il risultato, anche se all'interno degli individui catalogati come stony hard 3 non presentano il frammento di interesse (Figura 7.3).

I bulk sono stati anche analizzati con 8 combinazioni di primer AFLP, utilizzando il sequenziatore capillare. In media si sono ottenuti 25 frammenti totali ed 1 frammento polimorfico (presenza o assenza in uno dei due parentali) per combinazione. Degli 8 frammenti polimorfici nessuno segregava all'interno dei bulk come ricercato.

Figura 7.2 Risultati BSA, sono evidenziati i profili d'interesse in cui Yumyeong (Y) ed i 2 bulk 'stony hard' (Sh1,Sh2) presentano una banda, mentre O'Henry (O'H) ed i 2 bulk non 'stony hard' (NSh1, NSh2) ne sono privi

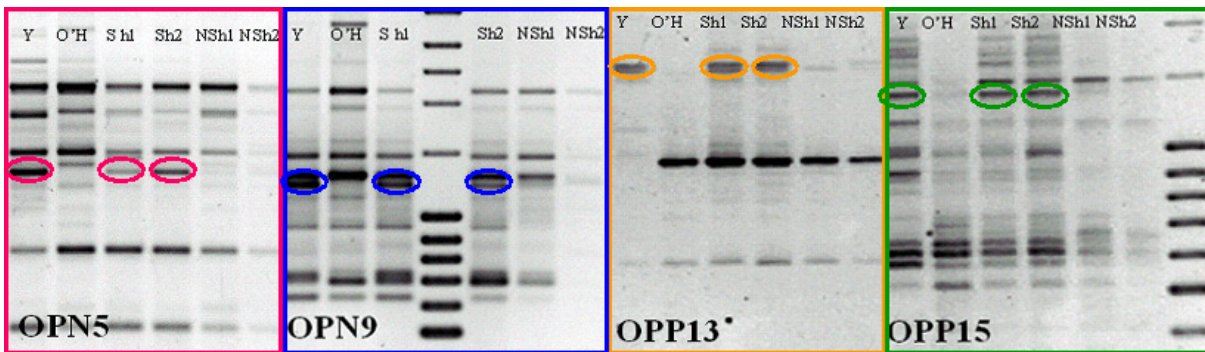
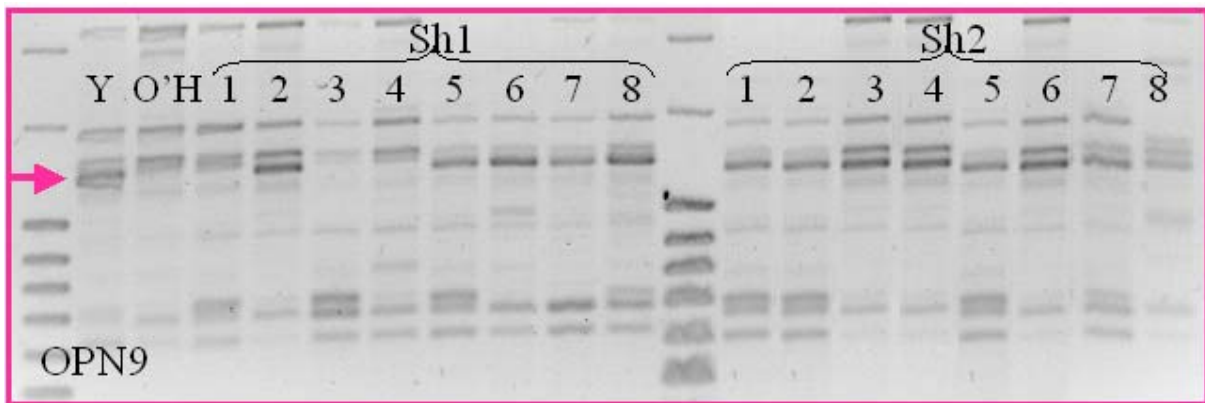


Figura 7.3 Verifica polimorfismo presenza frammento OPN9₁₁₀₀ nei singoli individui dei bulk Sh1 e Sh2

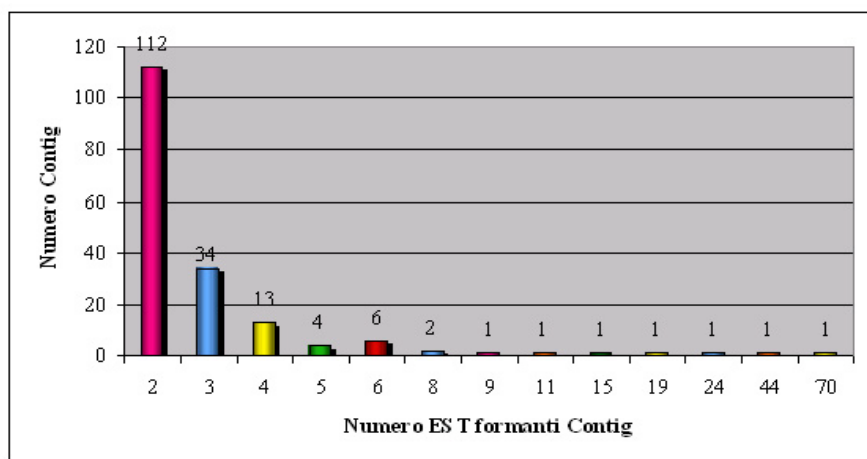


7.3 INDIVIDUAZIONE DI EST CORRELATE ALLA QUALITÀ DEL FRUTTO

Al fine di individuare dell'EST (Expressed Sequence Tag), correlate con la qualità del frutto, è stata costruita una libreria di cDNA da mesocarpo allo stadio S4 della cultivar Yumyeong (portatrice del carattere 'stony hard'); la libreria ottenuta è costituita da 2,4 milioni di cloni e le dimensioni medie dei frammenti, inseriti al suo interno, sono di circa 1200 bp. Sono stati sequenziati 1765 cloni della libreria e la percentuale di successo è stata del 90,7% pari a 1729 sequenze. La dimensione media dei frammenti è stata di 839 bp; circa il 10% (83 bp) di ogni sequenza prodotta apparteneva al vettore di clonaggio utilizzato. Secondo il protocollo descritto nel paragrafo 6.3.4 le sequenze sono state analizzate per eliminare quelle che avevano meno di 100 basi di alta qualità (228; 12%), quelle in cui erano presenti più del 5% di basi sconosciute (7; 0,4%) e quelle sequenze che possedevano più di 40 Adenine o Timine ripetute (5; 0,25%).

In totale sono state prodotte 1525 EST (86,4% delle sequenze iniziali), con in media una lunghezza di 573 bp e con 443 basi di alta qualità. Le sequenze sono state assemblate ottenendo 883 Singlet e 178 Contig; all'interno dei Contig sono raggruppate 642 EST. Il numero delle EST che compongono ogni singolo Contig è compreso fra 2 e 70 (Figura 7.4), con una media di 3,6 EST per Contig. L'assemblaggio delle sequenze ha portato ad una riduzione della ridondanza di circa il 30%.

Figura 7.4 Grafico che illustra il numero di EST per ciascun Contig ottenuto



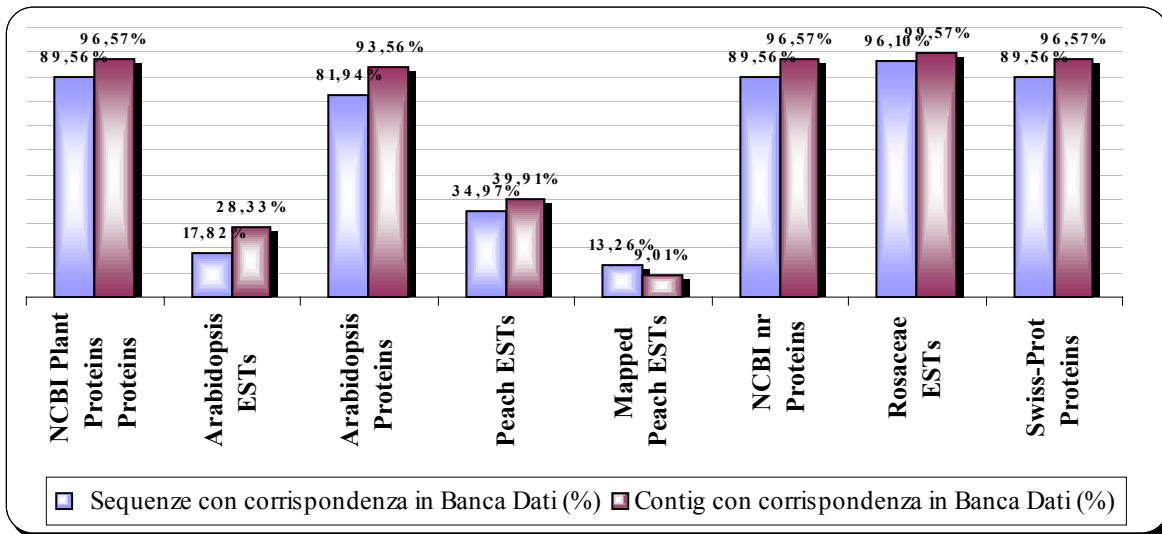
Tutte le EST ottenute ed i contig sono stati confrontati con differenti banche dati, al fine di ricercare il maggior numero di similarità che potessero illustrare funzioni reali o putative dei geni individuati. Nella figura 7.5 sono riportate le percentuali

di similarità del pool di EST e dei contig con 8 banche dati differenti, le percentuali variano da un minimo del 10%, delle EST mappate su una mappa fisica di pesco, ad un massimo del 99%, della banca dati specifica per le rosacee. Sono state individuate 64 nuove EST mai individuate prima all'interno delle *Rosacee*, la funzione di queste nuove sequenze è per l'80% sconosciuta.

Queste EST sono state rese pubbliche inserendole in tre differenti banche dati: l' NCBI database www.ncbi.nlm.nih.gov/; il GDR database www.genome.clemson.edu/gdr/, specifico per le rosacee

e l'ESTree database linuxbox.itb.cnr.it/ESTree/index.php, che contiene esclusivamente EST provenienti da librerie ottenute da mesocarpo di pesco.

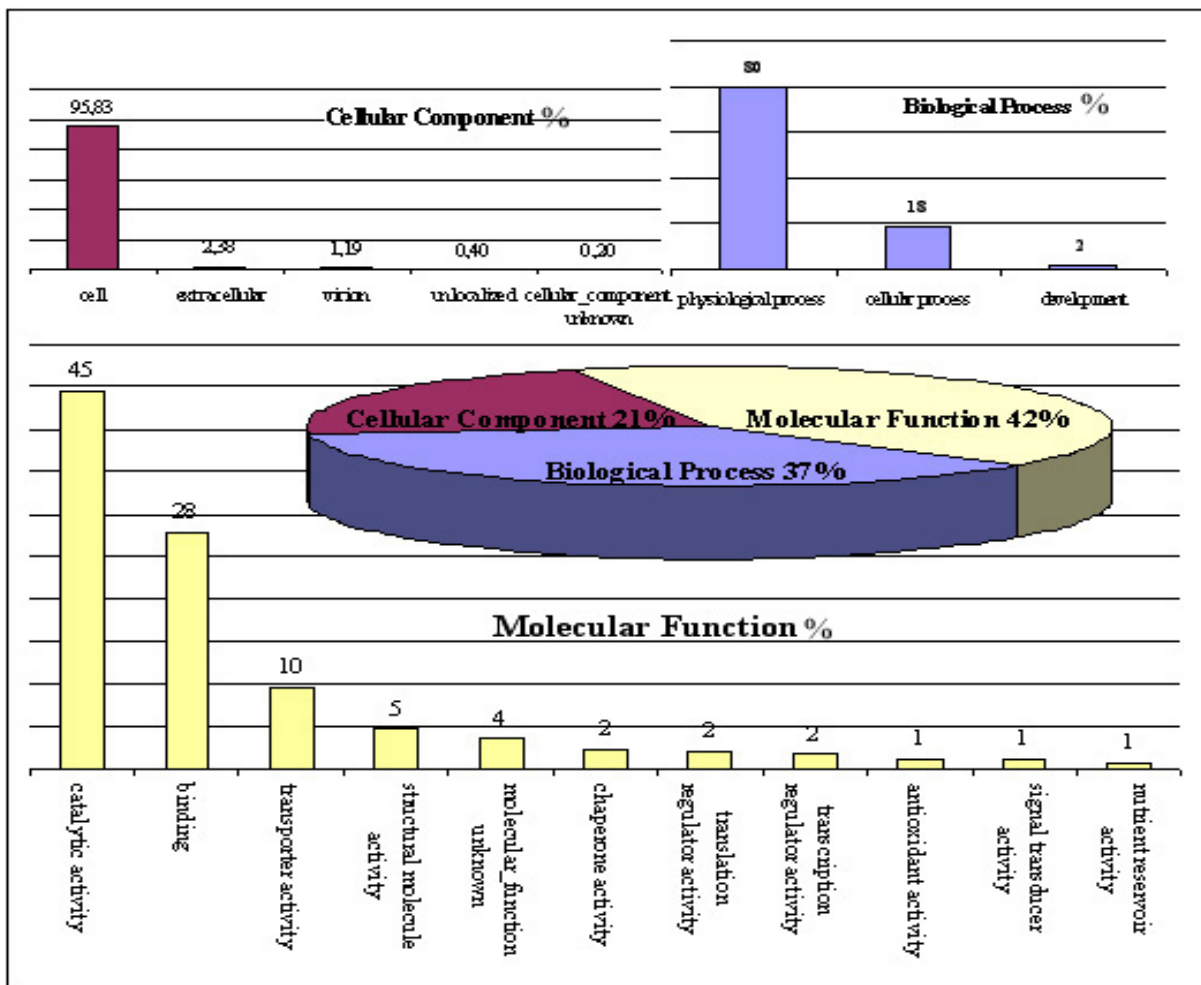
Figura 7.5 Percentuali di similarità delle EST e dei Contig con le maggiori banche dati pubbliche accessibili



All'interno del sito ESTree, le EST prodotte, sono state annotate secondo la GENE ONTOLOGY (GO). L'uso di un'ontologia o "lessico controllato" è molto importante per cercare di ottenere uniformità all'interno delle milioni di sequenze oggi disponibili. GO descrive ogni prodotto genico facendolo ricadere all'interno di tre categorie principali: "Processi Biologici", "Componenti Cellulari" e "Funzioni Molecolari". Queste classi a loro volta contengono altre categorie (Figura 7.6).

Il 42 % delle EST prodotte ricade nella categoria "Funzioni Molecolari"; in questa classe i prodotti genici vengono catalogati in base all'attività senza specificare il tipo di molecola che la svolge, dove, come ed in quale contesto. Il 37 % è stato inserito all'interno dei "Processi Biologici"; la differenza fra questa categoria e la precedente è che un processo viene identificato da diversi passaggi correlati gli uni agli altri. Infine il 21% fa parte dei "Componenti Cellulari" in cui ricadano prodotti genici che devono necessariamente far parte di strutture complesse.

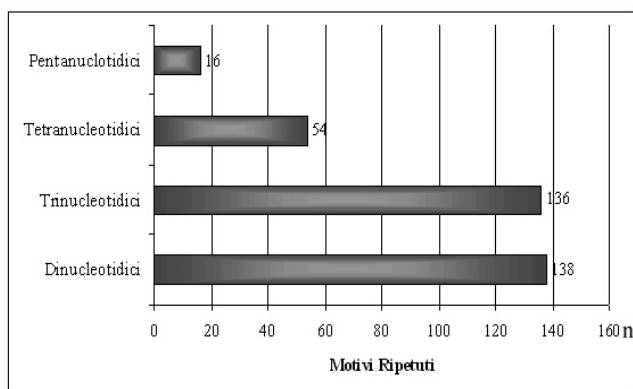
Figura 7.6 Catalogazione EST secondo GENE ONTOLOGY



7.3.1 Individuazione e caratterizzazione di EST-SSR

All'interno delle EST sono state individuate 395 sequenze microsatelliti. Circa il 70% degli SSR è composto da motivi di- e trinucleotidici (Figura 7.7). Utilizzando 32 delle sequenze individuate

Figura 7.7 Distribuzione motivi SSR all'interno delle EST



sono state disegnate delle specifiche coppie di primer che sono state saggate sia in pesco sia in specie appartenenti al genere *Prunus*. Per quello che concerne il pesco i marcatori SSR sono stati saggiati sui parentali e sul relativo ibrido F₁ di tre popolazioni segreganti: una Progenie F₂ (O'H x Y), un Backcross1 (*Prunus persica* x *Prunus ferganensis*, P x F) ed una popolazione F₂ derivante da un ibrido

pesco-mandorlo (T x E) la medesima utilizzata per la costruzione della mappa di riferimento del

genere. I marcatori sono stati anche testati su una varietà di mandorlo (cv Ferragnes), due varietà di albicocco (cv Ninfa, cv Laycot), una varietà di susino europeo (cv Firenze 90), due di susino giapponese (cv Angeleno, cv Green Sun) e due varietà di ciliegio dolce (cv Burlat C1, cv Santa Rosa). Tutte le coppie di primer tranne 3 hanno dato dei prodotti di amplificazione (EPPISF008, EPPISF022, EPPISF024), altre 3 coppie di primer sono state eliminate poiché davano delle amplificazioni di scarsa qualità (EPPISF003, EPPISF005, EPPISF028). Nella tabella 7.5 è riportata la posizione dei microsatelliti all'interno di ogni sequenza e la proteina con la quale è stata rivelata la massima similarità (E-value $\leq 1e^{-6}$).

Tabella 7.5 Posizione del microsatellite all'interno della sequenza e proteina con la quale la sequenza ha la maggior similarità

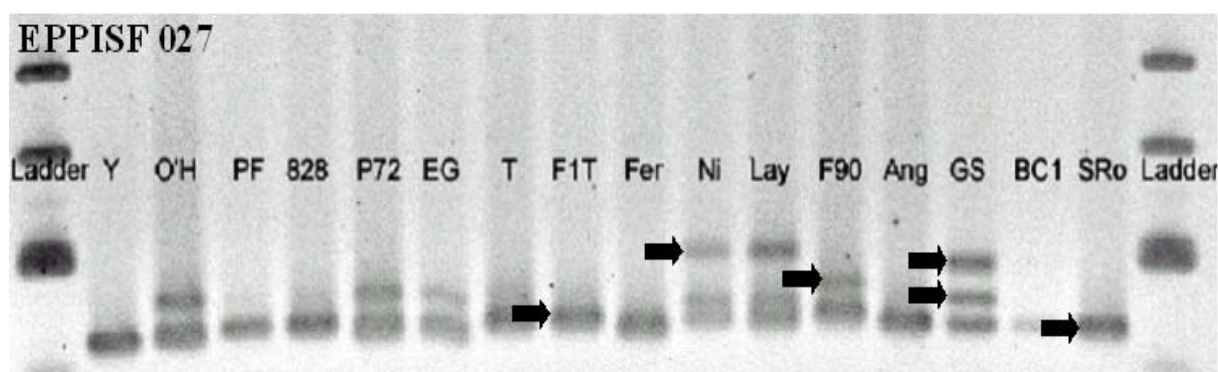
Sigla SSR	Organismo	Proteina	Posizione nella Sequenza
EPPISF001	<i>D. carota</i>	proteina sconosciuta	3'UTR
EPPISF002	no similarità		ORF
EPPISF003	<i>A. thaliana</i>	histone H2B	ORF
EPPISF004	no similarità		ORF/3'UTR
EPPISF005	no similarità		ORF
EPPISF006	<i>N. tabacum</i>	bZIP transcription factor BZI-2	ORF
EPPISF007	<i>A. thaliana</i>	proteina sconosciuta	3'UTR
EPPISF008	<i>Oryza sativa</i>	proteina sconosciuta	ORF
EPPISF009	<i>A. thaliana</i>	proteina sconosciuta	ORF
EPPISF010	<i>Oryza sativa</i>	abscisic acid-inducible protein kinase	5'UTR
EPPISF011	<i>A. thaliana</i>	translocase of inner mitochondrial membrane TIM23	ORF
EPPISF012	<i>Oryza sativa</i>	hypothetical protein	3'UTR
EPPISF013	<i>N. tabacum</i>	early tobacco anther 1	3'UTR
EPPISF014	<i>A. thaliana</i>	proteina sconosciuta	ORF
EPPISF015	<i>A. thaliana</i>	proteina sconosciuta	ORF
EPPISF016	<i>A. thaliana</i>	proteina sconosciuta	5'UTR
EPPISF017	<i>P. persica</i>	MADS6; PpMADS6	5'UTR
EPPISF018	<i>Malus x domestica</i>	cysteine protease inhibitor cystatin	ORF
EPPISF019	no similarità		ORF
EPPISF020	no similarità		5'UTR
EPPISF021	<i>A. thaliana</i>	putative jasmonic acid regulatory protein	ORF
EPPISF022	no similarità		ORF
EPPISF023	<i>H. brasiliensis</i>	latex profilin Hev b 8	5'UTR
EPPISF024	<i>Petunia x hybrida</i>	actin-depolymerizing factor 1	ORF
EPPISF025	<i>A. thaliana</i>	At4g19170	ORF
EPPISF026	<i>A. thaliana</i>	DNA binding protein ACBF, putative	ORF
EPPISF027	no similarità		ORF
EPPISF028	<i>A. thaliana</i>	DNAJ-like protein	ORF
EPPISF029	<i>A. thaliana</i>	calmodulin-like protein	ORF
EPPISF030	<i>A. thaliana</i>	proteina sconosciuta	3'UTR
EPPISF031	<i>T. cacao</i>	carboxypeptidase type III	ORF
EPPISF032	<i>A. thaliana</i>	thiazole biosynthetic enzyme precursor	ORF

Le dimensioni dei prodotti di PCR sono comprese fra 151 e 276 bp, con una media di 204 bp. I 26 primer individuano 61 alleli totali, con una media 2,3 alleli per marcatore. Il marcatore più informativo è EPPISF018 che individua 7 loci differenti, mentre 8 marcatori sono monomorfici. Questo dato può dipendere da due fattori, il primo di carattere tecnico, poiché gli SSR sono stati risolti in gel d'agarosio ed il secondo di carattere teorico, gli EST-SSR venendo da regioni conservate del genoma sono generalmente meno polimorfici. Per tutti marcatori è stata calcolata l'eterozigosità osservata ed attesa (Tabella 7.5).

Quattordici sono risultati essere polimorfici nei parentali della popolazione T x E, 12 sono polimorfici in P x F e 4 in O'H x Y.

Gli SSR hanno sempre amplificato anche nelle diverse specie saggiate, dimostrando la possibilità di un utilizzo interspecifico (Figura 7.7).

Figura 7.7 Corsa Elettroforetica marcatore SSR EPPISF027



(Yumyeong, Y; O'Henry, O'H; *Prunus ferganensis*, PF; *Prunus persica*, 828; ibrido P x F 828; P72; EarlyGold, EG; Texas, T; Ibrido pesco-mandorlo, F1T; Ferragnes, Fer; Ninfa, Ni; Laycot, Lay; Firenze 90, F90; Angeleno, Ang; Green Sun, GS; Burlat C1, BC1; Santa Rosa, SRo)

Tabella 7.6 Sigla Nuovi Primer (EPPISF) e relativo accession number, sequenza dei primer, dimensione prodotti di PCR (pPCR), Temperatura di annealing(Ta), numero alleli putativi (Na), eterozigosità osservata (Ho).

Locus	SSR Motivo	Sequenza Primer 5' => 3'	pPCR (bp)	Ta (C°)	Na	Ho	He
EPPISF001 DN676679	(GATG)5	F: TCCCAAGTGGCTACGAGTGT R: AGGATGTGGAAGAAGGCAAG	243	57	2	0,00	0,5
EPPISF002 DN676920	(AG)7/(GTGGT)3	F: CGACGTGTGACCAAAGGAC R: GCAACTCCATCCACATTTCTC	196	56	5	0,33	0,50
EPPISF004 DN676733/DN676878	(CCA)5 / (TCTCA)9	F: CTGAACCACACAAGCCACAT R: TGAGATGAGATGAGAGCAAGC	180	57	1	0,00	0,00
EPPISF006 DN676689	(ATC)5	F: AAACCCATGAACCCAACAAC R: CACAAACACCACCAATCCAA	235	56	1	0,00	0,00
EPPISF007 DN676717	(TG)7	F: CCTCAGCCACCAATCATTCT R: GGGTTTGGAACATTTTCATC	176	57	2	0,00	0,28
EPPISF009 DN676752	(TCT)5(TCC)3	F: TCATAAATAAACCATGCCATT R: TGACGTTGTGGTGGAAAGAG	182	56	1	0,00	0,00
EPPISF010 DN676766	(AG)10	F: GCCTTGAAAAACCCATGAAG R: CCAAAATTCCCAGACCCAAT	151	57	2	0,17	0,15
EPPISF011 DN676780	(TC)7	F: TGCCCTAACCCTCTAACCCCTA R: GACGTTGGGAGCTGGTAGAG	192	57	4	0,50	0,49
EPPISF012 DN676788	(TC)22	F: AAGTGCCAGACCAAGAATGG R: GATTGCAAGCAAACCCTGAT	197	56	5	0,33	0,29
EPPISF013 DN676814	(TCTGA)3	F: CTGACATCGGAGCTGCTGTA R: CCAGAACCGTGGATCTTCAT	239	57	1	0,00	0,00
EPPISF014 DN676828	(CAG)CCA(CAG)6	F: GCAGAGAGAAAAGCAAAAGCA R: AAAGTGGAAAGAGCCGACCT	246	58	3	1,00	0,28
EPPISF015 DN676843	(TTC)6	F: AGCAATGGCATCAACTTCACT R: ATCCACCAGCCTGCTCCTAT	187	57	2	0,00	0,28
EPPISF016 DN676849	(CTT)7	F: ATGCGAGTCTCTTGGGTCAT R: TGGAAATGAGTTTGGGGTGA	157	57	1	0,00	0,00

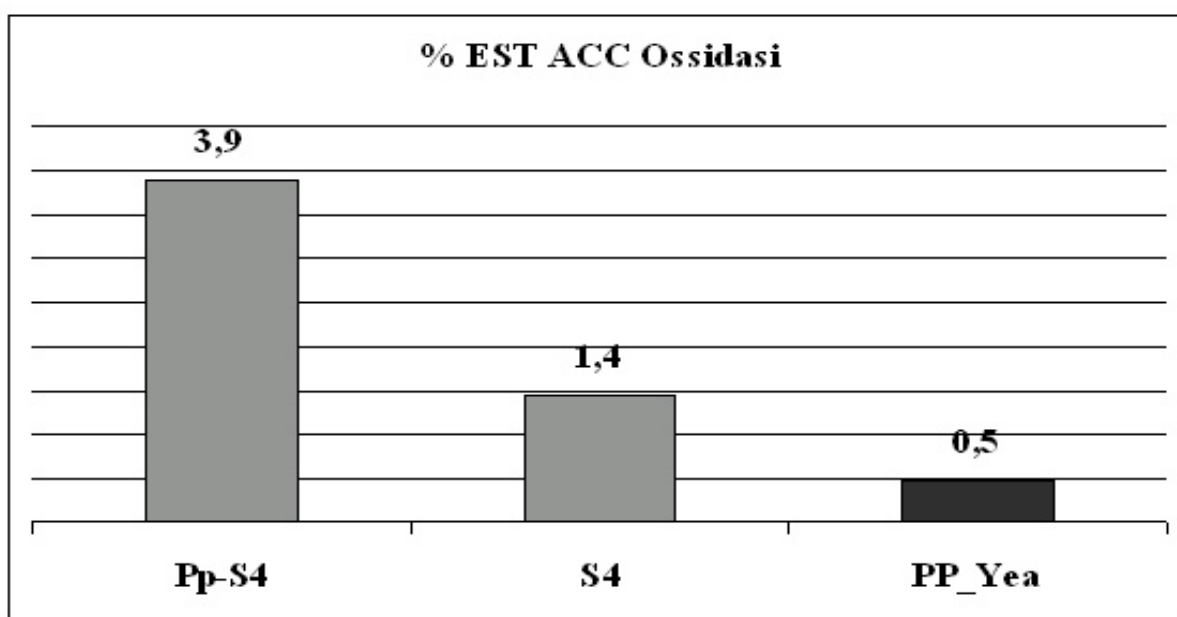
Tabella 7.6 continua

Locus	SSR Motivo	Sequenza Primer 5' => 3'	pPCR (bp)	Ta (C°)	Na	Ho	He
EPPISF017 DN676906	(TTCT)5	F: TTAGAAGAAAGAGAGAGAAGAAGCA R: CCCGCCTGTTGATCTAGTATT	276	57	2	0,17	0,15
EPPISF018 DN676935	(TCT)5(TCC)3	F: TTGGATGAAATTGAAATCGAAA R: GTGATCTTGGACAGCGAAGC	232	56	7	0,67	0,64
EPPISF019 DN676958	(GCT)12	F: GGGGGATAATTTTGCTCTGC R: CCTAACCAACCTCCACCTCCA	225	57	2	1,00	0,50
EPPISF020 DN677047	(CT)9(TC)3	F: TGATTTCGCTTCATCTCACCA R: TCAGAACCACACGAGATTGC	175	56	3	0,50	0,57
EPPISF021 DN677054	(TC)5	F: AGCACCGACTCCTCCACA R: GATGATGGCAACGGGAAG	168	56	1	0,00	0,00
EPPISF023 DN677123	(AGTTC)4	F: GGCAAACATACTCAGAATCA R: ACACACTTCCGTCGTGACC	215	57	2	0,33	0,44
EPPISF025 DN677202	(TC)16	F: TTGGCTTCCTCTTTCCTATCC R: AGTTTTTGGAGGCTGTGGTG	168	57	1	0,00	0,00
EPPISF026 DN677210	(TC)7	F: ACCCCACCACAAGTCCAATA R: CAGCAGCGTAAGGCATGTAA	224	58	1	0,00	0,00
EPPISF027 DN677213	(CAG)6 / (CAG)5	F: TTACATGCCTTACGCTGCTG R: CACCCAGATGGTCCTGATCT	172	57	3	0,50	0,38
EPPISF029 DN677236	(CGT)6	F: GGGGGTTTTTCATTTTTGGTC R: GTCCCGATACCACCAGCTTC	223	57	2	0,17	0,15
EPPISF030 DN677259	(AG)7	F: AGCATCGAGGACTTGGAGAA R: ACCACGCCTAACCTTTACCC	209	57	2	1,00	0,50
EPPISF031 DN677262	(TTC)5	F: CGACCGATGTCTCTTCCTGT R: GACGAGTTTGGGGAGATTGA	243	57	1	0,00	0,00
EPPISF032 DN677272	(ACCC)3/(TC)8	F: TCCCCACAGATATTTTCAGC R: GTCGAGGAGAGAGGGCTTTT	203	57	3	0,33	0,34

7.3.2 *Analisi dell'Espressione delle Endopoligacturonasi (EndoPG) e dell' 1-aminociclopropano-1-carbossilato ossidasi (ACO)*

All'interno delle EST sono stati ricercati i trascritti solitamente espressi del mesocarpo di pesco allo stadio S4 (maturazione fisiologica): Espansine, Endopoligalatturonasi, Pectato Liasi, Pectin Metilesterasi, Endo- β -1,4-glucanasi, β -galattosidasi, β -1,3-glucanasi e 1-aminociclopropano-1-carbossilato Ossidasi (Aco) . Non sono state ritrovate EST correlate a questi trascritti ad eccezione di 6 EST correlate con le espansine e 8 EST corrispondenti all'Aco.

Figura 7.8 Percentuali di EST correlate con Aco in tre diverse librerie provenienti da mesocarpo di pesco allo stadio S4 :Pp-S4 (set EST 1); S4 (set EST 2); PP_Yea (set EST da Yumyeong)



Come già descritto in precedenza l'Aco è uno degli enzimi chiave della biosintesi dell'etilene che è l'ormone responsabile della maturazione nei frutti climaterici. Solitamente è abbondantemente espresso in frutti allo stadio di maturazione fisiologica ed in altre collezioni di EST, provenienti da mesocarpo di pesco, la percentuale di sequenze correlate a questo enzima è compresa fra il 3,9 e l'1,4 %. All'interno del pool delle EST ottenute solo lo 0,5% sono correlate con Aco (Figura 7.8). Per verificare il livello di espressione di Aco si è proceduto con un'analisi northern utilizzando l'RNA estratto dal mesocarpo dalla cultivar Yumyeong (portatrice carattere 'stony hard') e dalla cultivar Françoise con polpa fondente, come controllo positivo, e l'estremità 3'-UTR del gene, come sonda. Inoltre per verificare se la consistenza della polpa di Yumyeong fosse dovuta non alla deficienza di etilene, ma alla non sintesi dell'endopoligalatturonasi, sui medesimi campioni, si è effettuata un'analisi northern, utilizzando, come sonda, l'estremità 3'-UTR del gene per le endo-PG. Il risultato ottenuto mostra come ci sia un basso livello d'espressione di Aco in Yumyeong; questo

causa la non espressione delle endopoligalatturonasi, geni sotto il diretto controllo dell'etilene (Figura 7.9).

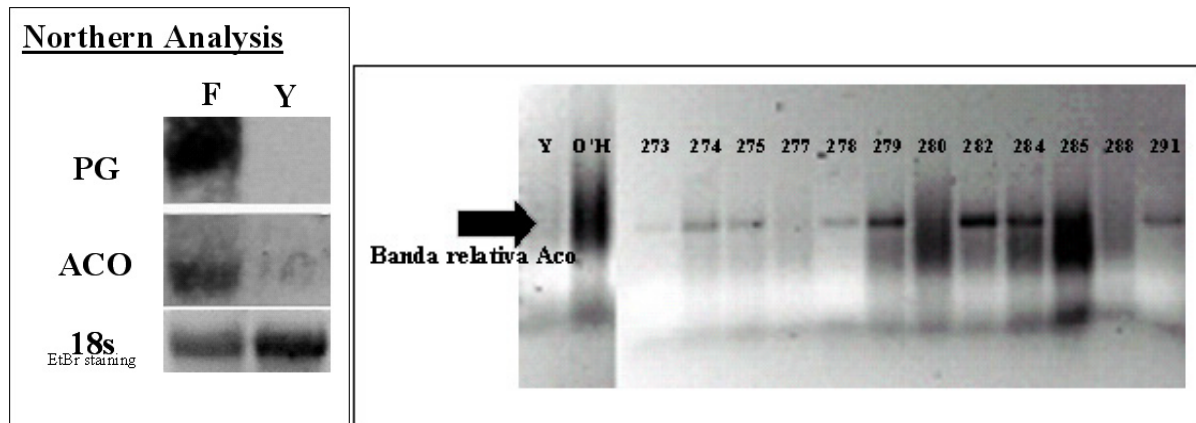


Figura 7.9 Analisi Northern su Françoise (F) e Yumyeong (Y) (ACC Ossidasi, Aco; EndoPoligalatturonasi, PG); Reverse Trascrittase PCR utilizzando primer ACO (Yumyeong, Y; O'Hery O'H, e 12 individui popolazione F2)

Con l'intento di collocare all'interno della mappa d'associazione Aco, due primer disegnati per l'amplificazione della regione 3'-UTR del gene sono stati saggiati sull'RNA totale dei parentali (Yumyeong, O'Henry) e di 40 individui che compongono le due popolazioni putative. Non è stato possibile estrarre l'RNA da tutti i frutti a causa della bassa produttività della popolazione ancora giovane. La presenza/assenza della banda corrispondente all'Aco è stata letta come un marcatore dominante con segregazione del tipo 3:1, ma a causa della bassa numerosità ed al profilo, a volte, di difficile lettura il marcatore non è stato inserito in mappa.

8 DISCUSSIONE

8.1 COSTRUZIONE DELLA MAPPA D'ASSOCIAZIONE Y x O'H

8.1.1 *Marcatore Molecolari*

In accordo con quanto già riportato in bibliografia sulla possibilità di amplificazioni di loci microsatelliti interspecifici (Dirlewanger et al. 2002; Cipriani et al. 1999; Lambert et al. 2004), sono stati amplificati con successo 300 loci SSR isolati in diverse specie del genere *Prunus*. La percentuale di amplificazioni ottenute è dell'84%.

Come riportato anche in Verde et al. (2005) il set UDP si è rivelato il più polimorfico, rispetto agli altri microsatelliti testati, con una percentuale del 73,3% (fra O'Henry e Yumyeong) a fronte di un valore medio del 40,7%. I set BPPCT e CPPCT, anche se provenienti entrambi da librerie genomiche ottenute dalla cultivar di pesco O'Henry ed arricchite in CT, hanno evidenziato diverse percentuali di polimorfismo, rispettivamente del 36,6% e del 21,2%, confermando ciò che era stato evidenziato in precedenza (Verde et al., 2005) Il livello più basso di polimorfismo è stato individuato dai set UDA (mandorlo) ed EmpaS (ciliegio selvatico) in accordo con quanto riportato per amplificazioni effettuate in specie differenti da quella in cui sono stati isolati i microsatelliti (Ellegren et al., 1997; Huang et al., 1998). Considerando i 24 SSR segreganti all'interno della popolazione il livello del polimorfismo rilevato è stato solo del 8 %, drammaticamente più basso rispetto a quanto riportato da altri autori sia in incroci intra che interspecifici (Foulongne et al. 2003; Aranzana et al. 2003; Dirlewanger et al. 2004b; Lambert et al. 2004, Verde et al., 2005). In particolare questo valore è un quarto di quanto evidenziato da Verde et al. (2005), che riportava una percentuale di polimorfismo pari al 33,3%. Le ipotesi che si possono formulare per spiegare questo fenomeno sono : la popolazione F₂ non deriva dall'ibrido F₁ fornita, l'utilizzo di un incrocio intraspecifico e la tecnica di risoluzione utilizzata. L'individuazione di marcatori SSR polimorfici tra i parentali ed eterozigote nella F₁, ma monomorfici nella popolazione F₂ analizzata, supportano l'ipotesi che l'ibrido F₁ fornito, pur essendo compatibile con la provenienza dall'incrocio O'H x Y (non sono mai stati individuati alleli estranei), possa non essere stato utilizzato per la costituzione della popolazione. Un'altra evidenza a supporto di questa tesi è la presenza del 23,3 % di marcatori AFLP che pur segregando correttamente in F₂ non evidenziano la banda segregante nell'ibrido F₁. Un'altra spiegazione del basso livello di polimorfismo è intrinseco dell'incrocio intraspecifico; infatti i dati riportati in bibliografia si riferiscono ad incroci interspecifici all'interno del genere

Prunus. Infine la metodica utilizzata, separazione su gel d'agarosio dei frammenti, pur essendo più agevole ha un minor potere risolutivo se paragonata a quelle che utilizzano come substrato l'acrilammide.

Gli SSR, saggiati sulla popolazione completa con segregazione dominante erano il 45%, di questi il 68,75% presentava una fortissima distorsione; dopo aver eliminato gli individui considerati fuori popolazione 11 dei 16 loci SSR dominanti sono risultati non segreganti. Quindi la percentuale dei microsatelliti con segregazione di tipo dominante si è attestata al 24%, in accordo con quanto osservato da Aranzana et al., (2003) in una progenie F2.

8.1.2 Mappa D'Associazione

Contrariamente a quanto atteso per il pesco ($x = 8$), nonostante l'elevato numero di marcatori utilizzato e dopo aver eliminato i probabili fuori popolazione, sono stati individuati 16 gruppi d'associazione maggiori. Questo dato mette in luce ancora una volta che probabilmente all'interno della popolazione 1 sono presenti degli individui che, pur venendo dallo stesso incrocio, non fanno parte della progenie F2 analizzata. Probabilmente è per le stesse motivazioni che i caratteri pesco/nettarina e stony hard risultano non associati.

La mappa ottenuta copre 635 cM del genoma del pesco, apparentemente questo valore, corretto della compressione dovuta al programma d'analisi utilizzato (Verde et al., 2005), è più alto di quello della mappa di riferimento del genere T x E, 522 cM (Dirlewanger et al., 2004). Le mappe ottenute da incroci interspecifici sono, generalmente, più corte perché presentano delle riduzioni nelle frequenze di ricombinazione (Causse et al. 1994), in particolare T x E mostra delle notevoli soppressioni di ricombinazione rispetto ad altre mappe di *Prunus* (Foulogne et al., 2003; Verde et al., 2005) Questa mappa, quindi, copre circa la metà del genoma del pesco stimato intorno ai 1000 cM considerando un valore di compressione della mappa T x E pari a 0,5 (Verde et al., 2005).

Sei dei gruppi ottenuti presentano dei loci in comune con altre mappe del genere *Prunus*, questo ha consentito di assegnare il numero al gruppo in accordo con quanto già pubblicato. All'interno di questi gruppi sono stati localizzati dei loci SSR mai mappati in precedenza, 3 di essi si sono collocati nel G2 (MA26, CPSCT009, pchgms26); 1 nel G6 (EPPISF002) e un altro nel G4 (EPPISF032). Poiché due porzioni del G4 risultavano associate ad un LOD appena superiore al valore soglia, si è validato questo risultato mappando il microsatellite EPPISF032 nei BIN della popolazione T x E (Howad et al., 2005).

Per i gruppi G1 e G2 è possibile effettuare un'analisi comparativa. Nel G1 sono presenti 3 (M4, CPPCT026, BPPCT016) marcatori in comune con la mappa di Mirabolano x mandorlo-pesco (Dirlewanger et al. 2004b); l'ordine di questi loci risulta essere identico nelle due mappe. Anche nel G2 3 marcatori (M1, UDP013, M7) sono in comune con la stessa mappa, ma in questo caso l'ordine

dei loci non è stato mantenuto essendosi verificata un'inversione tra UDP013 e M1; considerando la ridotta distanza coperta dai 3 marcatori (O'H x Y è di 7 cM; P2175 x GN22 è di 6,2 cM) probabilmente questa non è una reale inversione cromosomica, ma è un artefatto di mappatura.

8.2 BULKED SEGRAGANT ANALYSIS PER L'INDIVIDUAZIONE DI UN MARCATORE ASSOCIATO AL CARATTERE 'STONY HARD'

La Bulk Segregant Analysis (BSA) è una metodica veloce e semplice per individuare marcatori associati a specifici caratteri, *conditio sine qua non* è il possesso di una popolazione segregante per i caratteri in esame. Il successo di questo tipo di approccio, è strettamente correlati alla divergenza dei parentali rispetto ai caratteri studiati; per studiare il colore della polpa, per esempio, si deve disporre di una popolazione segregante derivante dall'incrocio di una cultivar a polpa bianca e di una a polpa gialla.

Il principio della BSA è quello di raggruppare gli individui informativi per il carattere in esame in modo da poter studiare particolari regioni del genoma senza essere influenzati dal 'rumore di fondo' prodotti dai loci non associati al carattere e quindi alla regione di interesse (Michelmore et al, 1991). La dimensione minima del bulk può essere determinata dalla frequenza con la quale loci non associati possono essere individuati come segreganti all'interno dei bulk. Questo valore dipende dal tipo di marcatore utilizzato (dominante o codominante) e dal tipo di popolazione utilizzata (F₂, Backcross, ecc). Utilizzando marcatori RAPD ed una popolazione F₂ la probabilità che un bulk, di n individui, posseda una banda non associata con il carattere e che un secondo bulk, composto del medesimo numero di individui, non abbia la banda viene definito dalla funzione $2(1-(1/4)^n)(1/4)^n$.

Quindi avendo costituito un bulk composto da 8 individui, la probabilità che un marcatore polimorfico sia non-associato al carattere è pari a 3×10^{-5} . Per questo tipo di analisi sono stati scelti i marcatori RAPD poiché individuano un elevato numero di loci, infatti testando circa 200 primer sono state rilevate più di 1000 bande. Solo il 15% dei marcatori evidenziava, nei due parentali (Yumyeong, O'Henry) un polimorfismo del tipo presenza/assenza a conferma dell'elevato livello di omozigosi in pesco (Testolin et al., 2000), Questa condizione è dovuta da un lato all'autogamia della specie e dall'altro alla bassa variabilità genetica tra le varietà dovuta alla loro origine che si può far risalire ad un ristretto pool di genotipi derivanti dalla varietà 'Chinese Cling' (Scorza et al., 1985). Dei quattro profili d'interesse inizialmente individuati, solo uno (OPN9) è stato confermato dopo aver ripetuto la reazione di PCR, questo a conferma della scarsa riproducibilità dei profili ottenuti con marcatori RAPD. Risulta evidente, ancora una volta, come la scelta di un marcatore molecolare sia un compromesso che deriva dalla valutazione dei vantaggi e degli svantaggi del marcatore stesso in relazione alla problematica biologica per la quale lo si vuole applicare.

Il primer polimorfico nei bulk è stato saggiato anche nei singoli individui che li compongono, tre campioni catalogati 'stony hard' non possedevano il frammento OPN9₁₁₀₀ una probabile spiegazione per questa discrepanza è forse dovuta alla difficile identificazione del carattere. In futuro questa banda potrà essere subclonata ed utilizzata come marcatore specifico per l'individuazione di genotipi 'stony hard'.

8.3 INDIVIDUAZIONE DI EST CORRELATE ALLA QUALITÀ DEL FRUTTO

8.3.1 *Assemblaggio delle sequenze ed individuazione di EST-SSR*

Il set di EST sequenziato ed analizzato si discosta dai valori di altri set generati a partire da differenti specie appartenenti alla famiglia delle Rosacee (2763 EST di mandorlo, 9853 EST di pesco, 4668 EST di Rosa) (Jung et al., 2005); in particolare la percentuale di EST che si raggruppano per formare i contig è del 42% contro una media del 63, 53% e di conseguenza la riduzione della ridondanza si attesta intorno al 30% contro un valore medio del 53%. Questi valori sono probabilmente da imputare alla differente numerosità dei set confrontati, infatti il set di EST prodotto è sei volte e mezzo più piccolo.

All'interno delle EST sono state isolate circa 400 sequenze microsatelliti, con una percentuale del 26% rispetto al totale considerando tutte le sequenze SSR > di 10bp, se si considerano solo le sequenze > di 18bp, al fine di poter effettuare un confronto con quanto riportato in bibliografia, la percentuale scende al 7,6%. Questo valore è comunque superiore a quelli riportati che si attestano, intorno al 4% per le Rosacee ed al 2,5% in Arabidopsis (Jung et al., 2005). Questo dato conferma l'utilità delle EST come una buona fonte di marcatori SSR. In accordo con quanto riportato da differenti autori nel caso delle specie vegetali (Miyao et al. 1996; Cardle et al. 2000; Kantety et al. 2002), i motivi dinucleotidici più rappresentati sono AG/GA/TC/CT, mentre i più rari sono GC/CG. Per quello che riguarda i motivi trinucleotidici, pur variando questi da specie a specie (Toth et al., 2000), i più rappresentati nel set di EST analizzato sono AAG/AGA/GAA/CTT/ TTC/TCT in contrasto con quanto riportato in pesco (Jung et al., 2005). Si può anche notare come la distribuzione dei motivi di-, tri-, tetra- e pentanucleotidici sia diversa rispetto ad altri EST-SSR derivanti da pesco, infatti pur rimanendo i motivi trinucleotidici i più rappresentati la loro percentuale è intorno al 50% contro il 65% riportato da Jung et al. (2005). La differenza più importante riguarda gli SSR pentanucleotidici che rappresentano il 12% del totale valore 2,5 volte maggiore rispetto a quanto riportato in pesco (Jung et al., 2005).

La ragione più probabile dell'abbondanza dei motivi trinucleotidici nelle EST e negli esoni è che mutazioni in questi siti possono essere meno dannose rispetto a variazioni di motivi dinucleotidici; l'inserimento di un trinucleotide all'interno di un esone comporta la sintesi di un aminoacido in più,

mentre l'inserimento, la duplicazione o la delezione di un dinucleotide avrebbe come conseguenza lo slittamento di tutto il 'frame' di traduzione.

Alcune di queste sequenze ripetute sono state utilizzate per identificare dei loci microsatelliti. Delle 32 coppie di primer disegnate 3 (10%) non hanno mai dato amplificazione. Dal confronto di questi loci con banche dati, si è osservato che si trovano a cavallo fra due esoni, quindi nel DNA genomico la presenza di un introne porta alla distruzione della sequenza riconosciuta dal primer. La caratterizzazione dei primer SSR disegnati ha messo in evidenza la loro capacità di amplificare anche in specie differenti da quella in cui sono stati originati, la percentuale di amplificazione è pari al 90% considerando un totale di 32 coppie di primer, ma se si considerano le ragioni della non amplificazione delle tre coppie di primer SSR già citate, la percentuale di trasferibilità è del 100%.

Il numero di alleli evidenziati è più basso se riferito a loci SSR isolati a partire da DNA genomico, di conseguenza anche l'eterozigosità osservata ed attesa hanno valori inferiori. Questo dato può dipendere essenzialmente da due fattori: da un lato l'uso dell'agarosio, che ha un potere discriminante inferiore dell'acrilammide, per la visualizzazione delle reazioni di PCR, e dall'altro l'origine stessa degli SSR. Gli EST-SSR hanno una duplice peculiarità: sono meno polimorfici rispetto ai microsatelliti di origine genomica, ma sono più trasferibili. La probabile ragione di questo fenomeno è che il trascrittoma è più conservato.

8.3.2 *La consistenza della polpa: melting, nonmelting e stony hard*

L'intento del sequenziamento di EST dalla cultivar Yumyeong era quello di aver una maggiore definizione del carattere 'stony hard'. L'analisi delle EST relative ai geni noti per essere correlati con i processi di ammorbidimento della polpa, all'interno della cultivar Yumyeong, ha evidenziato una realtà inaspettata; nessuno di questi geni sembrava essere presente, inoltre uno dei geni chiave della maturazione, l'ACC ossidasi, era rappresentato a livelli molto bassi se si confrontano le EST correlate a questo gene in altre librerie. Le informazioni ottenute hanno portato a concludere che le pesche 'stony hard' non producono etilene, a conferma di quanto affermato da Haji et al. (2001). Più in dettaglio si può ipotizzare che questa tipologia di frutti non riesca ad attivare la reazione autocatalitica dell'etilene, che è responsabile del climaterio. Questo dato giustifica, anche l'assenza d'espressione di tutti i geni coinvolti nelle modificazioni della polpa, poiché è noto che questi sono sotto il controllo positivo dell'etilene (Trainotti et al., 2003). Le EST isolate all'interno della libreria hanno una similarità del 99% con le sequenze delle Aco individuate in Pesco (Reperti et al., 2001) ed in Albicocco, ciò porta ad ipotizzare che la mancata sintesi dell'etilene non sia dovuta ad una mutazione dell'enzima. Questo dato viene confermato anche da quanto ottenuto da Haji et al. (2003), che somministrando, alle pesche stony hard, Acido 1-aminociclopropano 1-carbossilico, substrato dell'enzima, hanno osservato la produzione di etilene endogeno. Si può quindi supporre

che il blocco della sintesi avvenga o a causa di una mutazione nella regione del promotore di Aco, o a mutazioni in altri geni che concorrono alla biosintesi dell'etilene come ACS, o a uno dei molti fattori di trascrizione coinvolti.

Un altro dato importante emerso durante queste indagini riguarda i caratteri consistenza della polpa ed aderenza della polpa al nocciolo. Come dimostrato da Peace et al. (2005) questi due tratti sono controllati dallo stesso gene, quello delle *endoPG*, e sono definiti dalla combinazione di 3 alleli differenti. Haji et al. (2005) hanno dimostrato come la cultivar stony hard Yumyeong sia fondente, contrariamente a quanto si pensasse, mentre Peace (comunicazione personale) la cataloga come fondente e aderente. Come è noto il gene delle *endoPG* è sotto il controllo positivo dell'etilene, quindi in assenza dell'ormone non dovrebbe essere espresso portando ad un fenotipo polpa croccante ed aderente. Di conseguenza la classe fenotipica polpa stony hard-spicca non dovrebbe essere presente all'interno della popolazione in esame. Inaspettatamente i dati osservati evidenziano come ci sia una segregazione di tipo indipendente fra questi due caratteri ($\chi^2 = 2,2$ non significativo; gl 3); infatti come già illustrato in precedenza il carattere aderenza al nocciolo e stony hard non sono associati. Una possibile spiegazione di questo fenomeno, da validare e ipotizzata anche da Peace et al. (2005), è la presenza di più copie del gene delle *endoPG* all'interno della stessa regione del genoma, che posseggano promotori differenti. Questi promotori potrebbero essere di due tipologie: una indipendente e l'altra dipendente dall'etilene.

9 CONCLUSIONI

E' stata costruita una mappa d'associazione (O'H x Y, Yumyeong portatore carattere 'stony hard'), all'interno della quale sono stati localizzati due caratteri morfologici strettamente legati alla qualità del frutto. Il colore della polpa e l'aderenza della polpa al nocciolo. Sono stati, inoltre, collocati cinque nuovi loci SSR che potranno essere utilizzati come "ancore" in altre mappe d'associazione all'interno del genere *Prunus*. Il carattere aderenza della polpa cosegrega con e il marcatore endoPG1, confermando l'ipotesi proposta recentemente (Peace et al. 2005).

Analizzando bulk composti da individui con fenotipo 'stony hard', è stato individuato un frammento RAPD, OPN9₁₁₀₀, che sembra essere associato con il carattere in esame. Questo risultato, quando validato, potrà essere un utile strumento nei programmi di miglioramento genetico per la selezione e la valutazione precoce dei semenzali.

Sono state prodotte 1500 EST da mesocarpo di frutto, a maturazione fisiologica della cultivar Yumyeong. Utilizzando le sequenze microsatelliti presenti all'interno delle EST sono stati disegnati e caratterizzati 26 loci SSR. Questi marcatori hanno dato luogo ad amplificazioni in differenti specie del genere *Prunus* ed hanno evidenziato dei polimorfismi nei parentali di tre popolazioni segreganti (F₂ T x E; BC₁ P x F; F₂ O'H x Y). Questi nuovi SSR potranno essere utilizzati da quanti operano nell'ambito della genetica del genere *Prunus*.

Infine è stato confermato, a livello molecolare, ciò che era stato evidenziato da Haji et al. (2001), per via chimica. La cultivar Yumyeong, con fenotipo 'stony hard', non produce etilene ed è questa deficienza che conferisce la peculiare croccantezza della polpa. L'individuazione, all'interno delle EST, ottenute a partire dalla cv Yumyeong, di alcune sequenze omologhe alla porzione codificante dell'enzima ACC ossidasi ha confermato che questo enzima chiave della via biosintetica dell'etilene, è funzionale. Questa evidenza ha permesso di affermare che lo stop nella via biosintetica è probabilmente dovuto a modificazioni della regione promotrice del gene Aco o a mutazioni in uno dei passaggi a monte della sintesi.

Infine, un'importante risultato, anche se ancora da validare, riguarda due caratteri correlati con la qualità del frutto: la consistenza della polpa e l'aderenza della polpa al nocciolo. Si è evidenziato che questi due caratteri, pur dipendendo dallo stesso gene (endopoligalatturonasi), hanno una sensibilità diversa all'etilene; avallando l'ipotesi di più copie del gene nella medesima regione del genoma. Il carattere aderenza della polpa sembra essere etilene indipendente, poiché all'interno della popolazione analizzata (O'H x Y), evidenzia una segregazione di tipo indipendente rispetto al carattere 'stony hard'.

Questo dottorato ha chiarito alcuni aspetti relativi ai caratteri che concorrono a determinare la qualità del frutto in pesco, ma allo stesso tempo, ha posto nuovi interrogativi che attendono di essere risolti.

10 BIBLIOGRAFIA

- Adams M D** et al. (1991) Complementary DNA sequencing: "expressed sequence tags" and the human genome project. *Science*, 252: 1651-1656.
- Aerts J**, Wetzels Y, Cohen N, Aerssens J (2002) Data mining of public SNP databases for the selection of intragenic SNPs. *Human Mutation*, 20: 162 – 173.
- Alexander L** and Grierson D (2002) Ethylene biosynthesis and action in tomato: a model for climacteric fruit ripening. *Journal of Experimental Botany*, 53: 2039-2055.
- Altschul SF**, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J. Mol. Biol.*, 215: 403-410.
- Altschul SF**, Madden, TL, Schaffer, AA, Zhang, J, Miller, W, Lipman, DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, 25: 3389-402. Review.
- Andersen JR**, Lübberstedt (2003) Functional markers in plants. *Trends Plant Sci*, 8: 554–559.
- Aranzana MJ**, Garcia-Mas J, Carbò J, Arús P (2002) Development and variability of microsatellite markers in peach. *Plant Breed*, 1221: 87-92.
- Aranzana MJ**, Pineda A, Cosson P, Dirlewanger E, Ascasibar J, Cipriani G, Ryder CD, Testolin R, Abbott AG, King GJ, Iezzoni AF, Arús P (2003) A set of simple-sequence repeat (SSR) markers covering the *Prunus* genome. *Theor Appl Genet*, 87: 805-815.
- Ayub R**, Guis M, BenAmor M, Gillot L, Roustan JP, Latche A, Bouzayen M, Pech JC (1996) Expression of ACC oxidase antisense gene inhibits ripening of cantaloupe melon fruit. *Nature Biotechnology*, 14: 862-866.
- Bachem CWB**, Oomen RFJ, Visser RGF (1998) Transcript imaging with cDNA-AFLP: a step-by-step protocol. *Plant Mol Biol Rep*, 16: 157–173.
- Bachem CWB**, van der Hoeven RS, de Bruijn SM, Vreugdenhil D, Zabeau M, Visser RGF (1996) Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. *Plant J*, 9: 745–753.
- Bailey JS**, French AP (1941) The genetic composition of peaches. *Mass. Agr. Expt. Sta. Bul.*, 378: 91.
- Bailey JS**, French AP (1949) The inheritance of certain fruit and foliage characters in the peach. *Mass. Agr. Expt. Sta. Bul.* 452.
- Baird WV**, Estager AS, Wells, J (1994) Estimating nuclear DNA content in peach and related diploid species using laser flow cytometry and DNA hybridization. *J Amer Soc Hortic Sci*, 119: 1312-1316.

- Ballester J**, Boskovic R, Batlle I, Arús P, Vargas F, de Vicente MC (1998) Location of the self-incompatibility gene on the almond linkage map. *Plant Breed.*, 117: 69–72.
- Ballester J**, Socias I Company R, Arus P, De Vicente MC (2001) Genetic mapping of a major gene delaying blooming time in almond. *Plant Breeding*, 120: 268-270.
- Barry CS**, Blume B, Bouzayen M, Cooper W, Hamilton AJ, Grierson D (1996) Differential expression of the 1-aminocyclopropane-1-carboxylate oxidase gene family of tomato. *The Plant Journal*, 9: 525-535.
- Barry CS**, Llop-Tous MI, Grierson D (2000) The regulation of 1-aminocyclopropane-1-carboxylate acid synthase gene expression during the transition from system-1 to system-2 ethylene synthesis in tomato. *Plant Physiology*, 123: 979-986.
- Bartolozzi F**, Warburton ML, Arulsekhar S, Gradziel TM (1998) Genetic characterization and relatedness among California almond cultivars and breeding lines detected by randomly amplified polymorphic DNA (RAPD) analysis. *J. Am. Soc. Hort. Sci.*, 123: 381-387.
- Beckman JS**, Soller M (1983) Restriction fragment length polymorphisms in genetic improvement: Methodologies, mapping and cost. *Theor. Appl. Genet.*, 67: 35-43.
- Blake MA** (1932) The J.H. Hale as a parent in peach crosses. *Proc Amer. Soc. Hort. Sci.*, 29: 131-136.
- Bliss FA**, Arulsekhar S, Foolad MR, Becerra V, Gillen AM, Warburton ML, Dandekar AM, Kocsisne GM, Mydin KK (2002) An expanded genetic map of *Prunus* based on an interspecific cross between almond and peach. *Genome*, 45: 520–529.
- Bonghi C**, Ferrarese L, Ruperti B, Tonutti P, Ramina A (1998) Endo-b-1,4-glucanases are involved in peach fruit growth and ripening, and regulated by ethylene. *Physiologia Plantarum*, 102: 346-352.
- Boškovic R**, Tobutt KR (1996) A linkage map of isoenzyme loci in the cherry progeny *Prunus avium* ‘Napoleon’ x *P. nipponica*. *ISHS Acta Horticulturae* 484: Eucarpia Symposium on Fruit Breeding and Genetics.
- Bostein D**, White RL, Skolnick MH, Davis RW (1980) Construction of a genetic map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.*, 32: 314-331.
- Brown GR**, Kadel EE III, Bassoni DL, Kiehne KL, Temesgen B, van Buijtenen JP, Sewell MM, Marshall KA, Neale DB (2001) Anchor reference loci in loblolly pine (*Pinus taeda* L.) for integrating pine genomics. *Genetics*, 159:799–809.
- Brown PO**, Botstein D (1999) Exploring the new world of the genome with DNA microarrays. *Nature Genetics*, 21: 33-37.

- Brummell DA**, Dal Cin V, Crisosto CH, Labavitch JM (2004) Cell wall metabolism during maturation, ripening and senescence of peach fruit. *Journal of Experimental Botany*, 55: 2029–2039.
- Brunel D**, Froger N, Pelletier (1999) Development of amplified consensus genetic markers (ACGM) in *Brassica napus* from *Arabidopsis thaliana* sequence of known biological function. *Genome*, 42: 387–402.
- Bushman W**, Thompson JF, Vargas L, Landy A (1985) Control of Directionality in Lambda Site Specific Recombination. *Science*, 230: 906-911.
- Callahan AM**, Scorza R, Bassett C, Nickerson M, Abeles FB (2004) Deletions in an endopolygalacturonase gene cluster correlate with non-melting flesh texture in peach. *Functional Plant Biology*, 31: 159–168.
- Cardle L**, Ramsay L, Milbourne D, Macaulay M, Marshall D, Waugh R (2000) Computational and experimental characterization of physically clustered simple sequence repeats in plants. *Genetics*, 56:847–854.
- Carpita NC**, Gibeaut DM (1993) Structural model of primary cell walls in flowering plants consistency of molecular structure with the physical properties of the walls during growth. *The Plant Journal*, 3: 1-30.
- Chang C**, Stewart RC (1998) The two-component system: regulation of diverse signaling pathways in prokaryotes and eukaryotes. *Plant Physiology*, 117: 723-731.
- Chaparro JX**, Werner DJ, O'Malley D, Sederoff RR (1994) Targeted mapping and linkage analysis of morphological, isozyme and RAPD markers in peach. *Theor Appl Genet*, 87: 805-815.
- Chen S**, Chung M K, Martin D, Rozich R, Tchou P J, Wang Q (2002) SNP S1103Y in the cardiac sodium channel gene *SCN5A* is associated with cardiac arrhythmias and sudden death in a white family. *J. Med. Genet.*, 39: 913-915.
- Ching A**, Caldwell KS, Jung M, Dolan M, Smith OS, Tingey S, Morgante M, Rafalski AJ (2002) SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines *BMC Genetics*, 3: 19-25.
- Ciardi JA**, Tieman DM, Lund ST, Jones JB, Stall RE, Klee HJ (2000) Response to *Xanthomonas campestris* pv. *Vescicatoria* in tomato involves regulation of ethylene receptor gene expression. *Plant Physiology*, 123: 81-92.
- Cipriani G**, Lot G, Huang WG, Marrazzo MT, Peterlunger E, Testolin R (1999) AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* (L) Batsch]: isolation, characterization and cross-species amplification in *Prunus*. *Theor Appl Genet*, 99: 65–72.

- Claverie M**, Bosselut N, Lecouls AC, Voisin R, Lafargue B, Poizat C, Kleinhentz M, Laigret F, Dirlewanger E, Esmenjaud D (2004) Location of independent root-knot nematode resistance genes in plum and peach *Theor Appl Genet*, 108: 765–773.
- Connors CH** (1920) Some notes on the inheritance of unit characters in peach. *Proc Amer. Soc. Hort. Sci.*, 16: 24-36.
- Cooke R**, Raynal M, Laudie M, Grellet F, Delseny M, Morris PC, Guerrier D, Giraudat J, Quigley F, Clabault G et al. (1996) Further progress towards a catalogue of all Arabidopsis genes: analysis of a set of 5000 non redundant ESTs. *Plant J*, 9: 101–124.
- Cordeiro GM**, Casu R, McIntyre CL, Manners JM, Henry RJ (2001) Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to erianthus and sorghum. *Plant Sci*, 160: 1115–1123.
- Curtoni ES**, DallaPiccola B, De Marchi M, Mattiuz P, Momigliano Richiardi P, Piazza A (1991) *Manuale di Genetica UTET*.
- Dawson DM**, Melton LD, Watkins CB. 1992. Cell wall changes in nectarines (*Prunus persica*). Solubilization and depolymerization of pectic and neutral polymers during ripening and in mealy fruit. *Plant Physiology*, 100: 1203–1210 (and correction *Plant Physiology*, 102: 1062–1063).
- Decroocq V**, Favè MG, Hagen L, Bordenave L, Decroocq S (2003) Development and transferability of apricot and grape EST microsatellite markers across taxa. *Theor. Appli. Gen.*, 106: 912-922.
- DellaPenna D**, Lincoln JE, Fischer RL, Bennet AB (1989) Transcriptional analysis of polygalacturonase and other ripening associated genes in Rutgers, *rin*, *nor*, and Nr tomato fruit. *Plant Physiology*, 90: 1372-1377.
- Dettoni MT**, Quarta R, Verde I (2001) A peach linkage map integrating RFLPs, SSRs, RAPDs and morphological markers. *Genome*, 44: 783-790.
- Di Mascio P**, Kaiser S, Sies H (1989) Lycopene as the most efficient biological single oxygen quencher. *Arch.Biochem. Biophys.*, 274: 532-538.
- Di Vaio C**, Ritieni A (2003) Indici di maturazione e qualità nel pesco. Atti IV convegno Nazionale sulla Peschicoltura Meridionale <http://unipa.it/medpeach/atti/>.
- Dirlewanger E**, Cosson P, Howad W, Capdeville G, Bosselut N, Claverie M, Voisin R, Poizat C, Lafargue B, Baron O, Laigret F, Kleinhentz M, Arus P, Esmenjaud D (2004b) Microsatellite genetic linkage maps of myrobalan plum and an almond-peach hybrid-location of root-knot nematode resistance genes. *Theor Appl Genet*, 109: 827-38.

- Dirlewanger E**, Cosson P, Tavaud M, Aranzana MJ, Poizat C, Zanetto A, Arús P, Laigret F (2002) Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry. *Theor Appl Genet*, 105: 127-138.
- Dirlewanger E**, Graziano E, Joobeur T, Garriga-Calderé F, Cosson P, Howad W, Arús P (2004a) Comparative mapping and marker-assisted selection in Rosaceae fruit crop. *Proc Natl Acad Sci U S A*, 101: 9891-9896.
- Dirlewanger E**, Pascal T, Zuger C, Kervella J (1996) Analysis of molecular markers associated with powdery mildew resistance genes in peach (*Prunus persica* (L.) Batsch) x *Prunus davidiana* hybrids. *Theor Appl Genet*, 93: 909-919.
- Dirlewanger E**, Pronier V, Parvery C, Rothan C, Guye A, Monet R (1998) Genetic linkage map of peach [*Prunus persica* (L.) Batsch] using morphological and molecular markers. *Theor Appl Genet*, 97: 888-895.
- Doyle JJ**, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus*, 12(1): 13-15.
- Du Toit R**, Volsteadt Y, Apostolides Z (2001) Comparison of the antioxidant content of fruit, vegetables and teas measured as vitamin C equivalents. *Toxicology*, 166: 63-69.
- Ellegren H**, Moore S, Robinson N, Byrne K, Ward W, Sheldon BC (1997) Microsatellite evolution - a reciprocal study of repeat lengths at homologous loci in cattle and sheep. *Mol Biol Evol*, 14: 854-860.
- Etienne C**, Rothan C, Moing A, Plomion C, Bodénès C, Svanella-Dumas L, Cosson P, Pronier V, Monet R, Dirlewanger E (2002) Candidate genes and QTLs for sugar and organic acid content in peach [*Prunus persica* (L.) Batsch]. *Theor Appl Genet*, 105: 145-159.
- Ewing B**, Green P (1998) Base-calling of automated sequence traces using phred II. Error probabilities. *Genome Res.*, 8: 186-194.
- Ewing B**, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequence traces using phred I. Accuracy assessment. *Genome Res.*, 8: 175-185.
- FAOSTAT (2005)** FAO Statistical Databases: Agriculture. <http://apps.fao.org>. Accessed.
- Fishman ML**, Levaj B, Gillespie D, Scorza R (1993) Changes in the physico-chemical properties of peach fruit pectin during ontree ripening and storage. *Journal of the American Society for Horticultural Science*, 118: 343-349.
- Fodor S.p.a.** et al. (1991) Light-directed, spatially addressable parallel chemical synthesis. *Science*, 251: 767-773.
- Foolad MR**, Arulsekar S, Becerra V, Bliss, FA (1995) A genetic map of *Prunus* based on an interspecific cross between peach and almond. *Theor Appl Genet*, 91: 262-269.

- Foulongne M**, Pascal T, Arús P, Kervella J (2003) The potential of *Prunus davidiana* for introgression into peach [*Prunus persica* (L.) Batsch] assessed by comparative mapping. *Theor Appl Genet*, 107: 227-238.
- Fourmann M**, Barret P, Froger N, Baron C, Charlot F, Delourme R, Brunel D (2002) From *Arabidopsis thaliana* to *Brassica napus*: development of amplified consensus genetic markers (ACGM) for construction of a gene map. *Theor Appl Genet*, 105: 1196– 1206.
- Fraser LG**, Harvey CF, Crowhurst RN (2004) EST-derived microsatellites from *Actinidia* species and their potential for mapping. *Theor Appl Genet*, 108: 1010–1016.
- Fulton TM**, Van der Hoeven R, Eannetta NT, Tanksley SD (2002) Identification, analysis and utilization of conserved ortholog set markers for comparative genomics in higher plants. *Plant Cell*, 14: 1457–1467.
- Gaziano JM**, Hennekens CH (1993) the role of β -carotene in cardiovascular disease. In *Carotenoids in human health*. Carotenoid in human health, NY Acad. Sci: 148-155.
- Georgi LL**, Wang Y, Reighard GL, Mao L, Wing RA, Abbott AG (2003). Comparison of peach and *Arabidopsis* genomic sequences: fragmentary conservation of gene neighborhoods. *Genome*, 46: 268-276.
- Germano J**, Klein AS (1999) Species-specific nuclear and chloroplast single nucleotide polymorphisms to distinguish *Picea glauca*, *P. mariana* and *P. rubens*. *Theor Appl Genet*, 9:37–49.
- Gibbs AJ**, McIntyre GA (1970) The diagram, a method for comparing sequences. Its use with aminoacid and nucleotide sequences. *Eur. J. Biochem.*, 16: 1-11.
- Giordano M**, Mellai M, Hoogendoorn B, Momigliano-Richiardi P (2001) Determination of SNP allele frequencies in pooled DNAs by primer extension genotyping and denaturing high-performance liquid chromatography. *J Biochem Biophys Methods.*, 47: 101-110.
- Giuliano G**, Al-Babili S, von Lintig J (2003) Carotenoid oxygenases: cleave it or leave it. *Trends Plant Sci*, 8: 145-149.
- Gordon D**, Abanjian C, Green P (1998) Consed: A graphical tool for sequence finishing. *Genome Research*, 8: 195-202.
- Gribskov M**, McLachlan AD, Eisenberg D (1987) Profile analysis: detection of distantly related proteins. *Proc. Nat. Acad. Sci.*, 84: 4355-4358.
- Gupta PK**, Varshney RK, Prasad M (2002) Molecular markers: principles and methodology. In: Jain SM, Brar DS, Ahloowalia BS (eds) *Molecular techniques in crop improvement*. Kluwer, Dordrecht.

- Habu Y**, Fukada-Tanaka S, Hisatomi Y, Iida S (1997) Amplified restriction fragment length polymorphism-based mRNA fingerprinting using a single restriction enzyme that recognizes a 4-bp sequence. *Biochem Biophys Res Commun*, 234: 516–521.
- Haji T**, Yaegaki H, Yamaguchi M (2001) Changes in Ethylene Production and Flesh Firmness of Melting, Nonmelting and Stony Haed Peacjes after Harvest. *J. Japan. Soc. Hort. Sci.*, 70: 458-459.
- Haji T**, Yaegaki H, Yamaguchi M (2003) Softening of stony hard peach by ethylene and the induction of Endogenuos Ethylene by 1-aminociclopropane-1-carboxylate acid (ACC). *Journal of Japanese Society for Horticultural Science*, 72: 212-217.
- Haji T**, Yaegaki H, Yamaguchi M (2005) Inheritance and expression of fruit texture melting, non-melting and stony hard in peach. *Scientia Horticulturae*, 105: 241-248.
- Hamada H**, Petrino MG, Kakunaga T (1982) A novel repeated element with Z-DNA-Forming potential is widely found in evolutionary diverse eukaryotic genomes. *Proc. Natl. Acad. Sci. USA*, 79: 6465-6469.
- Hamilton AJ**, Bouzayen M, Grierson D. (1991) Identification of tomato gene for the ethylene forming enzyme by expression in yeast. *Proceeding of the National Academy of Sciences, USA*, 88: 7434-7437.
- Hardys H**, Balick M, Shierwater B (1992) Application of randomly amplified polymorphic DNA (RAPD) in molecular ecology. *Mol. Ecol.*, 1: 55-63.
- Harker FR**, Redgwell RJ, Hallett IC, Murray SH, Carter G. (1997). Texture of fresh fruit. *Horticultural Reviews*, 20: 121–224.
- Harry DE**, Temesgen B, Neale DB (1998) Codominant PCR-based markers for *Pinus taeda* developed from mapped cDNA clones. *Theor Appl Genet*, 97: 327–336.
- Hartl L**, Weiss H, Stephan U, Zeller FJ, Jahoor A (1993) *Theor. Appl. Genet* 90: 601–606.
- Hayama H**, Shimada T, Ito A, Yoshioka H, Kashimura Y. (2001). Changes in the levels of mRNAs for putative cell wall-related genes during peach fruit development. *Scientia Horticulturae*, 91: 239-250.
- Hayden MJ**, Good G, Sharp PJ (2002) Sequence tagged microsatellite profiling (STMP): improved isolation of DNA sequence flanking target SSRs. *Nucleic Acids Res*, 30: 129
- Hedrick UP**, Howe GH, Taylor OM, Tubergen CB (1917) *The peaches of New York* .Albany, JB Lyon Co., printers.
- Herner RC** and Sink K (1973) Ethylene production and respiratory behavior of the rin tomato mutant. *Plant Physiology*, 52: 38-42.

- Hesse C** (1975) Peaches. In *Advances in fruit breeding*. Edited by J. Janick and J. Moore. Purdue University Press, West Lafayette, Ind. pp. 285–335.
- Howad W**, Yamamoto T, Dirlwanger E, Testolin R, Cosson P, Cipriani G, Monforte AJ, Georgi L, Abbott AG, Arús P (2005) Mapping with a Few Plants: Using Selective Mapping for Microsatellite Saturation of the *Prunus* Reference Map Genetics: Published Articles Ahead of Print as 10.1534/genetics.105.043661.
- Huan X**, Madan A (1999). CAP3: A DNA sequence assembly program. *Genome Research*, 9: 868-877.
- Huang WG**, Cipriani G, Morgante M, Testolin R (1998) Microstellite DNA in *Actinidia chinensis*: isolation, characterization, and homology in related species. *Theor Appl Genet*, 97: 1269-1278.
- Huettel B**, Santra D, Muehlbauer FJ, Kahl G (2002) Resistance gene analogues of chickpea (*Cicer arietinum* L.): isolation, genetic mapping and association with a *Fusarium* resistance gene cluster. *Theor Appl Genet*, 105: 479–490.
- Hurtado MA**, Romero C, Vilanova S, Abbott AG, Llácer G, Badenes M L (2002) Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.), and mapping of PPV (sharka) resistance. *Theor Appl Genet*, 105: 182-191.
- Hurtado MA**, Romero C, Vilanova S, Abbott AG, Llácer G, Badenes ML (2002). Genetic diversity in apricot cultivars based on AFLP markers. *Euphytica*, 127: 297-301.
- Jarvis P**, Lister C, Szabo V, Dean C (1994) *Plant Mol. Biol.*, 24: 685–687.
- Jauregui B** (1998) Localizacion de marcadores moleculares ligados a caracteres agronomicos en un cruzamiento interespecifico almendro-melocotonero. PhD thesis, University of Barcelona, Spain.
- Jeffreys AJ**, Wilson V, Thein SL (1985) Hypervariable “minisatellite” regions in human DNA. *Nature*, 314: 67-73.
- Johnson MS**, Doolittle RF (1986) A method for the simultaneous alignment of three or more aminoacid sequences. *J. Mol. Evol.*, 23: 267-268.
- Joobeur T**, Viruel MA, de Vicente MC, Jáuregui B, Ballester J, Dettori MT, Verde I, Truco MJ, Messeguer R, Batlle I, Quarta R, Dirlwanger E, Arús P. (1998) Construction of a saturated linkage map for *Prunus* using an almond x peach F₂ progeny. *Theor Appl Genet*, 97: 1034-1041.
- Joobeur T**, Periam N, de Vicente MC, King GJ, Arús P (2000). Development of a second generation linkage map for almond using RAPD and SSR markers. *Genome*, 43: 649-655.
- Jourdren C**, Barret P, Brunel D, Delourme R, Renard M (1996) Specific molecular marker of the genes controlling linolenic acid content in rapeseed. *Theor Appl Genet*, 93: 512–518.

- Jun JH**, Chung KH, Jeong SB, Lee HJ (2002) Development of RAPD and SCAR markers linked to flesh adhesion gene in peach. XXVI International Horticultural Congress. Toronto, Canada. p. 335.
- Jung S**, Abbott A, Jesudurai C, Tomkins J, Main D (2005) Frequency, type, distribution and annotation of simple sequence repeats in Rosaceae ESTs.
- Kantety RV**, La Rota M, Matthews DE, Sorrells ME (2002) Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Mol Biol*, 48: 501–510.
- Kantety V**, Rota L, Matthews E, Sorrells E (2002) Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice sorghum and wheat. *Plant Mol Biol*, 48: 501–510.
- Karling S**, Altschul SF (1990) Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc. Nat. Acad. Sci.*, 87: 2264-2268.
- Kashi Y**, King D, Soller M (1997) Simple sequence repeats as a source of quantitative genetic variation. *Nucleic Acid Research*.
- Kende H** (1993) Ethylene biosynthesis. *Annual Review of Plant Physiology and Plant Molecular Biology*, 44: 283-307.
- Kieber JJ**, Rothenberg M, Roman G, Feldman KA, Ecker JR (1993) CTR1, a negative regulator of the ethylene response way in Arabidopsis, encodes a member of the Raf family of protein kinases. *Cell*, 72: 427-441.
- Koch JL**, Nevins DJ (1989) Use of purified tomato polygalacturonase and pectinmethylesterase to identify developmental changes in pectins. *Plant Physiology*, 91: 241-255.
- Konicyzn A**, Ausubel F M (1993) *Plant J.*, 4: 403–410.
- Kosambi DD** (1944) "The estimation of map distances from recombination values." *Ann. Eugen.*, 12: 172-75.
- Lambert P**, Hagen L S, Arús P, Audergon JM (2004) Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.) compared with the almond Texas x peach Earlygold reference map for *Prunus*. *Theor Appl Genet*, 108: 1120-1130.
- Lanahan MB**, Yen HC, Giovannoni JJ, Klee HJ (1994) The Never ripe mutation blocks ethylene perception in tomato. *The Plant Cell*, 6: 521-530.
- Landy A** (1989). Dynamic, Structural, and Regulatory Aspects of Lambda Site-specific Recombination. *Annu. Rev. Biochem.*, 58: 913-949.
- Lauria M**, Rupe M, Guo M, Kranz E, Pirona R, Viotti A, Lund G (2004) Extensive Maternal DNA Hypomethylation in the Endosperm of *Zea mays* *The Plant Cell*, 16: 510-522.

- Lecoals AC**, Rubio MJ, Cabetas JC, Minot R, Voisin A, Bonnet G, Salesses E, Dirlewanger E, Esmenjaud D (1999). RAPD and SCAR markers linked to the Ma1 root-knot nematode resistance gene in Myrobalan plum (*Prunus cerasifera* Ehr.). *Theor. Appl. Genet.*, 99: 328-335.
- Lester DR**, Sherman WB, Atwell BJ. (1996). Endopolygalacturonase and the Melting Flesh (M) locus in peach. *Journal of the American Society for Horticultural Science*, 121: 231–235.
- Lester DR**, Spiers J, Orr G, Brady CJ. (1994). Peach (*Prunus persica*) endopolygalacturonase cDNA isolation and mRNA analysis in melting and nonmelting peach cultivars. *Plant Physiology*, 105: 225–231.
- Liang P**, Pardee AB (1992) Differential display of eukaryotic mRNAs by means of polymerase chain reaction. *Science*, 257: 967–971.
- Litt M**, Luty JA (1989). A hypervariable microsatellite revealed by in vitro amplification of dinucleotide repeat within the cardiac muscle action gene. *Am.J. Hum. Genet.*, 44: 397-401.
- Liverani A**, Giovannini D, Brandi F (2002) Increasing Fruit Quality of Peaches and Nectarines: the Main Goals of ISF-FO (Italy). *Acta Hort.*, 592: 507-514.
- Llop-Tous I**, Barry CS, Grierson D. (2000) Regulation of ethylene biosynthesis in response to pollination in tomato flowers. *Plant Physiology*, 123: 971-978.
- Lu ZX**, Sosinski B, Reighard GL, Baird WV, Abbott AG (1998) Construction of a genetic linkage map and identification of AFLP markers for resistance to root-knot nematodes in peach rootstocks *Genome*, 41: 199-207.
- Lu ZX**, Sossey-Alaoui K, Reighard GL, Baird WV, Abbott AG (1999) Development and characterization of a codominant marker linked to root-knot nematode resistance, and its application to peach rootstock breeding *Theor Appl Genet*, 99: 115-122.
- Lu ZX**, Reighard GL, Baird WV, Abbott AG, Rajapakse S (1996) Identification of peach rootstock cultivars by RAPD markers. *HortScience*, 31: 127-129.
- Lui Y**, Hoffman NE, Yang SF. (1985) Promotion by ethylene of the capacity to convert 1-aminocyclopropane-1-carboxylic acid to ethylene in preclimacteric tomato and cantaloupe fruit. *Plant Physiology*, 77: 407-411.
- Lunt DH**, Hutchinson WF, Carvalho GR (1999) An efficient method for PCR-based Identification of microsatellites arrays (PIMA). *Molecular Ecology*, 8: 893-894.
- Maniatis T**, Fritsch EF, Sambrook J, Nolan C (1982) *Molecular Cloning: A Laboratory Manual* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Manning K** (1990) Isolation of nucleic acids from plants by differential solvent precipitation. *Anal. Biochem*, 195: 45-50.

- Margiotti K**, Kim E, Pearce CL, Spera E, Novelli G, Reichardt JK. (2002) Association of the G289S single nucleotide polymorphism in the HSD17B3 gene with prostate cancer in Italian men. *Prostate*. Sep 15; 53(1): 65-8.
- Martin GB**, Williams JGK, Tanksley SD (1991) *Proc. Natl. Acad. Sci. USA*, 88: 2336–2340.
- Martínez-Gómez P**, Arulsekar S, Potter D, DeJong TM, Bliss FA, Gradziel TM (2001). SSR variation within and between peach and almond cultivars. *Plant & Animal Genome IX Conference*. San Diego, USA. <http://www.intl-pag.org>. Accessed December 15, 2002.
- Messina R**, Lain O, Marrazzo MT, Cpriani G, Testolin R (2004) New set of microsatellite loci isolated in apricot. *Molecular Ecology Notes*, 4.
- Michelmore RW**, Paran I, Kesseli RV (1991). Identification of markers linked to disease resistance genes by bulk segregant analysis. A rapid method to detect markers in specific genomic regions by using segregating populations. *Proc. Natl. Acad. Sci.*, 88: 9828-9832.
- Miller P** (1768) *The gardeners dictionary*, 8th edn. London.
- Miyao A**, Zhong S, Monna L, Yano M, Yamamoto K, Havukkala I, Minobe Y, Sasaki T (1996) Characterization and genetic mapping of simple sequence repeats in the rice genome. *DNA Res*, 3: 233–238.
- Mnejja M**, Garcia-Mas J, Howad W, Arús P (2005) Development and transportability across *Prunus* species of 42 polymorphic almond microsatellites *Molecular Ecology Notes*, 5: 531-536.
- Mnejja M**, Garcia-Mas J, Howad W, Badenes ML, Arús P (2004) Simple-sequence repeat (SSR) markers of Japanese plum (*Prunus salicina* Lindl) are highly polymorphic and transferable to peach and almond. *Molecular Ecology Notes*, 4: 163-166.
- Mohler V**, Klahr A, Wenzel G, Schwarz G (2002) A resistance gene analog useful for targeting disease resistance genes against different pathogens on group 1S chromosomes of barley, wheat and rye. *Theor Appl Genet*, 105: 364–368.
- Monet R** (1989) Peach genetics: past, present and future. *Acta Hort.*, 254: 49–53.
- Montgomery J**, Pollard V, Deikman J, Fischer RL. (1993) Positive and negative regulatory regions control the spatial distribution of polygalacturonase transcription in tomato fruit pericarp. *The Plant Cell*, 5: 1049-1062.
- Mullis K**, Faloona S, Scharf S, Saiki R, Horn G, Erlich H. (1986). Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. *Cold Spring Harbor Symp. Quant. Biol.*, 51: 263-273.
- Nakatsuka A**, Murachi S, Okunishi H, Shiomi S, Nakano R, Kubo Y, Inaba A. (1998) Differential expression and internal feedback regulation of 1-aminocyclopropane-1-carboxylate synthase, 1-

- aminocyclopropane-1-carboxylate oxidase, and ethylene receptor genes in tomato fruit during development and ripening. *Plant Physiology*, 118: 1295-1305.
- Needleman SB**, Wunsch CD (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.*, 48: 443-453.
- Nicholas FJ**, Smith CJS, Schuch W, Bird CR, Grierson D. (1995) High-levels of ripening-specific reporter gene-expression directed by tomato fruit polygalacturonase gene-flanking regions. *Plant Molecular Biology*, 28: 423-435.
- Oeller PW**, Wong LM, Taylor LP, Pike DA, Theologis A. (1991) Reversible inhibition of tomato fruit senescence by antisense RNA. *Science*, 254: 437-439.
- Okie WR** (1998) Handbook of peach and nectarine varieties. USDA Agric. Handbook No. 714.
- Orr G**, Brady CJ. 1993. Relationship of endopolygalacturonase activity to fruit softening in a freestone peach. *Postharvest Biology and Technology*, 3: 121-130.
- Peace CP**, Crisisto CH, Gradziel TM (2005) Endopolygalacturonase: a candidate gene for Freestone and Melting in peach. *Molecular Breeding*, 16: 21-31
- Pearson WR**, Lipman DJ (1988) Improved tools for biological sequence comparison. *Proc. Nat. Acad. Sci.*, 85: 2444-2448.
- Pearson JD**, Lipman DJ (1988). Improved tools for biological sequence comparison. *Proceedings of the National Academy of Science, USA* 85.
- Peri C**, Giovannelli G, Peccarisi T (1997) Linee guida per la messa a punto di sistemi di qualità nelle aziende enologiche. Ed. Chirotti.
- Peters JL**, Constandt H, Neyt P, Cnops G, Zethof J, Zabeau M, Gerats T (2001) A physical amplified fragment-length polymorphism map of Arabidopsis. *Plant Physiol*, 127: 1579-1589.
- Picton S**, Barton SL, Bouzayen M, Hamilton AJ, Grierson D. (1993) Altered fruit ripening and leaf senescence in tomatoes expressing an antisense ethylene-forming enzyme transgene. *The Plant Journal*, 3: 469-481.
- Powell W**, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996) Comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol. Breed.*, 2: 225-238.
- Pressey R**, Avants JK. (1978). Difference in polygalacturonase composition of clingstone and freestone peaches. *Journal of Food Science*, 43: 1415-1423.
- Pressey R**, Hinton DM, Avants JK. (1971). Development of polygalacturonase activity and solubilization of pectin in peaches during ripening. *Journal of Food Science*, 36: 1070-1073.
- Ptashne M**. (1992). *A Genetic Switch: Phage (Lambda) and Higher Organisms* (Cambridge, MA: Cell Press).

- Qin L**, Prins P, Jones JT, Popeijus H, Smant G, Bakker J, Helder J (2001) GenEST, a powerful bidirectional link between cDNA sequence data and gene expression profiles generated by cDNA-AFLP. *Nucleic Acids Res*, 29: 1616–1622.
- Quarta R**, Dettori MT, Verde I, Marchesi U, Palombi A (2001) Characterization and evaluation of genetic diversity in peach germplasm using RAPDs and RFLPs *Acta Horti*, 546: 489-496.
- Quint M**, Mihaljevic R, Dussle CM, Xu ML, Melchinger AE, Lübberstedt L (2002) Development of RGA-CAPS markers and genetic mapping of candidate genes for sugarcane mosaic virus resistance in maize. *Theor Appl Genet*, 105: 355–363.
- Rafalski JA**, Tingey SV (1993). Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. *Trends Genomics*, 9(4): 645-648.
- Rafalski JA**, Tingey SV, Williams JGK (1991). RAPD-markers a new technology for genetic mapping and plant breeding. *AgBiotech. News Inf.*, 3: 645-648.
- Rajapakse S**, Belthoff LE, He G, Estager AE, Scorza R, Verde I, Ballard RE, Baird WV, Callahan A, Monet R, Abbott AG (1995) Genetic mapping in peach using morphological, RFLP and RAPD markers. *Theor Appl Genet*, 90: 503-510.
- Raz V**, Fluhr R (1993) Ethylene signal is transduced via protein phosphorylation events in plants. *The Plant Cell*, 5: 523-530.
- Robertson JA**, Meredith FI, Horvat RJ Senter SD (1990) Effect of cold storage and maturity on the physical and chemical characteristics and volatile constituent of peaches (cv Cresthaven). *J. Agric. Food Chem.*, 38: 620-628.
- Rombauts S**, de Peer YV, Rouze P (2003) AFLPinsilico, simulating AFLP fingerprints. *Bioinformatics*, 19: 776–777.
- Rozen S**, Skaletsky HJ (1998) primer 3 Code disponibile presso: www-genome.wi.mit.edu/genome_software/other/primer3.html.
- Ruperti B**, Bonghi C, Rasori A, Ramina A, Tonutti P. (2001). Characterization and expression of two members of the peach 1-aminociclopropane-1-carboxylate oxidase gene family. *Physiologia Plantarum*, 111: 336-344.
- Sandhu D**, Sidhu D, Gill KS (2002) Identification of expressed sequence markers for major gene-rich regions of wheat chromosome group 1 using RNA fingerprinting-differential display. *Crop Sci*, 42: 1285–1290.
- Savolainen V**, Chase MW (2003) A decade of progress in plant molecular phylogenetics. *Trends Genet*, 19: 717–724.
- Schlötterer C**, Tautz D (1992) Slippage synthesis of simple sequence DNA. *Nucleic Acid Research*, 20: 211-215.

- Scorza R**, Mehlenbacher SA, Lighner GW (1985) Inbreeding and coancestry of freestone peach cultivars of the eastern United States and implications for peach germplasm improvement. *J Am Soc Hortic Sci*, 110: 547-552.
- Scorza R**, Melnicenco L, Dang P, Abbott AG (2002) Testing a microsatellite marker for selection of columnar growth habit in peach [*Prunus persica* (L.) Batsch]. *ISHS Acta Horticulturae*, 592: V International Peach Symposium.
- Scorza R**, Sherman W (1996) Peaches In: Fruit breeding Janick J, Moore JN (Eds), *Fruit Breeding*. Wiley (New York), 325-440.
- Shimomura K**, Yamamoto S, Harayama S, Saga N (2002) Type II DNA topoisomerase (Top2) as promising molecular marker for phylogenetic analysis in Rhodophyta. *Bot Mar*, 45: 87–90.
- Sitrit Y**, Bennett AB. (1998) Regulation of tomato fruit polygalacturonase mRNA accumulation by ethylene: a reexamination. *Plant Physiology*, 116: 1145-1150.
- Sosinski B**, Gannavarapu M, Hager LD, Beck LE, King GJ, Ryder CD, Rajapakse S, Baird WV, Ballard RE, Abbott AG (2000) Characterization of microsatellite markers in peach [*Prunus persica* (L.) Batsch]. *Theor Appl Genet*, 101: 421-428.
- Spanu P**, Reinhardt D, Boller T. (1991) Analysis and cloning of the ethylene-forming enzyme from tomato by functional expression of its messenger-RNA in *Xenopus laevis* oocytes. *EMBO Journal*, 10: 2007-2013.
- Sugimoto KJ**, Kawamata N, Sakajiri S, Oshimi K. (2002) Molecular analysis of oncogenes, ras family genes (N-ras, K-ras, H-ras), myc family genes (c-myc, N-myc) and mdm2 in natural killer cell neoplasms. *Jpn J Cancer Res*. Nov, 93(11):1270-7.
- Sutherland GR**, Richards RI (1995), Simple tandem DNA repeats and human genetic disease, *Proc. Natl. Acad. Sci. USA*, Vol.92: 3636-3641.
- Tanksley SD**, Young ND, Paterson AH, Bonierbale MW (1989) RFLP mapping in plant breeding: new tools for an old science. *BioTechnology*, 7: 257-263.
- Tanksley SD**, Ganai MW, Martin GB (1995) Chromosome landing: a paradigm for map-based gene cloning in plants with large genomes. *Trends in Genetics*, 11: 63-68(6).
- Tao R**, Yamane H, Sassa H, Mori H, Gradziel TM, Dandekar AM, Sugiura A (1997) Identification of stylar RNases associated with gametophytic self-incompatibility in almond (*Prunus dulcis*). *Plant Cell Physiol.*, 38:304–311.
- Tarun AS**, Theologis A. (1998) Complementation analysis of mutants of 1-aminocyclopropane-1-carboxylate synthase reveals the enzyme is a dimer with shared active sites. *Journal of Biological Chemistry*, 273: 12509-12514.

- Tatsuki M**, Mori H. (2001) Phosphorylation of tomato 1-aminocyclopropane-1-carboxylic acid synthase, LE-ACS2, at the C-terminal region. *Journal of Biological Chemistry*, 276: 28051-28057.
- Tautz D** (1989). Hypervariability of simple sequence s as a general source for polymorphic DNA marker. *Nucleic Acid Res.*, 17 (16): 6463-6471.
- Testolin R**, Marrazzo T, Cipriani G, Quarta R, Verde I, Dettori MT, Pancaldi M, Sansavini S (2000) Microsatellite DNA in peach (*Prunus persica* L. Batsch) and its use in fingerprinting and testing the genetic origin of cultivars. *Genome*, 43: 512-520.
- Testolin R**, Messina R, Lain O, Marrazzo MT, Huang WG, Cipriani G (2004) Microsatellites isolated in almond from an AC-repeat enriched library. *Mol Ecol Notes*, 4: 459-461.
- Testoni A**, Eccher Zerbini P (1989) Picking time and quality in apple storage. *Acta Hort. (ISHS)*, 258: 445-454.
- Thamarus KA**, Groom K, Murrell J, Byrne M, Moran GF (2002) A genetic linkage map for *Eucalyptus globulus* with candidate loci for wood, fiber, and floral traits. *Theor Appl Genet*, 104: 379–387.
- Theologis A**, Oeller PW, Wong LM, Rottmann WH, Gantz DM. (1993) Use of a tomato mutant constructed with reverse genetics to study fruit ripening, a complex developmental process. *Developmental Genetics*, 14: 282-295.
- Toth G**, Gaspari Z, Jurka J (2000) Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Res.*, 10: 967–981.
- Trainotti L**, Zanin D, Casadoro G (2003) A cell wall-oriented genomic approach reveals a new and unexpected complexity of the softening in peaches. *Journal of Experimental Botany*, 54: 1821-1832.
- Tuatz D**, Renz M (1984) Simple sequences are ubiquitous repetitive components of eukariotic genomes. *Nucleic Acid Res.*, 12: 4127-4138
- Van Ooijen JW**, Sandbrink JM, Vrielink M, Verkerk R, Zabel P, Lindhout P (1994) An RFLP linkage map of *Lycopersicon peruvianum*. *Theor Appl Genet*, 89: 1007-1013.
- Van Ooijen JW**, Voorrips RE (2001) JoinMap® 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, the Netherlands.
- Vaughan SP**, Russel K (2004) Characterization of novel microsatellites and development of multiplex PCR for large-scale population studies in wild cherry, *Prunus avium*. *Molecular Ecology Notes*, 4: 429-431.

- Verde I**, Vendramin E, Dettori MT, Quarta R, Wang Y, Lecouls AC, Abbott AG (2004) Target SSR Development in Peach and SSRs Mapping in a Peach BC₁ Progeny. *Acta Hort.*, 663: 63-68.
- Verde I**, Lauria M, Dettori MT, Vendramin E, Balconi C, Micali S, Wang Y, Marrazzo MT, Cipriani G, Hartings H, Testolin R, Abbott AG, Motto M, Quarta R (2005) Microsatellite and AFLP markers in the *Prunus persica* [L. (Batsch)] x *P. ferganensis* BC₁ linkage map: saturation and coverage improvement. *Theoretical and Applied Genetics*, 111: 1013-1021.
- Verhage BA**, van Houwelingen K, Ruijter TE, Kiemeney LA, Schalken JA. (2004) Single-nucleotide polymorphism in the E-cadherin gene promoter modifies the risk of prostate cancer. *Int J Cancer*. 109(3):480-1.
- Vilanova S**, Romero C, Abbott AG, Llàcer G, Badenes ML (2003) An apricot (*Prunus armeniaca* L.) F₂ progeny linkage map based on SSR and AFLP markers, mapping a plum pox virus resistance and self-incompatibility traits. *Theor Appl Genet* 107: 239-247.
- Viruel MA**, Messeguer R, de Vicente MC, Garcia-Mas J, Puigdomènech P, Vargas F, Arús P (1995) A linkage map with RFLP and isozyme markers for almond. *Theor Appl Genet*, 91: 964-971.
- Vision TJ**, Brown DG, Shmoys DB, Durrett RT, Tanksley SD (2000) Selective mapping: a strategy for optimizing the construction of high-density linkage maps. *Genetics*, 155: 407-420.
- Vos P**, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res*, 23: 4407-4414.
- Wang Y**, Georgi L, Zhebentyayeva T, Reighard GL, Scorza R, Abbott AG (2002) High-throughput targeted SSR marker development in peach (*Prunus persica*). *Genome*, 45: 319–328.
- Warburton ML**, Bliss FA (1996) Genetic diversity in peach (*Prunus persica* L. Batch) revealed by randomly amplified polymorphic DNA (RAPD) markers and compared to inbreeding coefficients. *J. Am. Soc. Hort. Sci.*, 121: 1012-1019.
- Waterman MS** (1989) Mathematical analysis of molecular sequences (special issue). *Bull. Math. Biol.* Pergamon Press, NY.
- Waugh R**, Powell W (1992) Using RAPD markers for crop improvement. *Trends Biotechnol.*, 10: 186-191.
- Weisberg RA**, Landy A (1983) Site-Specific Recombination in Phage Lambda. In *Lambda II*, R. A. Weisberg, ed. (Cold Spring Harbor, NY: Cold Spring Harbor Press), pp. 211-250.
- Weising K**, Nybom H, Wolff K, Meyer W (1995) DNA Fingerprinting in Plants and Fungi (ed. Arber, A.) CRC Press, Boca Raton pp. 1–3.

- Wilkinson J**, Lanahan M, Yen H, Giovannoni J, Klee H. (1995) An ethylene-inducible component of signal transduction encoded by Never-ripe. *Science*, 270: 1807-1809.
- Williams JGK**, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic marker. *Nucleic Acid Res.*, 18 (22): 6531-6535.
- Yamamoto M**, Miki T, Ishiki Y, Fujinami K, Yanagisawa Y, Nakagawa H, Ogura N, Hirabayashi T, Sato T. (1995) The synthesis of ethylene in melon fruit during the early-stage of ripening. *Plant and Cell Physiology*, 36: 591-596.
- Yamamoto T**, Mochida K, Imai T, Shi YZ, Ogiwara I, Hayashi T (2002) Microsatellite markers in peach [*Prunus persica* (L.) Batsch] derived from an enriched and cDNA libraries. *Molecular Ecology Notes*, 2: 298-301.
- Yang SF**, Hoffman NE. (1984) Ethylene biosynthesis and its regulation in higher plants. *Annual Review of Plant Physiology*, 35: 155-189.
- Yang W**, dongBai X, Eaton C, Kamoun S, van der Knaap E, Francis D (2004) Discovery, mapping, and application of single nucleotide polymorphisms in *Lycopersicon esculentum*. www.oardc.ohio-state.edu/tomato/wypag247.pdf.
- Yoshida M** (1976) Genetical Studies on the fruit quality of peach varieties: texture and keeping quality. *Bulletin of Fruit Tree Research Station A3*, 1-16.
- Zane L**, Bargelloni L, Patarnello T (2002) Strategy for microsatellite isolation: a review. *Molecular Ecology*, 11: 1-16.
- Zarembinski TI**, Theologis A. (1994) Ethylene biosynthesis and action: a case of conservation. *Plant Molecular Biology*, 26: 1579-1597.
- Zegzouti H**, Jones B, Frasse P, Marty C, Maitre B, Latche A, Pech JC, Bouzayen M. (1999) Ethylene-regulated gene expression in tomato fruit: characterization of novel ethylenesponsive and ripening-related genes isolated by differential display. *The Plant Journal*, 18: 589-600.
- Zhu YL**, Song QJ, Hyten DL, Van Tassell CP, Matukumalli LK, Grimm DR, Hyatt SM, Fickus EW, Young ND, Cregan PB (2003) Single-nucleotide polymorphisms in soybean. *Genetics*, 163: 1123–1134.