



UNIVERSITA' DELLA CALABRIA



UNIVERSITÀ DELLA CALABRIA
Dipartimento di Ecologia
Dottorato di Ricerca in Biologia Vegetale
XXI ciclo

**CARATTERIZZAZIONE DEL PROTEOMA
DI FRAGOLA DURANTE LA MATURAZIONE
ED IN DIVERSI GENOTIPI**

BIO/11

**Coordinatore
Prof. Aldo Musacchio**

**Docente guida
Dr. Gaetano Perrotta**

**Dottoranda
Dr.ssa Linda Bianco**

ENEA

INDICE

SCOPO DELLA TESI		3
1	INTRODUZIONE	5
1.1	Il proteoma	5
1.2	La proteomica	6
1.2.1	La proteomica sistematica e l'elettroforesi bidimensionale	8
1.2.2	La proteomica differenziale e l'analisi DIGE	11
1.2.2.1	Strategie alternative alla DIGE per la quantificazione	15
1.2.3	Metodi proteomici non basati sull'elettroforesi bidimensionale	18
1.3	La spettrometria di massa in proteomica	24
1.4	Gli spettrometri di massa	27
1.4.1	Sorgenti ioniche	29
1.4.1.1	Ionizzazione laser assistita da matrice (MALDI)	29
1.4.1.2	Ionizzazione per elettro-nebulizzazione (ESI)	30
1.4.2	Gli analizzatori di massa	32
1.4.2.1	Analizzatori a tempo di volo	32
1.4.2.2	Quadruplo e Trappole ioniche	34
1.4.2.3	FT-ICR	37
1.4.2.4	Analizzatori ibridi	38
1.5	Identificazione delle proteine attraverso analisi MS	41
1.5.1	Approccio top-down	41
1.5.2	Approccio bottom-up	42
1.5.2.1	Peptide mass fingerprinting (PMF)	42
1.5.2.2	Sequenziamento peptidico	43
1.5.2.2.1	Sequenziamento peptidico in silico	43
1.5.2.2.2	Sequenziamento de novo	44
1.6	La proteomica vegetale	46
1.7	La maturazione dei frutti	49
1.7.1	La maturazione della fragola	52
1.7.1.1	Cenni generali e cambiamenti macroscopici	52
1.7.1.2	Cambiamenti microscopici	54
1.7.1.3	Cambiamenti biochimici e molecolari	54
1.7.2	L'approccio proteomico per lo studio della maturazione in fragola	57
2	MATERIALI E METODI	61
2.1	Materiale vegetale	61
2.2	Estrazione proteica	61
2.3	Elettroforesi monodimensionale	64
2.4	Elettroforesi bidimensionale	64
2.5	Analisi DIGE	65
2.6	Acquisizione delle immagini ed analisi dei dati	66
2.7	Procedure di digestione proteica	67
2.8	Analisi nLC-ESI-IT-MS/MS	68
2.9	Analisi 2-D nLC-ESI-IT-MS/MS	69
2.10	Identificazione proteica	69
3	RISULTATI	71
3.1	Mappa 2-D del frutto di fragola	71
3.2	Identificazione delle proteine mediante approcci shotgun	79
3.2.1	Distribuzione funzionale delle proteine identificare	83
3.3	Variazioni del proteoma durante la maturazione	85
3.3.1	Proteine regolate in R vs I	88
3.3.2	Proteine regolate in R vs T	89
3.3.3	Proteine regolate in T vs I	89
3.4	Variazioni del proteoma nel confronto tra genotipi	99

4	DISCUSSIONE	101
4.1	Panoramica sulla mappa bidimensionale	101
4.2	Analisi <i>shotgun</i> : identificazione proteica su larga scala	104
4.3	Proteine coinvolte nel processo di maturazione	108
4.4	Le proteine correlate ai genotipi	115
5	CONCLUSIONI	117
6	BIBLIOGRAFIA	119
7	PUBBLICAZIONI, PARTECIPAZIONE A CONGRESSI E CORSI	131

SCOPO DELLA TESI

Le principali caratteristiche organolettiche della fragola matura, quali sapore, colore, profumo e consistenza, così come le relative proprietà nutraceutiche, rappresentate dal contenuto in fibre, vitamine ed antiossidanti, sono caratteri qualitativi, la cui espressione è strettamente dipendente da alcuni processi fisiologici, tra cui la maturazione.

La maturazione dei frutti rappresenta un fenomeno complesso, poiché richiede l'espressione di diversi geni e l'azione coordinata di vari enzimi e macromolecole proteiche. Il fenotipo maturo è la somma di tutti quei cambiamenti, biochimici e fisiologici, che avvengono proprio durante lo sviluppo. Sebbene questi cambiamenti siano variabili da specie a specie, essi includono generalmente modificazioni nell'ultrastruttura della parete cellulare, la conversione dall'amido in zuccheri, l'accumulo di pigmenti e di sostanze aromatiche.

Il presente lavoro di Tesi si propone come obiettivo principale l'individuazione di proteine coinvolte nell'espressione di caratteri qualitativi nel frutto di fragola, attraverso la determinazione della composizione del proteoma, investigato durante la maturazione ed in diversi genotipi.

Nello specifico, una mappa delle proteine presenti nel frutto della cultivar (cv) di riferimento, Queen Elisa (QE), è stata ottenuta mediante elettroforesi bidimensionale (2D-E) delle proteine, estratte a tre diversi stadi di maturazione: immaturo (I), turning (T) e rosso (R). Attraverso la combinazione di cromatografia liquida bidimensionale (2D-LC) e SDS-PAGE, entrambe accoppiate alla spettrometria di massa tandem (nLC-ESI-IT-MS/MS), è stato identificato e descritto un vasto set di proteine (1370), espresse in fragola durante la maturazione.

La tecnologia DIGE (*differential in gel electrophoresis*), inoltre, è stata applicata all'analisi dell'espressione proteica differenziale, durante la maturazione e nel confronto tra la cv di riferimento QE ed i genotipi parentali, Miss (M) e USB35 (U). Ciò ha permesso di identificare proteine accumulate in maniera differenziale, verosimilmente correlate all'espressione di tratti qualitativi.

L'identificazione di probabili target molecolari, coinvolti nell'espressione di caratteri qualitativi in fragola, rappresenta un fattore di fondamentale rilevanza per i progetti di miglioramento genetico, finalizzati a conferire un valore aggiunto ad alimenti di origine vegetale. L'introduzione di tratti desiderati in una specie d'interesse agro-alimentare è, di fatto, subordinata alla conoscenza dettagliata dei flussi metabolici che ne controllano l'espressione.

Solo attraverso la comprensione approfondita di questi eventi, che rappresenta lo scopo principale della presente Tesi di Dottorato, è possibile incrementare il potenziale dei moderni sistemi di miglioramento genetico, mediante specifici programmi di breeding e/o sistemi di ingegneria metabolica.

1 INTRODUZIONE

1.1 Il Proteoma

Il termine “proteoma”, coniato nel 1994 (*Conference on Genome and Protein Maps*, Siena) come equivalente linguistico del concetto di genoma, descrive il *set* completo di proteine espresse in un sistema cellulare di un dato organismo, in un determinato momento, sotto una determinata condizione (Wilkins et al., 1995). Mentre il genoma di una cellula o di un organismo è finito e statico durante tutta la vita della cellula o dell’organismo stesso, il proteoma, al contrario, è altamente dinamico e costantemente variabile in risposta a precisi stimoli esterni e/o durante lo sviluppo. Pertanto, lo studio del proteoma, ovvero del “complemento proteico codificato dal genoma”, deve confrontarsi sul piano tecnico ed intellettuale con una serie di difficoltà generate dalla sua vasta complessità, nonché dalle caratteristiche intrinseche alle macromolecole proteiche (Wasinger et al., 1995). Le proteine, infatti, presentano strutture complesse, proprietà chimico-fisiche eterogenee e diverse possibilità di modificazioni post-traduzionali, per cui nessuna singola metodica di estrazione proteica permette di catturare l’intero proteoma. Un consistente numero di protocolli, che permutano trattamenti fisici, solventi e *buffer*, è presente in letteratura (Rabilloud, 1996). Si stima che, per ogni singola proteina codificata dal genoma, esistono approssimativamente 300 modifiche chimiche che possono influenzarne la funzionalità, la stabilità e le possibilità di interazione (<http://www.abrf.org/index.cfm/dm.home>). Il proteoma presenta, inoltre, differenze di livello nell’espressione all’interno di una cellula: esistono proteine abbondanti, presenti a livelli compresi tra 10^5 - 10^7 molecole per cellula e proteine rare, presenti nell’ordine di 10-100 molecole per cellula ma non esiste nessuna tecnica molecolare

capace di amplificarle, così come succede, ad esempio, per i geni ed i trascritti genici poco abbondanti, amplificabili mediante PCR (*Polymerase Chain Reaction*). Questo *range* molto ampio di espressione limita la visualizzazione e la caratterizzazione esaustiva delle proteine minoritarie negli estratti cellulari totali. Ciò significa che proteine poco abbondanti, come quelle regolatorie e di membrana, risultano fuori dalla portata delle tecniche proteomiche, a meno che non vengano sviluppate strategie specifiche. Una caratteristica fondamentale del proteoma è la dinamicità, la capacità cioè di alterarsi in risposta a specifici stati metabolici e alla ricezione di stimoli sia extra- che intracellulari. Il numero di proteine perciò presenti in una cellula in un dato momento oscilla tra 200.000 e 2.000.000 di unità (Rose et al., 2004). Lo studio del proteoma appare perciò sempre più legato allo sviluppo di tecnologie mirate, capaci di isolare, separare e identificare il più alto numero possibile di componenti.

1.2 La Proteomica

La parola “proteomica” identifica quella branca delle scienze che focalizza la propria attenzione sulle analisi dei proteomi (Blackstock and Weir, 1999). Tale termine è spesso utilizzato per descrivere le metodologie attraverso le quali lo studio delle componenti proteiche di una cellula è perseguito. La proteomica, in realtà, è una vera e propria disciplina scientifica che si propone di colmare il *gap* di conoscenze tra le annotazioni genomiche depositate in banche dati e il comportamento cellulare (Roberts, 2002). I geni, infatti, si perpetuano come sequenze nucleotidiche, ma funzionano tramite l’espressione in proteine. Le molecole proteiche sono responsabili delle caratteristiche strutturali e funzionali della cellula e rappresentano,

dunque, il tramite attraverso il quale il materiale genetico governa la struttura e la funzione della cellula (Figura 1.1). Nella sua accezione più ampia, l'analisi proteomica si propone lo studio dell'identità, della quantità, della struttura, della funzione, nonché dell'interazione, di tutte le proteine di un dato organismo, organo od organello, in relazione alle loro variazioni nel tempo e nello spazio, durante gli stadi fisiologici di sviluppo, o in risposta a precisi stimoli (Figura 1.1). Nello specifico, le indagini proteomiche possono essere distinte in: 1) sistematiche, quando mirano ad identificare tutte le proteine espresse nel campione d'interesse; 2) differenziali, quando analizzano l'espressione differenziale delle proteine nei campioni d'interesse; 3) funzionali, quando comprendono lo studio delle interazioni tra proteine (interattomica), lo studio delle interazioni tra una proteina ed i suoi substrati (metabolomica) e lo studio delle funzioni specifiche delle proteine (genomica enzimatica e biochimica).

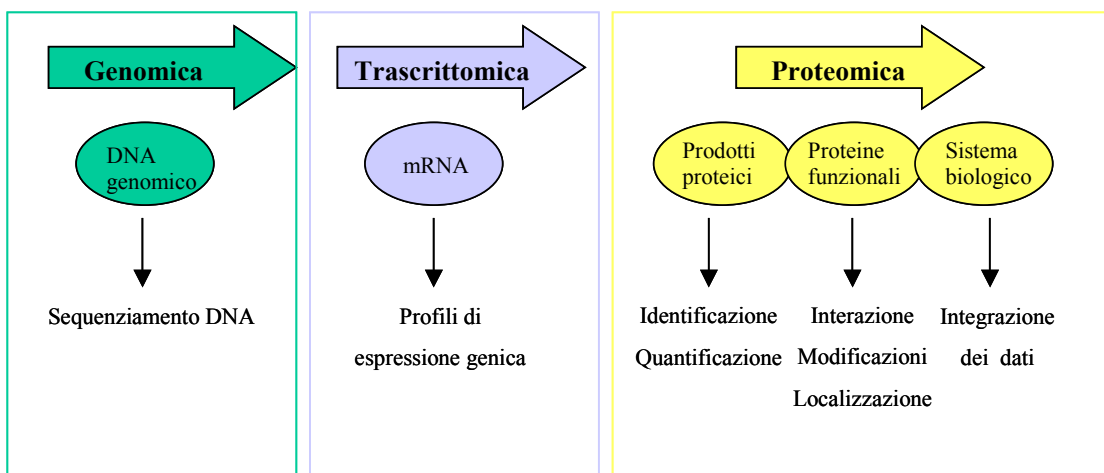


Figura 1.1 Dalla genomica alla proteomica. Integrazione delle informazioni dal genoma al proteoma, per una comprensione dettagliata dei sistemi biologici.

1.2.1 La proteomica sistematica e l'elettroforesi bidimensionale

La proteomica sistematica analizza su larga scala l'espressione delle proteine nel campione oggetto di studio. Essa nasce come logica conseguenza dei diversi programmi di sequenziamento genomico e del crescente sviluppo che ha investito il settore della bioinformatica. Tradizionalmente, la proteomica sistematica utilizza l'elettroforesi bidimensionale (2-DE) come tecnica principale per la separazione delle proteine.

Lo schema generale di un'analisi 2-DE prevede che le molecole proteiche siano inizialmente separate per focalizzazione isoelettrica (IEF). La IEF è un metodo elettroforetico che separa le proteine sulla base del loro punto isoelettrico (pI). Le proteine sono molecole anfotere, portano, cioè, carica netta positiva, negativa o pari a zero, a seconda della loro composizione aminoacidica e del pH dell'ambiente circostante. Il punto isoelettrico è lo specifico valore di pH in corrispondenza del quale la carica netta di una proteina è zero. Sotto l'influenza di un campo elettrico applicato in un gradiente di pH, una proteina si muove nel gradiente fino al punto in cui la carica netta posseduta è pari a zero. Una proteina con carica netta positiva migrerà verso il catodo, mentre una proteina carica negativamente migrerà verso l'anodo, diventando progressivamente meno cariche, nel percorso lungo il gradiente di pH, fino a perdere completamente la mobilità elettroforetica al valore di pH corrispondente al pI. La IEF concentra le proteine in corrispondenza del proprio pI, separando molecole caratterizzate da minime differenze nella carica.

In una analisi 2-DE, il campione proteico viene lasciato migrare lungo un sottile e stretto gel di poliacrilammide (*strip*), che ha un gradiente di pH immobilizzato (IPG) tra un anodo e un catodo. La migrazione di ciascuna proteina termina quando essa

raggiunge il proprio pI nel gradiente. La *strip* viene successivamente applicata ad un secondo gel di poliacrilammide, di dimensioni maggiori, sul quale si effettua una seconda separazione (SDS-PAGE - elettroforesi su gel di poliacrilammide in condizioni denaturanti), ortogonale alla prima, in base alla massa molecolare. La corsa su gel avviene in condizioni denaturanti, le proteine perdono, cioè, il loro naturale ripiegamento e vengono rivestite da molecole di sodio dodecil solfato (SDS), un surfattante anionico che, mascherando le cariche originali presenti sulle molecole proteiche, conferisce loro una carica netta negativa per unità di massa. Le proteine, caricate negativamente, sono lasciate migrare, in una cella di elettroforesi verticale, verso il polo positivo, posto in corrispondenza della parte inferiore del gel. La mobilità elettroforetica, legata alla lunghezza delle catene proteiche, consente alle molecole più piccole di muoversi più velocemente nel gel e perciò di separarsi dalle molecole più grandi. Una volta risolte, le proteine sono generalmente visualizzate mediante processi di colorazione al nitrato d'argento o al Blue Comassie (Rabilloud et al., 1994) ed eventualmente quantificate mediante l'uso di densitometri. Si ottiene così un elettroferogramma bidimensionale dell'estratto proteico totale, in cui ciascuno *spot* rilevato rappresenta teoricamente una singola molecola proteica, caratterizzata da uno specifico valore di massa e di punto isoelettrico (Gorg et al., 2000). Teoricamente, la 2-DE è capace di risolvere fino a 10.000 proteine su un singolo gel e rilevare e quantificare quantità inferiori ad 1 ng per spot. Il processo di separazione descritto richiede che le proteine vengano dapprima visualizzate attraverso specifiche procedure di rivelazione, successivamente gli *spot* corrispondenti vengano trattati attraverso un approccio proteomico *top-down*, che

consiste nel prelevarli dal gel e sottoporli a digestione. I peptidi così generati vengono identificati attraverso metodiche di analisi di spettrometria di massa (MS).

Affinché la proteina d'interesse possa essere individuata nel gel bidimensionale, essa deve risultare necessariamente visibile al processo di colorazione utilizzato. Molecole proteiche, presenti nel campione biologico in concentrazioni inferiori al limite di rivelazione del colorante scelto, non possono essere perciò evidenziate. Proteine estremamente acide o basiche, così come quelle idrofobiche, sono scarsamente risolte dalle separazioni 2DE. Inoltre, l'elettroforesi bidimensionale è un metodo limitato dalla quantità di campione che può essere applicato sulla prima dimensione (da 150 µg a pochi mg). E' difficilmente accoppiabile *on-line* con le moderne tecniche di MS, per cui interventi manuali o un complesso e costoso sistema robotizzato risultano necessari per implementare le analisi spettrometriche. Sebbene i migliori gel 2D riescano a risolvere circa 10.000 componenti, la tecnica resta comunque limitata alla rivelazione delle sole proteine abbondanti e la riproducibilità resta un problema con cui confrontarsi (Gygi et al., 2000).

Ciononostante, l'importanza degli studi sistematici basati sui gel bidimensionali è sottolineata dal crescente numero di banche dati di immagini 2-DE, relative a diversi organismi, dai mammiferi alle piante, pubblicate sul Web (<http://www.expasy.org/ch2d/2d-index.html>). Esse rappresentano un potente mezzo di scambio delle informazioni, poiché permettono la ricerca rapida di proteine o spot di interesse, mappati su 2-DE, attraverso semplici criteri di ricerca, come ad esempio il nome della proteina, numero di accesso, range di pH e peso molecolare (<http://www.expasy.org/world-2dpage>).

1.2.2 La proteomica differenziale e l'analisi DIGE

L'obiettivo principale della proteomica differenziale è lo studio dei cambiamenti nell'espressione del proteoma di un sistema biologico, in risposta a particolari stimoli o semplicemente durante le diverse fasi di sviluppo del sistema stesso. Quest'approccio analitico richiede saggi estremamente sensibili ed accurati per identificare le proteine in miscele complesse e quantificare le loro relative abbondanze. Molti studi hanno utilizzato l'elettroforesi bidimensionale per il profilo comparativo dell'espressione proteica, confrontando il *pattern* proteico prodotto da differenti campioni, analizzati su diverse mappe 2-D (Tonge et al., 2001). L'approccio descritto, oltre ad essere caratterizzato da una lenta processività, risulta poco accurato a causa dalle variazioni degli spot inter-gel e dalla limitata sensibilità dei coloranti utilizzati per la rivelazione delle proteine sulle singole mappe. Il blue Comassie rappresenta sicuramente il colorante più utilizzato, sia per la sua semplicità d'uso che per la sua compatibilità con altre tecniche analitiche, quali ad esempio la spettrometria di massa. Tuttavia, la sua scarsa sensibilità ne limita l'applicazione. La colorazione all'argento, pur rappresentando un metodo sicuramente più sensibile, offre, invece, un ristretto *range* dinamico, compromettendo l'accuratezza della quantificazione relativa.

L'analisi 2-D DIGE (*differential in-gel electrophoresis*) rappresenta una variante notevolmente più accurata della classica 2-DE (Unlu et al, 1997; Tonge et al., 2001). Diversi campioni proteici sono marcati con differenti coloranti fluorescenti, le cianine Cy2, Cy3 e Cy5, in modo da poter essere mescolati insieme, separati e visualizzati contemporaneamente sullo stesso gel 2-D. Operando in questo modo, il fenomeno della variazione degli spot da gel a gel, principale svantaggio della 2-D

classica, viene minimizzato (Tonge et al., 2001). Le diverse cianine possiedono masse molecolari pressoché uguali, perciò la posizione di una determinata proteina, in una mappa elettroforetica bidimensionale, sarà uguale per tutti i derivati fluorescenti. Poiché i fluorocromi utilizzati hanno spettri di fluorescenza distinguibili, differenti immagini per ciascun campione marcato possono essere registrate, mediante l'utilizzo di opportuni scanner (Tonge et al., 2001).

Dall'analisi delle immagini acquisite, è possibile evidenziare, utilizzando specifici software, differenze di espressione a livello dei singoli *spot* proteici. Gli *spot* di interesse vengono prelevati dal gel, digeriti con tripsina e sottoposti ad analisi di spettrometria di massa per l'identificazione proteica. L'analisi DIGE fornisce le basi per la *detection* di proteine il cui livello di espressione è alterato in un campione rispetto ad un controllo. L'abilità di evidenziare bassi livelli di proteina è perciò importante per la proteomica differenziale: la tecnologia DIGE supporta questa necessità, rilevando, grazie all'uso delle cianine, fino a 500 pg di proteina ed offrendo, inoltre, una risposta lineare alla variazione della concentrazione proteica di oltre 5 ordini di grandezza (10^5). Ciò permette a questa metodica di fornire una quantificazione più accurata rispetto al tradizionale *Silver-staining*, la cui risposta alla variazione della concentrazione proteica è di soli due ordini di grandezza.

Un'immagine rappresentativa di un'analisi DIGE è riportata nella Figura 1.2.

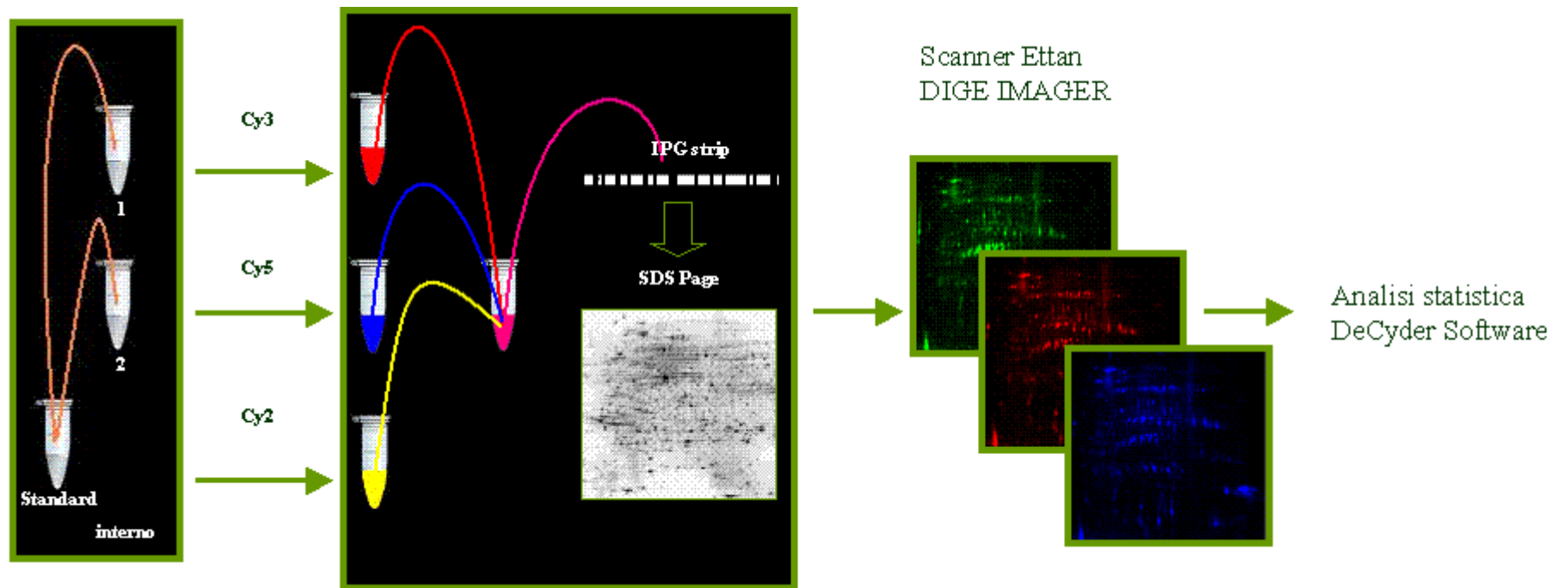


Figura 1.2 Rappresentazione schematica di un esperimento DIGE. Tre differenti campioni, marcati con differenti fluorofori (le cianine Cy3, Cy5 e Cy2), sono mescolati e separati sullo stesso gel bidimensionale. Le relative immagini sono acquisite con un sistema laser a lunghezza d'onda variabile. Le cianine utilizzate possiedono spettri di eccitazione ed emissione differenti; per tanto, mediante l'uso di opportuni software, è possibile quantificare e distinguere le proteine differenzialmente espresse.

Una limitazione fondamentale del metodo DIGE è rappresentata dalla fase di prelievo dello *spot* di interesse dal gel, per le successive analisi di spettrometria di massa. A causa della marcatura incompleta delle lisine (la procedura del *minimal labelling* marca, infatti, il 3% delle lisine totali), le proteine non marcate migrano leggermente più rapidamente rispetto a quelle marcate, utilizzate per la *detection* e il confronto quantitativo. Pur usando robot capaci di rilevare la fluorescenza e tagliare gli spot d'interesse, la quantità della proteina marcata può essere troppo bassa per l'identificazione proteica via spettrometria di massa. In alternativa, si può tentare di correlare il *pattern* elettroforetico del gel 2-D DIGE con il *pattern* di un parallelo gel preparativo, eventualmente colorato al Comassie colloidale. Questo approccio non è tuttavia completamente accurato, perché le differenti colorazioni mostrano variazioni nell'intensità da proteina a proteina; inoltre, il Comassie colloidale, per quanto rappresenti una colorazione estremamente sensibile, capace di rilevare proteine dell'ordine dei nanogrammi, è molto meno sensibile rispetto alle cianine, il cui limite di rivelazione si alza all'ordine dei picogrammi. Una seconda limitazione della tecnica è rappresentata dal fatto che proteine troppo poco abbondanti non vengono rilevate a causa della procedura di marcatura del *minimal labelling*. Tuttavia, il recente sviluppo di *tag* fluorescenti per le cisteine potrebbe minimizzare questo aspetto e rendere la 2-D DIGE ancora più sensibile. Come tutti i metodi proteomici basati sull'utilizzo dell'elettroforesi bidimensionale, essa risulta poco efficiente nel risolvere proteine idrofobiche e proteine con masse molecolari molto grandi o molto piccole. Inoltre, la quantificazione relativa si basa sull'assunzione che in ciascuno *spot* sia presente una sola proteina, per tanto fenomeni di co-migrazione possono compromettere le analisi via elettroforesi 2-D.

1.2.2.1 Strategie alternative alla DIGE per la quantificazione

Le metodiche di marcatura isotopica rappresentano oggi strategie complementari e/o alternative all'analisi differenziale via DIGE. L'utilizzo di isotopi permette la marcatura differenziale di diversi campioni, il loro mescolamento e la misura delle differenze relative nell'espressione proteica mediata dalla spettrometria di massa. Esistono diverse tecniche di marcatura. La marcatura biologica *in vivo* (Oda et al., 1999) consiste nel far crescere le cellule in un terreno arricchito in ^{15}N ed analizzare il campione proteico marcato, rispetto al campione controllo, attraverso SDS-PAGE, seguita da analisi di spettrometria di massa, per le successive quantificazioni. La metodica SILAC (*Stable Isotope Labelling by Amino acid in Cell culture*) rappresenta un'evoluzione dell'approccio descritto. SILAC si basa sull'incorporazione *in vivo* di specifici aminoacidi, come ad esempio ^{15}N -metionina o $^2\text{H}_3$ -leucina (Chen et al., 2002; Ong et al., 2002). Sebbene la marcatura biologica *in vivo* permetta di eliminare una serie di variabili sperimentali legati al trattamento dei campioni, i costi elevati, i tempi lunghi e l'impossibilità di analizzare campioni tissutali hanno determinato la necessità di introdurre marcature chimiche alternative. ICAT (*Isotope-Coded Affinity Tagging*), infatti, prevede la marcatura di un campione per derivatizzazione della catena laterale dei residui di cisteina con un reagente isotopico "pesante" (deuterato), mentre un secondo campione viene marcato con lo stesso reagente, ma "leggero" (non deuterato) (Figura 1.3) (Gygi et al, 1999).

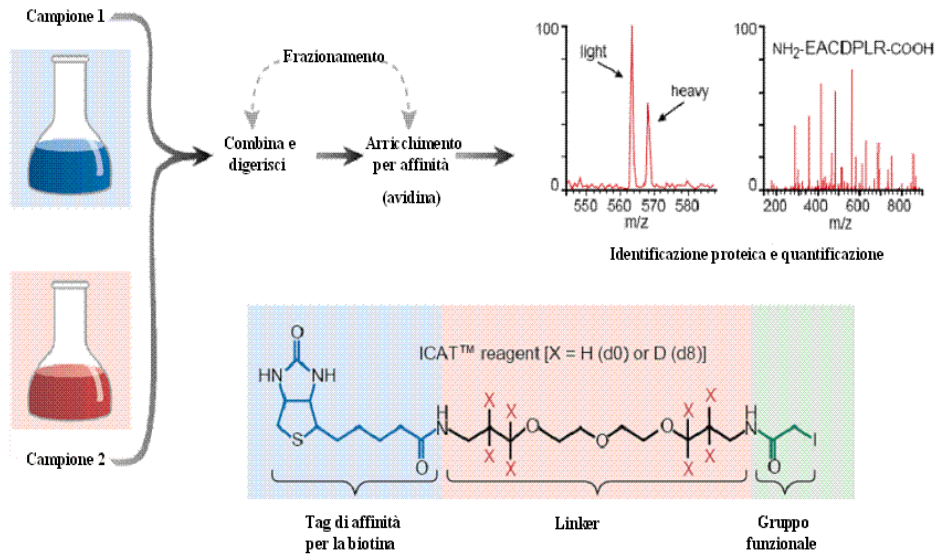


Figura 1.3 Rappresentazione schematica di un esperimento di proteomica quantitativa ICAT. Figura adattata da Patterson and Aebersold (2003).

Dopo la marcatura, i campioni vengono combinati insieme e digeriti con proteasi specifiche. Le peculiari strutture dei reagenti ICAT (Figura 1.3), caratterizzati da un *tag* per la cromatografia d'affinità, permettono il recupero dei peptidi marcati, mediante l'interazione avidina/biotina. Le coppie di peptidi marcati con i reagenti ICAT (Figura 1.3), leggero e pesante, sono chimicamente identiche e possono essere usate come standard interni per un'accurata quantificazione. L'abbondanza relativa dei peptidi, e conseguentemente l'abbondanza delle proteine nei campioni d'origine, è ottenuta come rapporto tra le intensità dei picchi della coppia peptidica diversamente marcata, rilevati dalla MS. Il principale svantaggio di questa tecnica è rappresentato dal fatto che solamente le proteine che contengono cisteina possono essere marcate.

I metodi appena descritti si basano sulla quantificazione mediata dalla MS. Esiste però una strategia analitica, denominata iTRAQ (Applied Biosystems), che permette

una quantificazione basata sulla spettrometria di massa tandem (MS/MS). Il reagente iTRAQ viene utilizzato per marcare campioni a valle della digestione proteolitica. Esistono 4 diversi reagenti iTRAQ, ciascuno contenente una regione *reporter* dal peso differente (da 114 a 117 Da; Figura 1.4). Dopo la reazione di marcatura, i campioni vengono mescolati insieme ed analizzati via MS. I peptidi provenienti dalla stessa proteina, espressa in maniera differenziale nei campioni in analisi, avranno uguale valore di massa, nonostante il differente valore della porzione del *reporter* del reagente iTRAQ (Figura 1.4). Questo è possibile perché il reagente iTRAQ è caratterizzato dalla presenza di una seconda regione, detta *balance* (che va da 31 a 28 Da), che bilancia le differenze di massa della regione *reporter*, generando un *tag* isobarico di 145 Da per tutti i differenti campioni (Figura 1.4). La frammentazione del peptide mediante MS/MS determina anche la rottura del reagente iTRAQ, che libera in forma ionica i diversi *reporter* a valori compresi tra 114 e 117 Da (Figura 1.4). A seconda del numero di campioni analizzati nel corso dell'esperimento, nella regione a masse basse dello spettro possono essere rilevati una coppia, una tripletta o una tetrade di segnali di intensità differente. La stima dei rapporti tra questi segnali *reporter* permette di ricavare l'abbondanza relativa del peptide, quindi della proteina a cui esso appartiene, nei campioni posti a confronto.

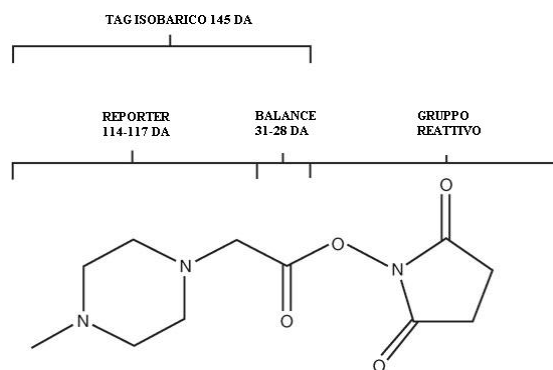


Figura 1.4 Rappresentazione schematica del reagente iTRAQ (Applied Biosystems).

1.2.3. Metodi proteomici non basati sull'elettroforesi bidimensionale

Le limitazioni intrinseche delle metodiche proteomiche tradizionali, basate sui gel di poliacrilammide e sull'elettroforesi bidimensionale, hanno determinato la ricerca di strategie analitiche alternative, da affiancare alle metodologie classiche, al fine di incrementare, non solo la processività dei campioni, ma anche la quantità e la qualità dei dati. In particolare, la cromatografia liquida ad alta risoluzione (HPLC) ha offerto una valida possibilità per il suo alto potere risolutivo, l'elevata riproducibilità e la possibilità di accoppiamento *on-line* con le attuali tecniche di spettrometria di massa. Essa si basa sulla differente distribuzione dei componenti di una miscela tra due fasi immiscibili tra loro: la fase stazionaria, immobilizzata nelle colonne cromatografiche, e la fase mobile, che fluisce all'interno della fase stazionaria. Il frazionamento e l'eluizione dei componenti avviene variando le caratteristiche della fase fluente (Wilson and Walker, 1994). L'ampia disponibilità di scelta per quel che riguarda le tipologie di entrambe le fasi rende la HPLC una tecnica versatile e di fondamentale importanza in molti settori delle scienze biologiche. Tuttavia, la maggior parte delle separazioni ottenute nel campo della proteomica sono state perseguite utilizzando la cromatografia liquida in fase inversa (RP-HPLC), per la sua elevata compatibilità con le moderne tecniche di MS (Simpson, 2003).

Il frazionamento mediante RP-HPLC si basa prevalentemente sulle interazioni tra le catene laterali degli amminoacidi costituenti il composto proteico in analisi e la superficie idrofobica della fase stazionaria. La fase mobile in RP-HPLC contiene una miscela di acqua e di un solvente organico con essa miscibile, in genere acetonitrile (ACN). L'ACN è un modificatore organico che mostra elevata affinità per composti idrofobici. Un componente acido (acido formico o acetico) è aggiunto alla fase

mobile per denaturare e caricare positivamente le componenti proteiche, nonché ridurre al minimo le interazioni ioniche non desiderate con la fase stazionaria (Simpson, 2003). Lo sviluppo della colonna si ottiene variando nel tempo la concentrazione del solvente organico eluotropo (eluizione in gradiente). Il frazionamento risulta basato sulla idrofobicità relativa delle proteine e dei peptidi, per cui vengono eluite per prima le componenti più idrofiliche. Le componenti idrofobiche vengono eluite successivamente al variare del gradiente di concentrazione del modificatore organico. Piccole variazioni nella composizione della fase mobile influenzano grandemente la separazione (Wilson and Walker, 1994). I supporti cromatografici utilizzati in fase inversa sono rappresentati da composti silicati (in genere gel di silice) derivatizzati con gruppi alchilici, in particolare con butile (C₄), ottile (C₈), o decaottile (C₁₈). Tali resine sono chiamate rispettivamente C₄, C₈, C₁₈. Le catene più corte, C₄ e C₈, sono preferite per campioni più idrofobici (proteine intere), mentre quelle più lunghe, tipo C₁₈, sono adatte per campioni nettamente più idrofilici (peptidi) (Simpson, 2003).

La RP-HPLC può essere usata anche per eliminare i sali da campioni proteici e peptidici provenienti da differenti procedure estrattive, da reazioni chimiche, quali riduzione e alchilazione in presenza di urea o guanidinio idrocloruro, o *cleavage* con bromuro di cianogeno.

La RP-HPLC è di fatto utilizzata per la separazione di peptidi da miscele complesse prima della spettrometria di massa e massa tandem (MS/MS). Il digerito proteico viene caricato nel sistema HPLC, separato e l'eluato analizzato *on-line* attraverso MS e MS/MS. Miscele complesse contenenti un elevato numero di componenti proteiche possono essere analizzate direttamente anche quando i livelli di concentrazione delle

differenti proteine variano di parecchi ordini di grandezza (Simpson, 2003). Sebbene siano stati esplorati numerosi metodi per accoppiare cromatografia liquida a fase inversa con la MS (Vestal, 1990), l'accostamento RP-HPLC e ionizzazione elettrospray (ESI) ha consentito di introdurre la cromatografia liquida accoppiata alla spettrometria di massa tandem (LC-MS/MS) nei laboratori come tecnica di routine. Tale approccio permette di ottenere, oltre ad informazioni sulla massa molecolare, anche informazioni di tipo strutturale.

L'uso di colonne capillari in cromatografia liquida ha notevolmente aumentato la sensibilità delle analisi MS/MS. Le microcolonne usate in RP-HPLC sono distinte sulla base del loro diametro interno (I.D.) e sono perciò classificate in: microbore (I.D. inferiore a 2.1mm), narrowbore (I.D. inferiore a 1.0mm) e capillari (I.D. inferiore a 0.5mm). L'introduzione delle colonne narrowbore e capillari ha ridotto le velocità di flusso a valori compresi tra i 50 e i 400 nl/min (nano-LC). La maggior parte delle nano-LC rileva quantità minime comprese tra 1 e 10 femtomoli (Gatlin et al., 1998). Poiché le colonne capillari possono essere usate anche come sorgenti di ionizzazione, ciò determina un aumento della sensibilità in seguito alla riduzione, nell'unità di tempo, del volume di solvente che può essere "nebulizzato" nello spettrometro di massa ESI. Martin e collaboratori (2000) hanno infatti rilevato quantità comprese tra 5 e 10 attomoli, in seguito all'introduzione dei sistemi di LC capillari completamente automatizzati (flussi variabili tra 5-200 nl/min).

Le separazioni multidimensionali rappresentano un sistema che ha notevolmente incrementato il potere delle risoluzioni cromatografiche. Esse consistono nell'accoppiare in LC due o più metodi di frazionamento tra loro indipendenti (o ortogonali). Ciò significa che le proprietà chimiche e fisiche, tramite le quali l'analita

è frazionato, devono essere differenti per ciascuna separazione (o dimensione) (Evans and Jorgenson, 2004). Sulla base di tale approccio, la RP-HPLC, che discrimina i componenti di una miscela peptidica sulla base dell'idrofobicità, è di fatto accoppiata, sia off-line che on-line, con la cromatografia liquida a scambio ionico (IEC).

La IEC si basa su interazioni di tipo elettrostatico che si verificano tra molecole cariche di segno opposto. Le proteine e i peptidi possiedono gruppi chimici ionizzabili e la presenza di carica può essere sfruttata per separare miscele di tali composti. La carica netta che mostra una molecola dipende dal suo pKa e dal pH della soluzione, secondo quanto descritto dall'equazione di Henderson-Hasselbach (Wilson and Walker, 1994). Le separazioni IEC sono condotte in colonne impaccate con resine scambiatrici di ioni. Esistono due tipi di resine: a scambio anionico e a scambio cationico, differenziate a seconda della carica che esse portano. Per le analisi delle miscele peptidiche sono generalmente utilizzate resine a scambio cationico forte (SCX). Esse sono rappresentate da matrici di cellulosa, destrano o polistirene opportunamente funzionalizzate con gruppi solfo (-SO₃-), solfometile (-CH₂SO₃-) e solfopropile (-CH₂CH₂CH₂SO₃-) (Wilson and Walker, 1994). Tanto maggiore è la carica della molecola da scambiare, tanto più forte risulterà il legame con la resina. I peptidi in miscela vengono normalmente intrappolati nelle colonne SCX a bassi valori di pH (@ 2), che conferiscono loro carica netta positiva. Una sola carica è in genere sufficiente per generare il legame con le resine a scambio cationico forte. Il frazionamento selettivo si ottiene variando nel tempo il pH e/o la forza ionica della fase mobile elettrolitica (Issaq et al., 2002).

Il potere di risoluzione di un sistema cromatografico 2D è dato dalla combinazione di due o più colonne. Teoricamente, la capacità di picco di un sistema 2D-LC è uguale al prodotto dei fattori di capacità relativi a ciascun sistema cromatografico utilizzato (Mitulovic et al., 2004). Essa misura la quantità di materiale che può essere risolto senza che si abbia sovrapposizione di picchi, indipendentemente dall'operazione di eluizione in gradiente. E' definita come il numero massimo di picchi che possono essere separati, l'uno accanto all'altro, con una risoluzione pari ad 1 unità di massa (Wilson and Walker, 1994). Ipoteticamente, se due tecniche hanno una capacità di picco ciascuna pari a 100, la combinazione di entrambe in un sistema 2D fornisce una capacità pari a 10.000. In una separazione 2D ideale, la distribuzione dei componenti sulla prima dimensione non dovrebbe correlarsi in alcun modo con la distribuzione nella seconda e nessuna risoluzione ottenuta nella prima separazione dovrebbe essere persa nella successiva (Evans and Jorgenson, 2004). Il metodo più semplice per realizzare un frazionamento 2D è accoppiare in serie due colonne con meccanismi ortogonali di separazione, come mostrato nella Figura 1.5

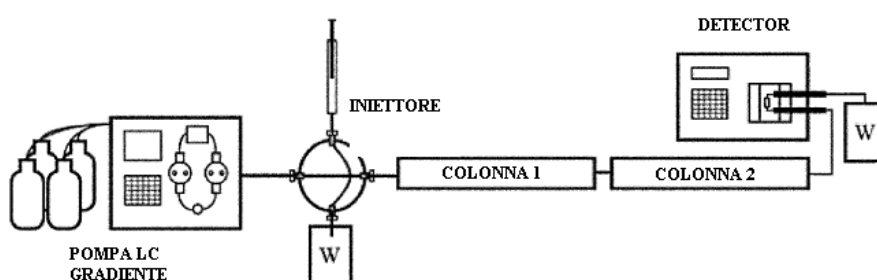


Figura 1.5 Rappresentazione schematica di un sistema cromatografico bidimensionale. La colonna 1 e 2 rappresentano due metodi di separazione, tra loro ortogonali.

Le frazioni fluiscono dalla prima colonna in seguito ad un processo di eluizione discontinuo (step di gradiente); a tal fine si usa un solvente dal potere eluotropo

crescente, specifico per il materiale d'impaccamento della colonna separatrice. Quando la prima frazione è eluita dalla colonna 1, il primo step di gradiente viene bloccato e la fase mobile appropriata alla colonna 2 viene applicata su di essa. I componenti della prima frazione vengono così separati nella seconda dimensione. Successivamente, un secondo step di gradiente è applicato alla prima colonna e il processo viene così ripetuto (Evans and Jorgenson, 2004). Il più promettente approccio che segue lo schema della LC multidimensionale è la tecnologia di identificazione proteica MudPIT (multidimensional protein identification technologies), sviluppata da Yates e collaboratori (Washburn et al., 2003; Wu et al., 2003). Tale metodica, che permette di catalogare centinaia di migliaia di proteine in un singolo esperimento, rientra nelle metodologie analitiche della “proteomica *shotgun*”.

Gli esperimenti di proteomica *shotgun* prevedono la digestione enzimatica di miscele proteiche complesse in peptidi, seguita da analisi di spettrometria di massa tandem (paragrafo 1.3). Algoritmi computerizzati permettono, successivamente, di “mappare” i peptidi così identificati sulle rispettive proteine, determinando le diverse proteine contenute nella miscela d'origine. Gli approcci di proteomica *shotgun* possono essere *gel-free*, se basati sulla cromatografia liquida mono- o bidimensionale, oppure *gel-based* se la popolazione proteica d'interesse è prefrazionata, utilizzando gel di poliacrilammide, generalmente SDS-PAGE (Lee et al., 2007).

1.3 La spettrometria di massa in proteomica

Tutte le metodologie descritte nei paragrafi precedenti utilizzano la spettrometria di massa (MS) come piattaforma tecnologica fondamentale per l'identificazione di macromolecole proteiche. La MS ha, di fatto, sostituito le procedure classiche di caratterizzazione delle proteine via degradazione di Edman, configurandosi come tecnica di elezione per l'indagine proteomica, poiché offre una tecnologia molto più sensibile, adatta allo studio di miscele complesse e ad analisi *high-throughput* (Yates, 1998).

La spettrometria di massa convenzionale è un approccio strumentale che permette di misurare la massa degli ioni che derivano dalle biomolecole d'interesse. La misura della massa sarà più o meno accurata a seconda della strumentazione utilizzata. La MS convenzionale fornisce, però, poche informazioni di tipo strutturale. Tali informazioni possono essere ottenute, invece, grazie alla spettrometria di massa tandem (MS/MS). Tale approccio permette di ricavare informazioni sulla sequenza peptidica. In pratica, lo spettrometro di massa può essere programmato per eseguire un *full-scan* delle masse peptidiche eluite in un preciso istante dalla colonna, e frammentare gli ioni precursori d'interesse (ioni parenti), rilevati nella relativa scansione di massa. La frammentazione degli ioni parenti può avvenire mediante l'utilizzo di diverse tecniche: CID (*Collision-induced dissociation*), ECD (*Electron capture dissociation*), ETD (*Electron transfer dissociation*) e IRMPD (*infrared multiphoton dissociation*) (Chen, 2008). Le tecniche elencate sono state sviluppate in associazione, più o meno specifica, con particolari tipi di analizzatori di massa, ciascuno dei quali è perciò capace di generare peculiari *pattern* di frammentazione

(Chen, 2008). Poiché le frammentazioni avvengono in maniera caratteristica e prevedibile, dai dati registrati è possibile risalire ad informazioni di sequenza.

Il metodo CID rappresenta, ad oggi, la tecnica più utilizzata per ottenere la frammentazione di proteine e peptidi. Esso si basa sull'urto anelastico tra gli ioni biomolecolari e molecole o atomi di un gas inerte. Durante la collisione, gli ioni d'interesse subiscono dapprima un aumento della loro energia cinetica, parte della quale viene successivamente convertita in energia vibrazionale, rapidamente distribuita lungo tutti i legami covalenti. La frammentazione avviene quando l'energia interna degli ioni supera la barriera di attivazione per la rottura dei legami. Il *pattern* di frammentazione, che dipende da alcuni fattori, come l'energia di collisione, la sequenza del peptide e il numero di cariche, contiene tutte le informazioni necessarie per determinare la sequenza aminoacidica. Appare, perciò, di fondamentale importanza avere un approccio sistematico per lo studio del *pattern* di frammentazione. La conoscenza di base della nomenclatura dei frammenti è essenziale per l'analisi proteomica. I frammenti possono essere divisi in due categorie: 1) frammenti che derivano dalla rottura dei legami chimici della catena peptidica; 2) frammenti che derivano dalla rottura delle catene laterali degli aminoacidi. Gli ioni prodotti in seguito a *cleavage* dei legami C α -C, C-N e/o N-C α presenti lungo la catena peptidica, sono chiamati rispettivamente ioni della serie a_n, b_n, c_n se la carica è ritenuta all'N-terminale, o ioni della x_n, y_n, z_n se la carica è ritenuta al C-terminale (*n* indica il numero di aminoacidi nel peptide; (Yates, 1998; Steen and Mann, 2004; Chen, 2008).

Lo schema rappresentativo della frammentazione e della nomenclatura di un peptide è illustrato nella Figura 1.6.

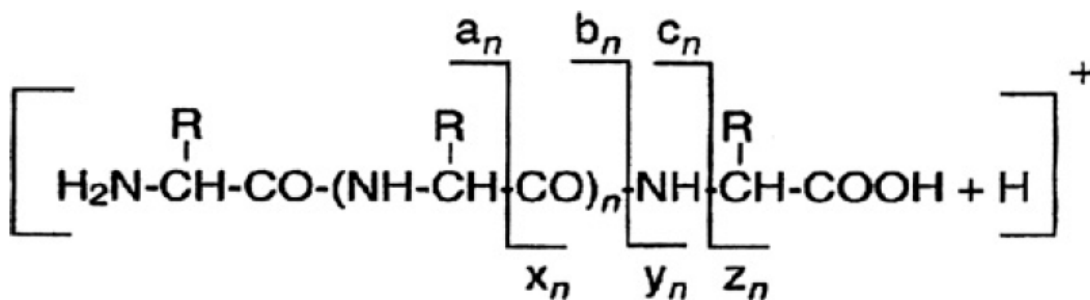


Figura 1.6 Schema di frammentazione di un peptide e relativa nomenclatura.

In linea di principio, la differenza di massa tra ioni consecutivi della stessa serie può essere usata per determinare l'identità degli aminoacidi, eccetto per leucina e isoleucina, perché isomeri. In realtà, anche la differenza di massa tra glutamina e lisina non può essere apprezzata perché molto piccola, a meno che non vengano usati strumenti ad altissima risoluzione. Esiste una differenza significativa tra il *pattern* di frammentazione prodotta da CID a bassa o ad alta energia di collisione (Chen, 2008). Per le basse energie di collisione, in genere inferiori a 200 eV, i frammenti generati sono soprattutto della serie b_n e y_n . Ad alte energie di collisione (dell'ordine di parecchi keV) avviene, invece, la rottura delle catene laterali, per cui è possibile distinguere anche gli isomeri leucina e isoleucina. Quando la frammentazione avviene tra gli atomi di carbonio β e γ , questi frammenti sono chiamati frammenti della serie d_n se la carica è ritenuta all'N-terminale, oppure frammenti della serie w_n se la carica è ritenuta invece al C-terminale (Chen, 2008). Nella maggior parte dei casi, gli spettri di massa tandem ottenuti mediante CID mostrano parecchie serie complete di ioni. Ciò richiedeva una complessa analisi manuale, oggi completamente automatizzata grazie all'introduzione di specifici software (Yates, 1998; Chen, 2008).

1.4 Gli spettrometri di massa

Gli spettrometri di massa sono strumenti capaci di formare, separare, rivelare gli ioni molecolari identificando il loro rapporto “massa-carica” (m/z). Essi derivano dall’assemblaggio di un set modulare di componenti essenziali:

- Sorgente ionica. Si tratta di un sistema capace di convertire e trasferire le molecole di analita in ioni in fase gassosa. La ionizzazione dei peptidi e delle proteine avviene per aggiunta di uno o più protoni alla biomolecola. Un peptide, che possiede una massa molecolare di 2000 dalton (M), avrà un rapporto m/z pari a 2001 dopo ionizzazione per addizione di un protone ($[M+H]1+$), e un valore di m/z pari a 1001 per addizione di due protoni ($[M+2H]2+$). Le sorgenti ioniche, oggi maggiormente in uso per le indagini proteomiche, sono rappresentate dalle sorgenti MALDI (matrix-assisted laser desorption/ionization) e da quelle ESI (elettrospray).
- Analizzatore di massa. Esso è dispositivo usato per separare gli ioni in fase gassosa. Ioni di un particolare valore di m/z sono separati l’uno dall’altro utilizzando proprietà fisiche, quali campi elettrici e magnetici. Tipicamente, le proprietà di campo vengono modificate in funzione del tempo. I principali tipi di analizzatori sono rappresentati dai quadrupoli, le trappole ioniche, gli analizzatori TOF (time of flight) e FT-ICR (Fourier transform-ion cyclotron resonance).
- Rivelatore. Gli ioni di particolari valori m/z , provenienti da un particolare tipo d’analizzatore di massa, vanno a colpire il sistema di rivelazione. La corrente ionica prodotta al *detector* (rivelatore) in funzione del tempo è usata

per determinare l'intensità (abbondanza) e il valore m/z degli ioni che vi arrivano. In generale, l'abbondanza relativa degli ioni generati nella sorgente è misurata nel rivelatore, il risultato è visualizzato come grafico (spettro di massa), che pone i valori di m/z sulle ascisse e i valori d'abbondanza ionica sulle ordinate. Gli spettri di massa sono generalmente normalizzati rispetto al picco più abbondante (picco base). In uno spettro normalizzato, il valore del picco base è regolato a 100 unità, le altezze relative di tutti gli altri picchi dello spettro sono, perciò, comprese tra 0 e 100 unità.

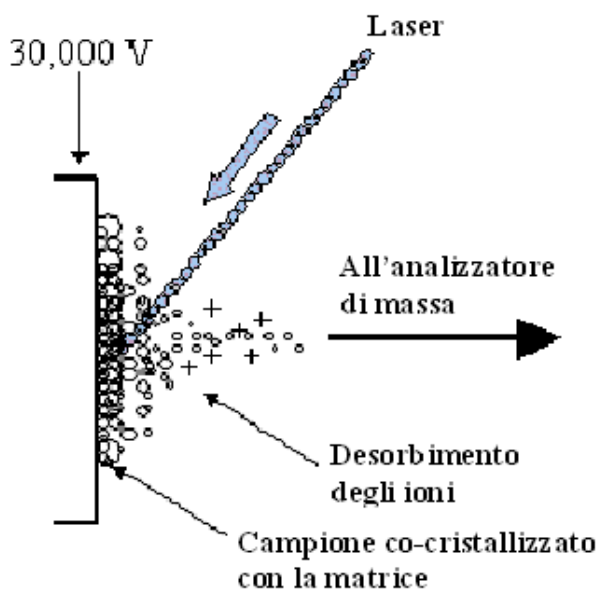
L'abilità di creare ioni in fase gassosa da macromolecole polari, zwitterioniche e non volatili, quali proteine e peptidi, senza originare fenomeni degradativi e di frammentazione, è di fondamentale importanza per tutte le analisi MS. Le tecniche di ionizzazione rappresentano, dunque, un aspetto cruciale connesso all'uso della spettrometria di massa negli studi proteomici. Tradizionalmente, il trasferimento di piccole molecole organiche in fase gassosa è stato perseguito attraverso vaporizzazione termica, seguita da impatto elettronico o ionizzazione chimica (Harrison and Cotter, 1990; Harrison, 1992). Il maggior ostacolo nell'applicare questi approcci all'analisi di proteine è stato la difficoltà di vaporizzare tali molecole senza decomposizione termica estensiva (Daves, 1979). Parecchi studi, compresi tra il 1950 e il 1970, dimostrarono che la polarità dei polipeptidi poteva essere ridotta mediante acilazione, esterificazione, riduzione o N- e O-metilazione (Morris et al., 1973). Grazie a queste reazioni, piccoli peptidi furono portati in fase gassosa per le analisi MS. Negli anni ottanta, il bombardamento con atomi veloci (FAB, *fast atom bombardment*) fornì un ulteriore passo avanti nello sviluppo delle procedure di MS

applicate allo studio delle componenti proteiche di un sistema biologico. Ciononostante, solo con l'introduzione dei cosiddetti metodi di ionizzazione *soft*, MALDI ed ESI, la spettrometria di massa è diventata una tecnica di routine per l'analisi di peptidi e proteine.

1.4.1 Sorgenti ioniche

1.4.1.1 Ionizzazione laser assistita da matrice (MALDI)

La ionizzazione MALDI è eseguita lasciando co-cristallizzare le molecole di analita, su un'apposita piastra di metallo, con un eccesso di matrice, sostanza capace di assorbire a specifiche lunghezze d'onda (Figura 1.7; Beavis and Chait, 1996).



Le matrici tipicamente usate per le analisi di peptidi e proteine includono, rispettivamente, l'acido α -ciano 4-idrossi-cinnammico e l'acido sinapinico. Gli ioni sono prodotti bombardando il campione con impulsi, di breve durata (1-10 ns), di luce UV (generalmente, UV337nm), generata da un laser ad azoto (Figura 1.7).

Figura 1.7 Rappresentazione schematica del processo MALDI.

L'interazione del laser pulsato con il campione fornisce la ionizzazione (per protonazione) sia della matrice sia dell'analita, attraverso un meccanismo di trasferimento d'energia dalla matrice all'analita stesso, piuttosto che attraverso una

ionizzazione diretta mediata dal laser (Simpson, 2003). MALDI produce, fondamentalmente, ioni a carica singola. Una spiegazione di come ciò avvenga è stata proposta da Karas e collaboratori (2000). Applicando un alto potenziale elettrostatico (circa 30 kV) tra la piastra MALDI, su cui sono depositati i cristalli di matrice e analita, e il foro d'ingresso (inlet) dell'analizzatore, gli ioni generati vengono accelerati verso l'analizzatore stesso.

1.4.1.2 Ionizzazione per elettro-nebulizzazione (ESI)

La ionizzazione ESI (*elettrospray*) consente di creare ioni in fase gassosa, applicando una tensione ionizzante di parecchi kilovolt ad un ago nebulizzatore collegato all'uscita di un sistema per l'introduzione del campione (Banks and Whitehouse, 1996). Esso può essere rappresentato da una siringa, per l'infusione diretta, o da un cromatografo HPLC. Creando una differenza di potenziale tra il capillare, attraverso cui fluisce il campione, e l'inlet dello spettrometro di massa, uno spray di goccioline cariche si forma a partire dall'analita in soluzione (Figura 1.8). Le gocce cariche sono desolvate per applicazione di un controflusso di gas e/o per riscaldamento (Mann et al., 2001). Quando la densità di carica superficiale supera il limite di stabilità di Rayleigh, cioè quando la repulsione elettrostatica diventa più grande della tensione superficiale della goccia, la goccia esplode ed espelle gli ioni nudi (ioni positivi). Essi sono convogliati nell'analizzatore di massa grazie ad un sistema di lenti (Figura 1.8).

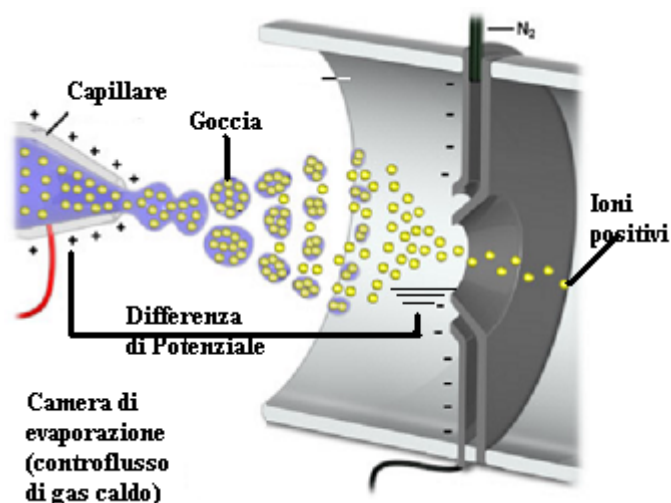


Figura 1.8 Rappresentazione schematica del processo ESI.

Regolando in maniera opportuna (*tuning*) i parametri elettrici per l'*elettrospray*, il desorbimento degli ioni avviene in maniera più o meno rapida. ESI può convertire prontamente molecole in soluzione in ioni in fase gassosa. Gli ioni prodotti, però, sono multicarica. La presenza di cariche multiple dipende da fattori quali la composizione ed il pH del solvente usato per la ionizzazione, ma anche dalla natura chimica dell'analita.

Per peptidi (massa inferiore a 2000 dalton), ESI genera ioni a singola o doppia carica; per proteine (massa superiore a 2000 dalton), il processo di ionizzazione genera specie a carica multipla (*charge envelope*). Il fenomeno delle cariche multiple ha l'effetto di abbassare i valori di m/z dell'analita ad un *range* che può essere facilmente misurato da differenti tipi di analizzatori di massa, quali analizzatori quadrupolari e a trappola ionica (Simpson, 2003).

1.4.2 Gli analizzatori di massa

1.4.2.1 Analizzatori a tempo di volo

Il principio su cui si basa un analizzatore a tempo di volo (TOF, Time of Flight) è relativamente semplice: esso misura il tempo che gli ioni impiegano per percorrere una regione priva di campi elettrici (detta, appunto, tubo di volo).

Nella sorgente ionica, tutti gli ioni sono accelerati allo stesso valore d'energia cinetica (E_k), attraverso l'applicazione di un campo elettrico (E); poiché E_k è funzione della massa, gli ioni più piccoli sono più veloci di quelli più grandi e percorrono in tempi minori il tubo di volo, raggiungendo prima il *detector* (Steen and Mann, 2004).

Nello specifico,

$$E_k = zeE$$

(dove z è il numero delle cariche ed e la carica di un elettrone)

$$E_k = 1/2mv^2$$

Le velocità degli ioni (v) sono dipendenti dalla loro massa (m),

$$v = (2zeE/m)^{1/2}.$$

Il tempo (t) necessario per attraversare la regione è registrato e convertito in uno spettro di massa:

$$D = vt$$

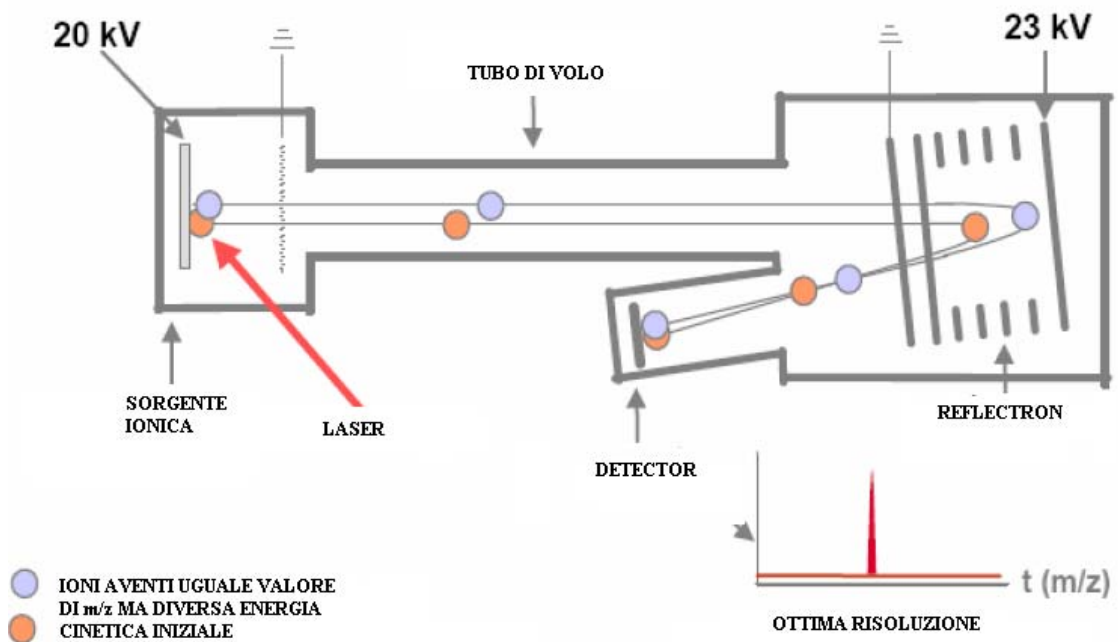
(D è la distanza coperta dallo ione)

$$t = (m/2zeE)^{1/2}D,$$

cioè

$$\mathbf{m/z = 2eE(t/D)^2}$$

Il distanza di volo per uno ione può essere aumentata senza aumentare la grandezza del tubo di volo, incorporando uno specchio o *reflectron* (Figura 1.9). Esso è costituito da una serie di griglie a voltaggi crescenti, posizionate alla fine della regione priva di campi. Gli ioni arrivano al *reflectron* e tornano indietro, curvando il loro percorso in modo tale da non scontrarsi con gli ioni che continuano ad arrivare al *reflectron* stesso. Il vantaggio principale di questo sistema consiste nell'incrementare notevolmente la risoluzione, poiché esso corregge le piccole variazioni di energia cinetica tra gli ioni dello stesso valore m/z , minimizzando le



variazioni nei tempi di volo.

Figura 1.9 Rappresentazione schematica di un analizzatore di massa TOF, equipaggiato di sistema reflectron.

Tradizionalmente l'analizzatore di massa TOF si interfaccia con una sorgente MALDI. Tuttavia, l'introduzione di TOF ad accelerazione ortogonale (oaTOF) ha permesso di interfacciare questo tipo di analizzatore con qualsiasi sorgente ionica,

purché sistemata in maniera perpendicolare al tubo di volo. Gli ioni vengono collimati tramite una lente elettrostatica e quindi trasferiti in una regione di immagazzinamento, da cui vengono successivamente estratti ed accelerati. Ciò influenza notevolmente la risoluzione e l'accuratezza di massa. Ulteriori sviluppi della tecnologia degli analizzatori TOF hanno portato alla creazione d'analizzatori TOF/TOF, che permettono di frammentare il peptide d'interesse, rilevarne i frammenti ed ottenere informazioni di sequenza. In questo tipo di strumenti, due tubi di volo sono posti in serie ed intervallati da una cella di collisione. Il primo TOF permette di selezionare uno ione precursore, che viene frammentato nella cella di collisione, grazie alla collisione con un gas inerte (CID, collision induced dissociation). Il secondo TOF esegue, invece, una scansione dei frammenti generati dalla rottura del precursore, generando spettri MS/MS.

1.4.2.2 Quadrupolo e Trappole ioniche

Un quadrupolo lineare è costituito da quattro barre metalliche, tra le quali è applicata una differenza di potenziale (il "quadrupolo"; Figura 1.10). Gli ioni entrano da una estremità e "viaggiano" con velocità costante in direzione parallela ai poli (direzione z), ma acquistano delle oscillazioni complesse lungo le direzioni x e y a causa dell'applicazione contemporanea ai poli di una tensione (U) a corrente continua (dc) e di una tensione (V) a radiofrequenza (rf) (Simpson, 2003).

Esiste una oscillazione stabile che consente ad un particolare ione di attraversare da un'estremità all'altra il quadrupolo, senza entrare in collisione con i poli; questa oscillazione dipende dal rapporto m/z dello ione. Perciò, soltanto ioni aventi un preciso rapporto m/z attraverseranno l'intera lunghezza del filtro ad una data serie di

condizioni. Tutti gli altri ioni acquisteranno delle oscillazioni instabili e collideranno con i poli, perdendosi. La scansione delle masse viene condotta variando sia rf che dc, mantenendo, però, costante il loro rapporto.

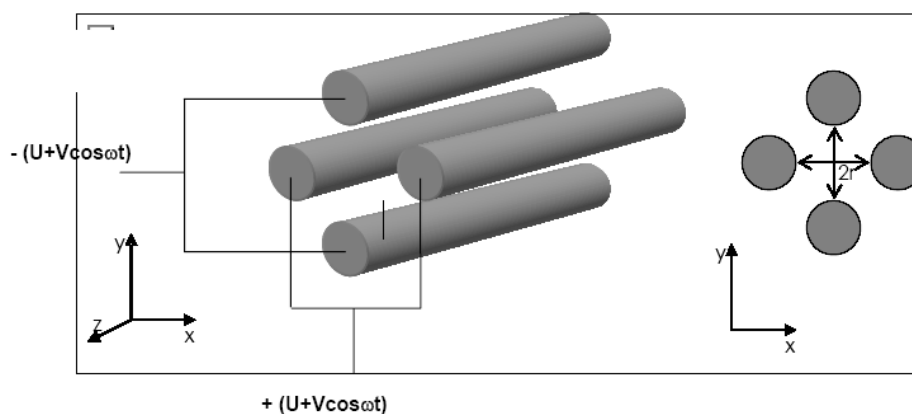


Figura 1.10 Rappresentazione schematica del quadrupolo.

Una trappola ionica consiste in una disposizione sferica del filtro di massa a quadrupolo lineare (Silverstein and Webster, 1999; Figura 1.11). Schematicamente, essa è dotata di un elettrodo ad anello (*hyperbolic ring electrode*) e di due elettrodi a calotta (*hyperbolic end-cap electrode*). Questi ultimi presentano piccoli fori, che permettono il passaggio degli ioni attraverso la trappola. Un alto potenziale di radiofrequenza (RF) è applicato all'elettrodo ad anello, mentre gli *end-cap* sono mantenuti a bassi valori. La differenza di potenziale che si stabilisce tra i due tipi di elettrodi forma sostanzialmente un campo quadrupolare. Gli ioni iniettati all'interno della trappola sopra un certo valore di m/z (determinato dall'intensità della radiofrequenza) hanno traiettorie stabili. La collisione con gas inerte (elio, presente nella IT a pressioni circa pari a 10^{-3} Torr) smorza la loro energia cinetica, in modo da intrappolarli, per azione della RF applicata, al centro dello strumento. Al fine di

ottenere la scansione delle masse, un voltaggio supplementare è applicato agli elettrodi a calotta, mentre la RF, applicata all'elettrodo ad anello, è progressivamente incrementata. Man mano che la frequenza applicata aumenta, gli ioni di valori massa-carica crescenti entrano in risonanza con essa, perciò diventano instabili e la loro velocità e traiettoria aumentano in direzione assiale finché non sono espulsi dalla trappola verso il detector (Simpson, 2003).

Un analizzatore così formato permette di effettuare anche analisi di spettrometria di massa tandem. Si tratta di una serie di eventi che portano alla selezione di uno ione precursore (parentale), seguita dall'analisi degli ioni "figlio", derivanti per frammentazione dallo stesso ione parentale. La frammentazione è in genere indotta da collisioni con gas inerte (CID, *collision-induced dissociation*).

Per l'analisi MS/MS in trappola ionica, gli ioni precursori, di un preciso rapporto m/z , possono essere isolati aumentando l'intensità della RF applicata all'elettrodo ad anello, fino al punto in cui tutti gli ioni sotto la massa selezionata sono espulsi. Durante questo processo, anche gli ioni con un rapporto m/z superiore allo ione precursore sono espulsi fuori dalla trappola, applicando una RF addizionale agli *end-cap*, la cui frequenza ed intensità viene settata appena sopra il valore specifico per ritenere la massa dello ione precursore. Man mano che la RF al ring electrode viene aumentata, gli ioni con un rapporto massa-carica maggiore rispetto allo ione d'interesse sono portati, in successione, in risonanza con il voltaggio supplementare, applicato agli elettrodi a calotta. Essi vengono via via espulsi, lasciando al centro della trappola lo ione parentale. Esso viene sottoposto a CID; la popolazione di frammenti risultanti è progressivamente analizzata, effettuando la scansione dei

rapporti m/z , come descritto in precedenza, per acquisire uno spettro, detto di frammentazione (Simpson, 2003).

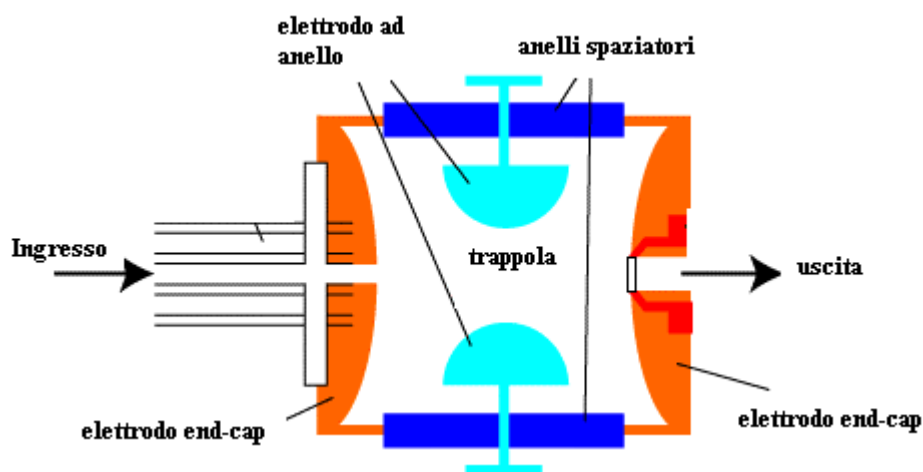


Figura 1.11 Rappresentazione schematica di un analizzatore di massa a trappola ionica

Gli analizzatori a trappola ionica sono compatti, robusti e relativamente economici. Tuttavia, essi sono strumenti a risoluzione unitaria che presentano, in pratica, bassa accuratezza nella determinazione della massa. Ciò è in parte dovuto al numero limitato di ioni che possono essere accumulati senza che si generino distorsioni “spazio-carica”.

1.4.2.3 FT-ICR

Gli strumenti FT-ICR (*Fourier Transform-ion Cyclotron Resonance*) utilizzano un campo magnetico statico, prodotto da un magnete superconduttore, per intrappolare gli ioni all'interno di una cella cubica. Gli ioni saranno forzati a mantenere un moto circolare con una caratteristica frequenza orbitale, dipendente dallo specifico rapporto m/z (Silverstein and Webster, 1999).

Grazie all'applicazione di un impulso di radiofrequenza, tra due piastre della cella che fungono da trasmettitore (piastre di eccitazione), gli ioni intrappolati producono un moto orbitale coerente. Gli ioni in moto orbitale generano un debole segnale, che viene raccolto dalle piastre della cella che agiscono come ricevitore (piastre di ricezione). La frequenza del segnale generato da un particolare ione è uguale alla sua frequenza orbitale, a sua volta inversamente correlata al rapporto m/z . L'intensità del segnale per una data frequenza è proporzionale al numero di ioni aventi un dato valore m/z . Il segnale totale viene amplificato e scomposto nelle varie componenti di frequenza, generando così lo spettro di massa (Simpson, 2003). In condizioni di vuoto spinto, il moto orbitale può essere mantenuto per tempi relativamente lunghi, per permettere di misurare la frequenza numerose volte, aumentando così l'accuratezza. Gli FT-ICR sono in grado di produrre spettri di massa con un'elevatissima risoluzione.

Gli esperimenti di MS/MS possono essere ottenuti in un analizzatore FT-ICR, lasciando collidere gli ioni con gas inerte (CID, oppure mediante esposizione a fotoni (IRMPD, infra-red multiphoton dissociation), o ad elettroni (ECD, electron capture dissociation).

1.4.2.4 Analizzatori ibridi

Negli ultimi anni, lo sviluppo tecnologico, unito al crescente interesse nei confronti della proteomica, ha determinato lo sviluppo di macchine ibride, caratterizzate da elevatissima risoluzione e accuratezza di massa. L'assemblaggio quadrupolo-TOF, meglio conosciuto come Q-TOF, è caratterizzato dall'accoppiamento tra un filtro di massa a quadrupolo, una cella di collisione esapolare e un oaTOF (Figura 1.12).

L'elevata sensibilità, velocità di scansione, risoluzione, accuratezza e stabilità della calibrazione dell'Orbitrap hanno determinato notevoli vantaggi rispetto agli strumenti di prima generazione, soprattutto quando si comparano gli esperimenti di massa tandem, ottenuti mediante CID.

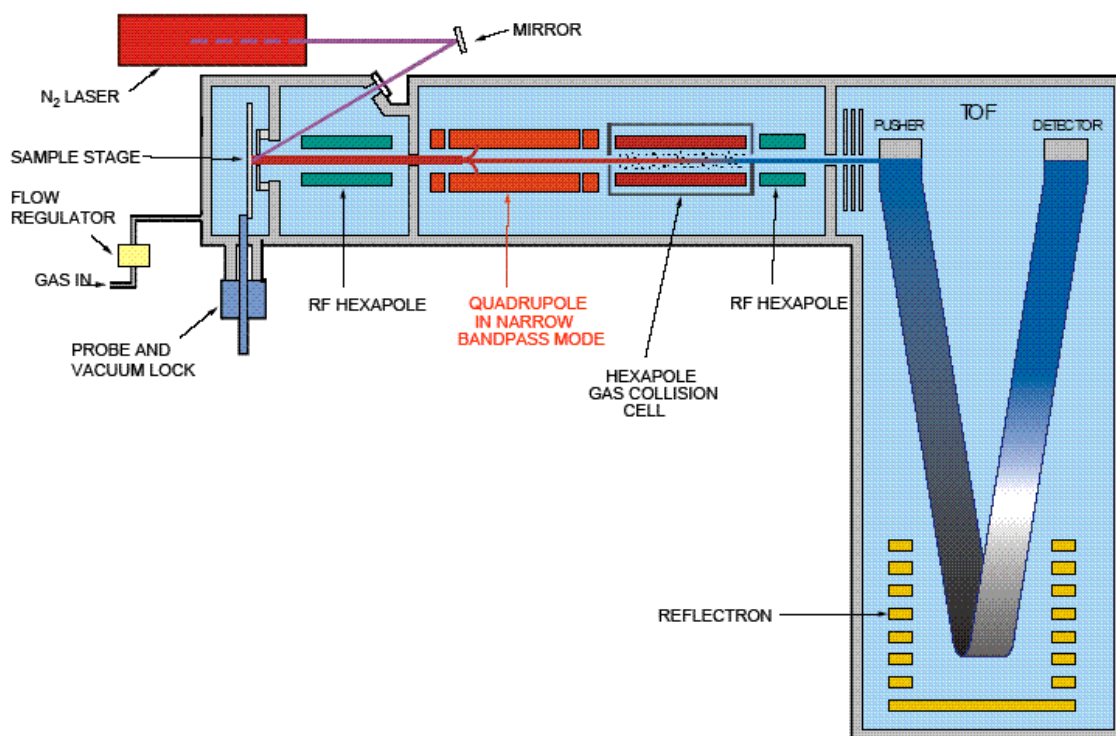


Figura 1.12 Rappresentazione schematica di un analizzatore di massa Q-TOF.

Un analizzatore di massa di ultima generazione è rappresentato dall'Orbitrap. Esso consiste di due elettrodi concentrici intorno ai quali gli ioni circolano. Non prevede l'uso di campi magnetici e la dimensione dello strumento è ridotta a pochi centimetri. Il principio su cui si basa è il movimento assiale degli ioni orbitanti intorno all'elettrodo centrale. Lo spettro di massa è ottenuto come trasformata di Fourier della corrente generata. Tale strumentazione è dotata di altissima risoluzione e accuratezza di massa.

Sulla base di questa tecnologia, è stato sviluppato un spettrometro di massa ibrido, il cui schema è riportato nella Figura 1.13.

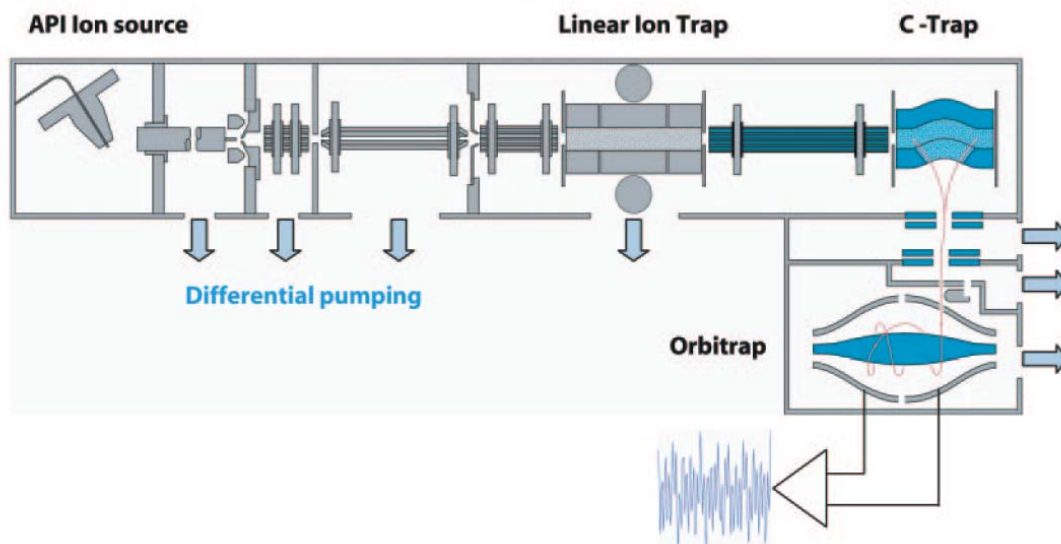


Figura 1.13 Rappresentazione schematica di un analizzatore di massa ibrido, che accoppia trappola lineare, C-Trap ed Orbitrap.

Esso consiste di una trappola ionica lineare (LTQ, *linear trap quadruple*), capace di fare MS e MSⁿ, con una alta sensitività ma una bassa risoluzione e accuratezza di massa (Figura 1.13). Gli ioni accumulati nella LTQ vengono poi trasferiti in una C-Trap, che deve il suo nome alla forma incurvata, molto simile ad una lettera C (Figura 1.13). Essa è un quadrupolo che accumula e conserva gli ioni. Nella C-Trap, gli ioni, per effetto di basse pressioni di azoto, vengono forzati ad accumularsi al centro. Da qui, vengono poi iniettati nell'*orbitrap* usando impulsi elettrici ad alto voltaggio (Figura 1.13). Gli ioni vengono catturati nell'*orbitrap* abbassando il potenziale elettrico dell'elettrodo interno, a forma di fuso. Gli ioni iniziano a circolare intorno ad esso, disegnando traiettorie ad anello che oscillano lungo l'asse z. La frequenza dell'oscillazione, inversamente proporzionale al quadrato del rapporto m/z degli ioni, è rilevata da un amplificatore differenziale. Un'onda sinusoidale viene elaborata per ogni singolo composto ed i differenti segnali generati

da una miscela di ioni vengono convertiti in spettri di massa mediante trasformata di Fourier. Nella combinazione LTQ Orbitrap, la trappola lineare accumula, isola e frammenta gli ioni. Sebbene sia possibile frammentarli anche nell'orbitrap, è più pratico e veloce manipolare gli ioni nella LTQ. Generalmente, il *pattern* di frammentazione, ottenuta della combinazione LTQ-Orbitrap è lo stesso di quello che si ottiene dalla sola LTQ, poiché l'Orbitrap funziona fondamentalmente da analizzatore di massa ad alta *performance*.

1.5 Identificazione delle proteine attraverso analisi MS

Tradizionalmente, le proteine erano identificate attraverso un approccio automatizzato di sequenziamento *de novo*, basato sul principio della degradazione di Edman. Questa metodica, oltre a richiedere grosse quantità di proteina purificata, era destinata a fallire ogni qual volta la proteina d'interesse presentava all'estremità N-terminale, un aminoacido modificato, o comunque bloccato alla reazione di Edman, che richiede necessariamente un aminoacido libero.

Oggi, i sequenziatori automatici, basati su tale principio, sono stati ampiamente sostituiti dagli spettrometri di massa, per la loro notevole sensibilità e rapidità. Esistono due differenti tipi d'approccio basati sull'utilizzo della MS per l'identificazione proteica, definiti rispettivamente con i termini "top-down" e "bottom-up".

1.5.1 Approccio *top-down*

L'approccio "top-down" prevede l'analisi di una proteina intatta, sottoposta a frammentazione in fase gassosa direttamente nello spettrometro di massa (Reid and

McLucky, 2002). Questo tipo di metodica richiede l'uso di strumentazioni molto accurate; generalmente, essa è realizzata attraverso l'utilizzo di strumentazioni FT-ICR oppure *orbitrap*, combinate con un sistema di frammentazione tipo ECD. Strumentazioni di questo tipo, oltre a preservare, grazie al sistema ECD, le modifiche post-traduzionali, garantiscono un'elevata accuratezza ed un'ottima risoluzione di massa.

Sebbene quest'approccio sia in forte crescita, esistono numerose limitazioni all'uso, legate soprattutto all'elevato costo della strumentazione, all'impossibilità di analizzare proteine ad alto peso molecolare ed alla necessità di utilizzare campioni estremamente puri. Inoltre, i supporti bioinformatici per l'analisi dei dati *top-down* necessitano ancora d'ulteriori sviluppi (Chen, 2008).

1.5.2 Approccio *bottom-up*

L'approccio *bottom-up* prevede la digestione enzimatica della proteina d'interesse in peptidi, mediante l'uso di una proteasi specifica, generalmente rappresentata dalla tripsina. I frammenti peptidici generati possono essere utilizzati direttamente per l'identificazione proteica mediante *peptide mass fingerprinting* (PMF), oppure possono essere ulteriormente frammentati nell'analizzatore di massa, per ottenere l'identificazione della proteina corrispondente mediante sequenziamento peptidico.

1.5.2.1 *Peptide mass fingerprinting* (PMF)

L'analisi PMF prevede la digestione della proteina d'interesse, generalmente separata su gel di poliacrilammide, con una proteasi specifica, per esempio tripsina. Le masse dei frammenti triptici così generati vengono misurate mediante MS e

successivamente confrontate con le masse teoriche, predette a partire da sequenze proteiche e/o nucleotidiche depositate in banche dati.

Nello specifico, le sequenze proteiche e/o nucleotidiche contenute nei *database* vengono digerite *in silico*, applicando le stesse regole di taglio dell'enzima utilizzato sperimentalmente, tenendo conto anche di altri eventuali parametri, come ad esempio i siti di taglio mancati dall'enzima (*missed cleavage*) o le modifiche introdotte nel corso delle analisi (eventuali alchilazioni). Per ciascuna sequenza depositata in banca dati, viene costruito uno spettro teorico delle masse peptidiche, che viene successivamente confrontato con quello sperimentale, allo scopo di ottenere il *match* e l'identificazione proteica. L'identificazione mediante PMF è generalmente ristretta a miscele peptidiche semplici, relative ad organismi completamente sequenziati.

1.5.2.2 Sequenziamento peptidico

Il sequenziamento peptidico si basa sull'utilizzo della spettrometria di massa tandem (MS/MS) per l'identificazione proteica. Sono stati sviluppati due differenti metodi che usano i dati MS/MS: il primo utilizza gli spettri MS/MS, ottenuti sperimentalmente e non interpretati, per interrogare un database (Sequenziamento peptidico *in silico*); il secondo consiste, invece, nell'interpretazione *de novo* degli spettri (*de novo sequencing*).

1.5.2.2.1 Sequenziamento peptidico *in silico*

Il set di frammenti ionici generati attraverso MS/MS rappresenta l'impronta digitale di uno specifico peptide. Due o più peptidi identificati in questo modo sono generalmente sufficienti per l'identificazione di una proteina. Nella dissociazione

indotta da collisione (CID), i peptidi frammentano in modo prevedibile, per cui è possibile desumerne il pattern di frammentazione. In una serie, ciascun frammento peptidico differisce da quello successivo per un aminoacido. In linea di principio, è perciò possibile determinare la sequenza aminoacidica, partendo dall'analisi delle differenze di massa tra picchi vicini, appartenenti alla stessa serie. Tuttavia, lo spettro di frammentazione di un peptide può non contenere tutte le informazioni necessarie per derivare la sequenza aminoacidica completa; esso, infatti, può contenere numerosi picchi, appartenenti a serie differenti, che complicano l'interpretazione. Ciononostante, esso può contenere informazioni sufficienti per generare una sequenza parziale, che può essere utilizzata come marcatore di sequenza (*sequence tag*) per la ricerca in database. Il confronto tra il *pattern* di frammentazione sperimentale e quello teorico, unito alla sequenza del marcatore, nonché al valore di massa dello ione parente, permette l'identificazione delle proteine via spettrometria di massa tandem. Esistono diversi algoritmi che permettono la ricerca descritta: PeptideSearch, Sequest, Mascot, ProteinProspector, Phoenix, SpectrumMill sono solo alcuni tra quelli attualmente disponibili. L'approccio descritto permette la ricerca contro banche dati che contengono informazioni genomiche incomplete (ad esempio, *database* di EST), oppure l'identificazione di proteine in miscele complesse.

1.5.2.2.2 Sequenziamento *de novo*

La ricostruzione di una sequenza peptidica basata solamente sull'analisi degli spettri MS/MS acquisiti è conosciuta come sequenziamento *de novo* (Mann and Wilm, 1994). L'interpretazione *de novo* è spesso impiegata quando l'identificazione per

interrogazione di banche dati fallisce. Questo succede quando la sequenza della proteina in analisi non è presente nel *database*, perché ancora sconosciuta oppure caratterizzata da modifiche post-traduzionali o chimiche, rispetto alla variante depositata.

L'analisi *de novo* si basa sul presupposto che i frammenti ionici della serie b e y siano, in uno spettro di frammentazione, più abbondanti rispetto ai frammenti delle altre serie. La presenza di ioni della serie a, frammenti interni o frammentazioni della catena laterale dipende fortemente dalla sequenza del peptide, dai metodi di ionizzazione e di frammentazione utilizzati. In linea di principio, la sequenza aminoacidica si ottiene per analisi le differenze tra le masse degli ioni della stessa serie (b o y). L'utilizzo di strumentazioni piuttosto sensibili permette di migliorare la qualità degli spettri MS/MS. La presenza di specifici amminoacidi (aa) in una sequenza può essere determinata grazie alla *detection* dei relativi ioni immonio. Essi compaiono a bassi valori di massa in uno spettro MS/MS e possono essere visualizzati solo grazie all'uso di analizzatori dotati di grande risoluzione a masse basse. L'utilizzo di alte energie di collisione permette, invece, di discriminare tra amminoacidi isobarici, quali leucina e isoleucina, attraverso la produzione di ioni w e d, mentre la differenziazione tra glutamina e lisina ($\Delta m/z < 0.1$ Da) è ottenuta solo mediante l'utilizzo di strumentazione FT-ICR.

Il sequenziamento di peptidi notevolmente lunghi può essere conseguito combinando i dati MS/MS, che derivano dall'utilizzo di diverse proteasi, in digestioni indipendenti. La sequenza completa viene ricavata per allineamento delle sequenze ottenute. Sebbene algoritmi computazionali per il sequenziamento *de novo* siano

implementati nei *software* per l'interpretazione dei dati, pacchetti che combinino tra loro i peptidi sequenziati sono ancora scarsamente disponibili (Chen, 2008).

Diverse strategie analitiche, che spingono la frammentazione verso *pattern* desiderati, sono state sviluppate per ridurre la complessità degli spettri. Gli approcci più impiegati sfruttano l'introduzione di gruppi funzionali che aumentano o diminuiscono la formazione di ioni di una particolare serie. Per il *knock-out* di una serie completa, sono stati utilizzati gruppi che introducono una carica negativa all'N o C terminale. Ciò consente di discriminare, rispettivamente, il *pattern* di ioni della serie b ed y (Keough et al., 2000; 2002). In maniera analoga, è possibile aumentare l'intensità di ioni di una certa serie, attaccando gruppi carichi positivamente all'N-terminale (Munchbach et al., 2000; Lindh et al., 2000). Inoltre, la conversione della lisina C-terminale in omoarginina mediante incubazione con O-metilisourea a pH alcalino (Hale et al., 2000), è ampiamente utilizzata per aumentare la ionizzazione di ioni della serie y.

Ulteriori metodi consistono nel marcare una serie ionica particolare, introducendo isotopi stabili nel peptide (Gu et al., 2002; Sonsmann et al., 2002). Si tratta di metodologie molto eleganti, ma scarsamente applicabili a causa dell'elevato costo dei reagenti e la necessità di utilizzare una discreta risoluzione di massa. Tuttavia, esse potrebbero essere utilizzate anche per la caratterizzazione quantitativa relativa (Cagney and Emili, 2002).

1.6 La proteomica vegetale

Parlare della proteomica come un campo nuovo della ricerca scientifica o un settore di recente introduzione è anacronistico e non più applicabile. La ricerca proteomica è

ormai entrata nella sua seconda decade di vita e, come dimostra la crescente quantità di dati reperibili nella letteratura internazionale, essa non ha specifici ambiti d'applicazione, ma si estende a numerosi settori della ricerca biologica, dalla medicina al campo agro-alimentare.

Gli studi proteomici sono facilitati da un crescente numero di opzioni tecniche complementari, come descritto in precedenza; tuttavia, il principio base su cui si fonda la ricerca proteomica è l'utilizzo della spettrometria di massa per l'identificazione proteica. Le metodologie analitiche legate alla MS permettono, infatti, l'identificazione delle proteine d'interesse attraverso un confronto tra gli spettri di massa elaborati sperimentalmente con quelli elaborati *in silico*, sulla base delle sequenze proteiche e nucleotidiche depositate nelle banche-dati. Tale approccio è, pertanto, strettamente dipendente dalla disponibilità di un cospicuo numero di sequenze genomiche annotate per l'organismo oggetto di studio.

Il sequenziamento massivo delle molecole di DNA complementari agli RNA messaggeri (cDNA) ha generato un largo numero di sequenze EST (expressed sequence tag; [www.ncbi.nlm.nih.gov/genomes/ PLANTS/Plantlist.html](http://www.ncbi.nlm.nih.gov/genomes/PLANTS/Plantlist.html)), che hanno rappresentato una risorsa di inaspettato valore per l'analisi sistematica dell'espressione genica, così come per l'identificazione delle proteine via spettrometria di massa.

Con il completamento di oltre 600 diversi progetti di sequenziamento genomico, tra cui quello umano (Consortium, 2004; Venter et al., 2001; Lander et al., 2001; www.genomesonline.org), l'attenzione principale è di fatto passata dal sequenziamento all'annotazione della funzione genica, nonché all'analisi dei pattern d'espressione proteica in cellule e tessuti di differente origine. Gli organismi

completamente sequenziati sono, per lo più, organismi modello, poiché la ricerca biologica ha spesso focalizzato la propria attenzione su sistemi riproducibili in laboratorio, generalmente caratterizzati da dimensioni e tempi di generazioni ridotti. Pertanto, la proteomica di tali organismi gode di uno stadio di sviluppo molto più avanzato rispetto a quella di organismi non-modello.

In ambito vegetale, due sono gli organismi completamente sequenziati: *Arabidopsis thaliana* (~120 Mb, 5 cromosomi; The Arabidopsis Genome Initiative, 2000) e *Oryza sativa* (~ 450 Mb, 12 cromosomi; Goff et al., 2002; Yu et al., 2002). I lavori di proteomica sistematica sviluppati su *Arabidopsis* e riso prima che il loro genoma fosse disponibile (Komatsu et al., 1999; Tsugita et al., 1996), hanno fornito l'identificazione di un numero relativamente basso di proteine. Studi successivi hanno portato, invece, all'identificazione di diverse centinaia di macromolecole proteiche. Circa 2500 proteine sono state identificate in riso, grazie al lavoro sviluppato da Koller e collaboratori (2002), che ha fornito il primo esempio di sequenziamento *shotgun* di proteine vegetali da tessuti interi. Baerenfaller e collaboratori (2008) hanno caratterizzato, invece, il proteoma di differenti organi in *Arabidopsis thaliana*. Complessivamente, sono state identificate 13,029 proteine. Esse corrispondono a circa la metà di tutti i geni annotati per *A.thaliana*. L'analisi del set proteico ha fornito, inoltre, 57 nuove annotazioni, tra le quali quelle di geni con differenti regioni terminali o trascritti da ORF differenti rispetto a quelli annotati, o geni creduti erroneamente pseudogeni o regioni intergeniche. Il lavoro descritto dimostra il potenziale dell'approccio proteomico nel "perfezionare" le annotazioni genomiche, mostrando l'esistenza di geni non catalogati prima e migliorando la conoscenza di quelli già noti.

Ciononostante, il potenziale investigativo della ricerca proteomica non è ancora pienamente sfruttato in campo vegetale. I dati pubblicati in letteratura che riguardano specie diverse da *Arabidopsis* e riso sono ancora scarsi (Guarino et al, 2007; Carpentier et al., 2007; Giribaldi et al., 2007; Jorin et al, 2007), sebbene l'approccio proteomico abbia dimostrato di essere un ottimo metodo per indagare diversi problemi biologici, tra i quali l'equivalenza sostanziale e la sicurezza alimentare di derivati da piante transgeniche (Ruebelt et al, 2006, a, b, c), l'espressione differenziale delle proteine durante lo sviluppo fisiologico della pianta (Macherel et al., 2007), la risposta a stress biotici o abiotici (Amme et al., 2006; Jones et al., 2007), nonché l'analisi di processi biologici complessi, quali la maturazione dei frutti (Faurobert et al, 2007).

1.7 La maturazione dei frutti

La maturazione è un fenomeno singolare, di straordinaria importanza per la biologia delle piante, per l'apporto nutrizionale che molti frutti danno all'alimentazione umana. Le principali caratteristiche organolettiche dei frutti maturi, quali colore, sapore, profumo e consistenza, nonché l'accumulo di sostanze ad elevato valore nutraceutico, sono strettamente dipendenti dal processo di maturazione dei frutti. Gli eventi che sottendono a tale processo sono piuttosto complessi, poiché richiedono l'espressione di diversi geni e l'azione coordinata di vari enzimi e macromolecole proteiche. Il fenotipo maturo è, infatti, la somma di tutti quei cambiamenti biochimici e fisiologici che avvengono durante lo sviluppo. Questi cambiamenti, sebbene variabili da specie a specie, includono modifiche nell'ultrastruttura della parete cellulare, conversione dell'amido in zuccheri, aumentata suscettibilità ai patogeni,

alterazioni nella biosintesi ed accumulo di pigmenti, aumentati livelli di sostanze aromatiche. Durante la maturazione, la parete cellulare si rigonfia ed i plastidi perdono i loro grani d'amido. L'aumentata idratazione della parete causa una disorganizzazione nella matrice della parete stessa che accompagna l'ammorbidimento del frutto (*softening*). Nel frutto maturo, le cellule sono costituite soprattutto da vacuoli, che riducono il citoplasma ad uno stato sottile, compreso tra il tonoplasto e la parete. I vacuoli accumulano, non solo composti fenolici (inclusi gli antociani), ma anche acidi organici e zuccheri che, uniti alle sostanze aromatiche prodotte, influenzano la palatabilità del frutto. I cambiamenti che riguardano il colore sono determinati da accumulo di pigmenti di nuova sintesi e degradazione di quelli esistenti, mentre i fenomeni legati alla senescenza insorgono, invece, negli ultimi stadi di sviluppo, quando il frutto si disintegra, permettendo la dispersione dei semi. La maturazione è perciò parte di un processo di sviluppo continuo, in cui diverse fasi fisiologiche possono sovrapporsi (Moing et al., 2001). Sebbene la maggior parte dei frutti mostri, durante la maturazione, modifiche che riguardano il colore, la consistenza, l'aroma e la suscettibilità ai patogeni, la classificazione generalmente accettata è basata sulla distinzione tra frutti climaterici e non-climaterici. I frutti climaterici, quali ad esempio pomodoro, banana, mela, si distinguono da quelli non climaterici, come fragola, uva e arancia, per un aumento dell'attività respiratoria durante la maturazione, accompagnata dalla produzione d'etilene, che funge da ormone stimolante la maturazione stessa (Lelievre et al., 1997). L'entità della respirazione è generalmente molto elevata nello stadio iniziale dello sviluppo, quando l'intensa divisione cellulare che accompagna la formazione di tutte le parti del frutto richiede molta energia, e si assesta su valori più bassi quando

l'accrescimento del frutto avviene per semplice espansione delle cellule già formate. Nei frutti non climaterici, la respirazione diminuisce lentamente fino alla senescenza ed alla morte del frutto stesso; mentre nei climaterici, si assiste ad un rialzo dell'attività respiratoria in prossimità della maturazione, accompagnata dalla produzione d'etilene. Anche i frutti non climaterici sembrano rispondere, in parte, all'azione di quest'ormone fitoregolatore (Alonso et al., 1995), tuttavia, esso non appare necessario per la maturazione delle specie così classificate. Al contrario, l'etilene è indispensabile per il completamento della maturazione nei frutti climaterici, come dimostrano esperimenti d'inibizione della biosintesi dell'ormone (Yen et al., 1995).

Gli studi effettuati su numerose specie vegetali hanno permesso di selezione dei sistemi - modello, a cui riferire la maggior parte delle informazioni sulla regolazione molecolare dello sviluppo e della maturazione. Il principale riferimento per i frutti climaterici è rappresentato da pomodoro. La sua importanza alimentare e nutrizionale, unita al patrimonio genetico diploide, insieme alla rapidità di propagazione dei suoi semi, alla sua efficienza di ibridazione sessuale, nonché ai brevi tempi di generazione con possibilità di crescita in serra, ha reso il pomodoro l'organismo modello per le ricerche sulla maturazione. *Arabidopsis* rimane, invece, il modello di riferimento incontrastato per la biologia vegetale, rappresentando, nello specifico, un modello per lo sviluppo del frutto deiscente. Su *A. thaliana*, infatti, sono stati condotti numerosi programmi di mutagenesi su larga-scala, che hanno fornito potenti strumenti per l'accertamento della funzione genica (Parinov and Sundaresan., 2000). Analisi funzionali sui geni AGAMOUS hanno permesso di identificare parecchi geni MADS-box che regolano lo sviluppo della siliqua

(Ferrandiz et al., 1999). Tra le specie non climateriche, particolare attenzione è stata data alla fragola, il cui processo di maturazione è stato oggetto di studio, sia pur in misura significativamente minore rispetto a pomodoro. Mediante screening differenziali, sono state identificate nuove funzioni geniche coinvolte nella maturazione (Wilkinson et al., 1995; Manning, 1998; Nam et al., 1999), mentre, attraverso studi di profilo comparativo dell'espressione genica è stato possibile analizzare variazioni dell'espressione genica globale, nel corso della maturazione ed in diversi genotipi (Aharoni et al., 2002; Carbone et al., 2006).

I dati disponibili in letteratura suggeriscono, però, l'esistenza d'eventi biochimici e meccanismi regolatori, comuni alle specie climateriche e non climateriche; essi coinvolgono, tanto in pomodoro quanto in fragola, i geni MADS-box ed agiscono nelle fasi iniziali dello sviluppo del frutto (Vrebalov et al., 2002). Sebbene la comprensione di tali meccanismi appare di fondamentale importanza per una migliore conoscenza dei processi di sviluppo, la maturazione dei frutti sembra avvenire secondo un programma di crescita specie-specifico, con caratteristiche peculiari legate alle differenze anatomiche e/o fisiologiche dei diversi frutti.

1.7.1 La maturazione della fragola

1.7.1.1 Cenni generali e cambiamenti macroscopici

La fragola appartiene alla famiglia delle Rosaceae, in particolare al genere *Fragaria*. La fragola commerciale (*Fragaria x ananassa*), caratterizzata da un *background* genetico ottoploide, è rappresentata, complessivamente, da un centinaio di cultivar (cv) differenti, coltivate in areali molto diversi per caratteristiche pedologiche ed ambientali, per tecniche di coltivazione e per standard varietale. Consumata

regolarmente in tutto il mondo come parte integrante dell'alimentazione umana, la fragola è apprezzata sul mercato tanto per le sue caratteristiche organolettiche (colore, sapore, profumo e consistenza), quanto per il contenuto in sostanze ad elevato valore nutraceutico (vitamine e antiossidanti).

Il frutto di fragola è un falso frutto, poiché esso è il prodotto dell'espansione del ricettacolo del fiore, piuttosto che dell'ovario. I frutti veri, gli acheni, si trovano all'esterno del ricettacolo, immersi nella polpa carnosa che diventa, nel corso della maturazione, rossa e succulenta. Il ricettacolo è formato, all'interno, da una parte midollare centrale (chiamata pennello), da uno strato di cellule corticali, da uno di ipodermiche ed, infine, all'esterno, da cellule epidermiche (Suutarinen et al., 1998). Gli acheni sono connessi alla parte interna del ricettacolo attraverso una serie di fasci cribro-vascolari, dai quali ricevono nutrimento.

Il frutto di fragola cresce rapidamente, raggiungendo la grandezza finale mediamente 30 giorni dopo l'antesi. Il tempo che il frutto impiega per diventare completamente maturo dipende dalla temperatura e può variare tra i 20 ed i 60 giorni. Ciò probabilmente riflette gli effetti complessivi della temperatura sulla velocità metabolica, come è stato evidenziato mediante studi sull'attività enzimatica (Manning, 1993). La grandezza del frutto a maturità può essere influenzata sia dalla posizione sull'infiorescenza che dal numero e dalla grandezza degli acheni presenti sul ricettacolo. I frutti malformati possono derivare da una scarsa impollinazione (è necessario, infatti, che almeno il 30% dei carpelli venga fecondato per avere un frutto ben formato), oppure da danni, di diversa origine e natura (insetti, gelo, etc) a carico degli acheni. Tutti questi eventi influenzano la sintesi di auxina, che regola la crescita del frutto (Given et al., 1988 Fait et al., 2008).

1.7.1.2 Cambiamenti microscopici

L'espansione del ricettacolo dopo la caduta dei petali è provocata, inizialmente, da una combinazione di divisione ed espansione cellulare. La divisione rappresenta il 15-20% della crescita totale e avviene prevalentemente prima dell'antesi. Il resto della crescita è il risultato dell'espansione cellulare, accompagnata da cambiamenti nella parete cellulare (Schwab and Raab, 2004).

Alla caduta dei petali, le cellule del ricettacolo hanno pareti dense e piccoli vacuoli. Il citoplasma delle cellule è ricco di ribosomi; l'apparato del Golgi è ben rappresentato, i plastidi, anch'essi numerosi, contengono grani di amido. Durante lo sviluppo e la maturazione del frutto, si assiste ad una graduale idratazione della parete cellulare, che si rigonfia, i plastidi rilasciano l'amido, mentre la lamella mediana, la cui funzione è tenere insieme le cellule epidermiche, ipodermiche e corticali grazie all'azione cementante della pectina, inizia a sfaldarsi (Schwab and Raab, 2004).

Nel frutto maturo, le cellule sono occupate prevalentemente dai vacuoli, che accumulano pigmenti, acidi organici e zuccheri, mentre il citoplasma è scarsamente rappresentato, ridotto ad uno strato sottile, tra tonoplasto e parete cellulare.

1.7.1.3 Cambiamenti biochimici e molecolari

I dati riportati in letteratura testimoniano numerosi cambiamenti biochimici e molecolari nel frutto di fragola durante la maturazione, molti dei quali legati all'azione regolatrice dell'auxina. Il ruolo svolto da quest'ormone nello sviluppo del frutto verde è ben documentato (Nitsch, 1950; Perkins-Veazie, 1995). L'auxina, prodotta dagli acheni e trasportata dal floema al peduncolo, regola l'espressione

genica durante la maturazione. I primi studi di espressione genica differenziale, basati sull'utilizzo d'ibridazioni sottrattive, *differential display* o sequenziamento di librerie frutto-specifiche, portarono all'identificazione di trascritti genici auxina-dipendenti, probabilmente coinvolti nella maturazione (Reddy et al., 1990). Nessuno di questi trascritti si dimostrò responsivo all'etilene.

Studi successivi (Wilkinson et al., 1995; Manning, 1998) condussero all'identificazione d'ulteriori trascritti genici, frutto-specifici, che si rivelarono, per omologia di sequenza, molto simili a geni coinvolti nell'espressione di tratti qualitativi del frutto, come, ad esempio, quelli associati alla biosintesi di antocianine nella via dei fenilpropanoidi (Manning, 1998). Mediante un approccio basato sull'utilizzo di librerie sottrattive di cDNA (Medina-Escobar et al., 1997; Medina-Escobar et al., 1998), venne poi identificato il gene per la pectato-liasi, il cui trascritto si accumulava in maniera specifica nel frutto maturo, ma poteva essere indotto nel frutto verde in seguito a rimozione degli acheni, indicando, perciò, un ruolo di modulazione negativa da parte dell'auxina.

L'utilizzo delle stesse librerie permise l'identificazione del trascritto codificante per la diidroflavonol-4 reduttasi, enzima chiave della biosintesi dei flavonoidi. L'analisi dei livelli del relativo RNA indicava che il trascritto, presente nel frutto verde, aumentava significativamente nel frutto rosso (Moyano et al., 1998).

Anche le espansine dimostrarono di essere regolate nel corso della maturazione. Partecipando alla rottura dei legami ad idrogeno a livello della parete cellulare, contribuiscono al *softening* del frutto (Civello et al., 1999). Furono caratterizzati, inoltre, diversi trascritti genici appartenenti alla famiglia delle endo 1-4- β -glucanasi, alcuni dei quali, presenti solamente nei frutti in maturazione, mostrarono

un'espressione notevolmente aumentata nel frutto rosso (Harpster et al., 1998; Trainotti et al., 1999).

Studi successivi basati sull'utilizzo di un approccio transgenico permisero di valutare i caratteri qualitativi del frutto, dopo repressione dell'espressione genica della pectato-liasi (Jimenez-Bermudez et al, 2002). I frutti transgenici si dimostrarono più consistenti rispetto a quelli controlli.

Mediante l'uso dei cDNA microarray, le variazioni dell'espressione genica su larga scala vennero indagate durante la maturazione ed in diversi genotipi (Aharoni et al., 2002; Carbone et al., 2006). La maggior parte delle variazioni riguardava trascritti associati alla pigmentazione, alla rottura della parte cellulare e alla biosintesi di composti aromatici. Vennero identificati, come differenzialmente espressi, ad esempio, trascritti relativi alla cinnamil alcol deidrogenasi (CAD) e cinnamoil CoA redattasi (CCR), coinvolti nella produzione di lignina (Salentijn et al., 2003).

La fragola non è un frutto climaterico, tuttavia il contenuto di etilene è abbondante nel frutto verde (Knee et al., 1977). Molto poco è conosciuto a proposito del ruolo che questo ormone gassoso potrebbe svolgere in fragola. Differenti approcci furono utilizzati per testare effettivamente un coinvolgimento dell'etilene nel processo di maturazione. L'uso di un inibitore competitivo dell'etilene, l'1-metilciclopropene (1-MCP), applicato a quattro diversi stadi di maturazione, permise la stima delle variazioni dell'espressione genica mediante un approccio cDNA-AFLP (Balogh et al, 2005). Fu così identificata una serie di trascritti genici regolati nello stadio verde, rispetto agli stadi più avanzati, tra cui vennero identificato alcuni trascritti menzionati in precedenza, associati alla maturazione: la pectato-liasi e l'endo 1,3-1,4- β -D-glucanasi. Inoltre, mediante l'utilizzo di PCR degenerata, furono clonati tre

diversi recettori per l'etilene (Trainotti et al, 2005), la cui espressione risultava indotta dall'etilene stesso. Gli autori interpretarono questi dati alla luce di un possibile meccanismo mediato dall'etilene, coinvolto nel processo di maturazione (Trainotti et al, 2005). Tuttavia, dati elaborati in precedenza, mediante la trasformazione di piante con un costrutto che codificava per un enzima, l' S-adenosilmetionina idrolisi, che inattivava il precursore dell'etilene, non confermarono un'influenza diretta dell'etilene sul processo di maturazione della fragola (Mathews et al., 1995).

1.7.2 L'approccio proteomico per lo studio della maturazione in fragola

La transizione del ricettacolo da tessuto verde fiorale a "frutto" rosso e succulento rappresenta un fenomeno fisiologico di notevole interesse, per le conseguenze che esso riporta sulle caratteristiche della fragola matura. La maturazione è, infatti, un processo accompagnato da cambiamenti nella consistenza, nella pigmentazione, nell'accumulo di sostanze aromatiche, zuccheri ed acidi organici, capaci di influenzare grandemente la qualità del frutto. La conoscenza dei meccanismi biochimici e molecolari alla base della maturazione è, pertanto, di fondamentale importanza, per migliorare i caratteri qualitativi del frutto stesso, ed eventualmente delineare la linea di demarcazione tra maturazione e deterioramento.

Negli ultimi anni, una quantità crescente di dati presenti in letteratura ha riguardato lo studio dell'espressione genica in fragola, durante la maturazione ed in diversi genotipi (Aharoni et al., 2002; Aharoni et al., 2004; Carbone et al., 2006; Almeida et al., 2007). La maggior parte di questi dati è stata ottenuta mediante approcci trascrittomici, che permettono lo *screening* massivo di centinaia di geni

simultaneamente. Il successo di questi approcci dipende grandemente dal progresso nel campo della genomica. Il potere delle tecniche analitiche che si basano sui trascritti genici è fortemente ridotto negli organismi non modello, a causa della carenza di informazioni genomiche o a causa della divergenza di sequenze dagli organismi modello correlati. Inoltre, le sequenze geniche sono raramente identiche da una specie all'altra, mentre i geni ortologhi sono generalmente densi di sostituzioni nucleotidiche. Un'alternativa all'analisi dell'espressione genica è perciò rappresentata dallo studio del prodotto finale dei geni, le proteine, le cui sequenze, essendo più conservate, permettono un'identificazione più efficiente per confronto con proteine ortologhe meglio conosciute (Carpentier et al., 2008).

Ciononostante, ben pochi dati presenti in letteratura riguardano lo studio della controparte proteica in fragola (Davis and Folta, 2006). Gli unici lavori in questo settore sono stati pubblicati piuttosto recentemente e riguardano aspetti diversi dalla maturazione, in particolare focalizzano l'attenzione sugli allergeni e sulla variazione del proteoma tra diverse varietà di fragola (Alm et al., 2007; 113. Hjernø et al., 2006).

Ciò può addebitarsi, almeno parzialmente, a problemi tecnici, legati alla composizione del frutto di fragola, che non rappresenta una fonte "facile" di proteine. La parete cellulare e il vacuolo costituiscono, infatti, la maggior parte della massa della cellula, mentre il citoplasma rappresenta solo l'1-2% del volume totale. La fragola è, pertanto, caratterizzata da un rapporto proteine/volume relativamente basso, cioè ha una bassa concentrazione proteica, inferiore all'1% (Clements RL, 1970). Inoltre, essa è ricca di proteasi e numerose sostanze che possono interferire severamente con la separazione delle proteine e le successive analisi. I frutti, ed in

particolare quelli maturi, rappresentano un esempio di tessuto vegetale recalcitrante, poiché contengono alti livelli di sostanze interferenti, fra i quali polisaccaridi, lipidi, composti fenolici oltre una vasta gamma di metaboliti secondari. Solo nel 2007 è stato pubblicato un protocollo ottimizzato per l'estrazione del contenuto proteico totale da frutto di fragola (Zheng et al., 2007).

Inoltre, la carenza di sequenze specie-specifiche di fragola nei database proteici ha per lungo tempo scoraggiato gli studi di proteomica su tale organismo. Tuttavia, il rapido aumento della disponibilità di sequenze nucleotidiche, associato allo sviluppo di diversi progetti di sequenziamento delle EST (www.ncbi.nlm.gov/genomes/PLANTS/Plantlist.html; Expressed sequence tag), assieme all'analisi per confronto con proteine ortologhe, ha aperto nuove prospettive per l'analisi proteomica di piante non modello, fornendo nuove opportunità per l'identificazione di proteine via spettrometria di massa (Carpentier et al., 2008).

Le analisi proteomiche rappresentano, oggi, un approccio informativo per approfondire le conoscenze su questioni biologiche complesse, come ad esempio la maturazione. In maniera simile ai profili comparativi d'espressione genica, la proteomica offre l'opportunità di esaminare cambiamenti globali nell'espressione proteica, mediante specifiche metodologie investigative. Ciò appare di notevole interesse, se si considera la possibile discrepanza tra i trascritti genici e gli effettori finali, cioè le proteine mature. La maggior parte delle funzioni biologiche nelle cellule sono eseguite dalle proteine, piuttosto che dagli mRNA. Pertanto, i profili d'espressione basati sui trascritti non sempre forniscono informazioni pertinenti alla descrizione di un sistema biologico (Carpentier et al, 2008). Esistono, infatti, una serie di meccanismi di controllo post-trascrizionali e post-traduzionali che

influiscono grandemente sul fenotipo, attraverso la modulazione della velocità di traduzione, dell'emivita dei messaggeri o delle proteine, il controllo delle modifiche proteiche e il traffico proteico intercellulare (Mata et al., 2005; Higashi et al., 2006). Proprio per i motivi elencati, un approccio proteomico applicato allo studio della maturazione della fragola risulterebbe di notevole valore, non solo per la comprensione meccanicistica dell'evento biologico, ma anche nel tentativo di stabilire correlazioni funzionali tra genotipo e fenotipo.

2 MATERIALI E METODI

2.1 Materiale vegetale

I genotipi di fragola analizzati nel corso del presente lavoro sono la *cultivar* Queen Elisa ed i suoi genotipi parentali, la selezione USB35 (Lateglow x Seneca) e la cv Miss [(Comet x Honeoye) x Dana]. Le piante, cresciute in tunnel, sono allevate nel Nord Italia. Gli stadi di maturazione utilizzati per l'estrazione proteica sono tre: immaturo (circa 20 giorni dopo l'antesi), *turning* (circa 25 giorni dopo l'antesi) e rosso (circa 30 giorni dopo l'antesi). I frutti sono raccolti da 3-4 piante, per ciascuno stadio di maturazione. Un'immagine rappresentativa dei frutti ai tre diversi stadi di maturazione è riportata nella Figura 2.1.



Figura 2.1 Immagine rappresentativa dei frutti della cv QE, a tre diversi stadi di maturazione. Da sinistra verso destra: stadio immaturo, turning e rosso.

Circa 10 frutti, per ciascuno stadio di maturazione e genotipo, sono raccolti, tagliati in quarti, congelati in azoto liquido e conservati a -80°C fino all'uso.

2.2 Estrazione proteica

Il contenuto proteico totale (circa 1mg per 5g di tessuto congelato) è estratto, per ciascuno stadio di maturazione, da un *pool* di frutti che viene sottoposto a triturazione meccanica, in azoto liquido, mediante l'utilizzo di apposito mortaio e

pestello. I campioni sono disciolti in 25ml di buffer d'estrazione, così composto: 0.7M saccarosio, 0.1M KCl, 0.5M Tris-HCl, pH 7.5, 50mM EDTA, 2% w/v β -mercaptoetanololo and 1mM PSMF). Dopo l'aggiunta di un uguale volume di fenolo basico, la miscela è agitata vigorosamente per 30 minuti e successivamente centrifugata a 5000g per 30 minuti, a 4°C. Dopo centrifugazione, la fase fenolica viene raccolta e lasciata precipitare *overnight*, a -20°C, per aggiunta di 3 volumi di una soluzione satura di acetato d'ammonio in metanolo. Mediante centrifugazione, le proteine precipitate sono raccolte a formare un *pellet*, che viene lavato prima con metanolo, poi con acetone, entrambi freddi. I *pellet* sono asciugati all'aria, a temperatura ambiente, e successivamente solubilizzati nei seguenti *buffer*: (1) 8M urea, 2% CHAPS, 10mM DDT, 0.5% IPG buffer; (2) 100mM bicarbonato d'ammonio/trifluoroetanololo; il primo è specifico per le analisi elettroforetiche, il secondo per le analisi 2-D nLC-ESI-IT-MS/MS. Prima dell'analisi elettroforetica su gel 2-D, i campioni sono purificati mediante l'utilizzo del *kit* commerciale *2-D Clean-Up Kit* (GE Healthcare, Uppsala, Svezia). La concentrazione dei campioni proteici è stimata utilizzando il *kit 2-D quant kit* (GE Healthcare, Uppsala, Svezia). I dati riportati per l'analisi *shotgun* descrittiva sono il risultato di una singola estrazione proteica, fatta da un pool di campioni, per ciascuno stadio di maturazione. Per la mappa di riferimento annotata, le proteine, estratte dai frutti a tre diversi stadi di maturazione, sono separate mediante gel 2-D. Grazie all'utilizzo d'opportuna strumentazione software (vedi paragrafo 2.6), le immagini relative a ciascuno stadio di maturazione sono sovrapposte, in modo da determinare gli *spot* comuni ai tre stadi. Le analisi DIGE, invece, sono state condotte utilizzando almeno tre estrazioni proteiche indipendenti (Tabella 2.1, a e b).

Tabella 2.1

- a) Disegno sperimentale dell'analisi DIGE per l'identificazione di proteine differenzialmente espresse nel corso della maturazione nel genotipo di riferimento QE. Quattro repliche biologiche sono analizzate per ciascuno stadio di maturazione (I, T ed R).

Numero Gel	Cy2 Standard Interno	Cy3 Campione	Cy5 Campione
1	50 µg (4.17 µg di ciascun campione I1-4, T1-4, R1-4)	50 µg campione T2	50 µg campione R1
2	50 µg (4.17 µg di ciascun campione I1-4, T1-4, R1-4)	50 µg campione I1	50 µg campione T3
3	50 µg (4.17 µg di ciascun campione I1-4, T1-4, R1-4)	50 µg campione R3	50 µg campione I4
4	50 µg (4.17 µg di ciascun campione I1-4, T1-4, R1-4)	50 µg campione I3	50 µg campione R2
5	50 µg (4.17 µg di ciascun campione I1-4, T1-4, R1-4)	50 µg campione T4	50 µg campione I2
6	50 µg (4.17 µg di ciascun campione I1-4, T1-4, R1-4)	50 µg campione R4	50 µg campione T1

- b) Disegno sperimentale dell'analisi DIGE per l'identificazione di proteine differenzialmente espresse nel corso del confronto tra genotipi (M, U e QE) allo stadio di maturazione rosso. Tre repliche biologiche sono analizzate per ciascun genotipo.

Numero Gel	Cy2 Standard Interno	Cy3 Campione	Cy5 Campione
1	50 µg (4.17 µg di ciascun campione M1-3, U1-3, QE1-3)	50 µg campione M1	50 µg campione U1
2	50 µg (4.17 µg di ciascun campione M1-3, U1-3, QE1-3)	50 µg campione U2	50 µg campione QE1
3	50 µg (4.17 µg di ciascun campione M1-3, U1-3, QE1-3)	50 µg campione QE2	50 µg campione M2
4	50 µg (4.17 µg di ciascun campione M1-3, U1-3, QE1-3)	50 µg campione U3	50 µg campione QE3
5	50 µg (4.17 µg di ciascun campione M1-3, U1-3, QE1-3)	50 µg campione M3	50 µg campione U4
6	50 µg (4.17 µg di ciascun campione M1-3, U1-3, QE1-3)	50 µg campione QE4	50 µg campione M1

2.3 Elettroforesi monodimensionale

Le proteine estratte dai frutti della cv QE a tre diversi stadi di maturazione (I, T e R) sono analizzate mediante elettroforesi 1-D. 100 µg di ciascun campione sono denaturati e analizzati mediante SDS-PAGE, seguendo il protocollo classico, proposto da Laemmli (1970). Il *pattern* elettroforetico è visualizzato mediante colorazione al Comassie colloidale (Neuhoff et al, 1988). Ciascuna *lane* del gel è tagliata in bande di uguale spessore, successivamente processate, come descritto nel paragrafo 2.7. L'immagine del gel 1-D ottenuto è riportata nella Figura 2.2.

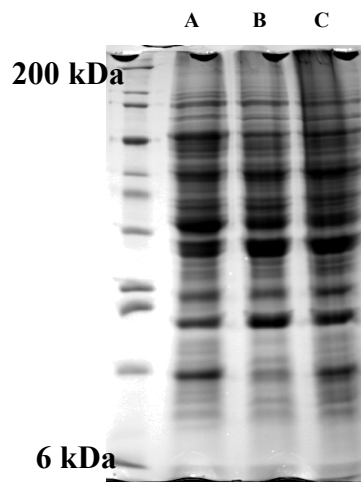


Figura 2.2 *Pattern* elettroforetico monodimensionale delle proteine estratte da QE allo stadio I (lane A), allo stadio T (lane B), allo stadio R (lane C).

2.4 Elettroforesi bidimensionale

Per la costruzione della mappa di riferimento annotata, 350 µg di proteine, estratte dai frutti di fragola a tre diversi stadi di maturazione (I, T e R) sono disciolti nel *buffer* così composto: 8M urea, 2% CHAPS, 10mM DDT, 0.5% IPG Buffer e caricate, mediante reidratazione passiva, sulle IPG strip, pH 3-10, 18 cm. Le proteine sono focalizzate utilizzando la strumentazione *Ettan IPGphor 3* (GE Healthcare,

Uppsala, Svezia), a 20°C, ad un massimo di 75 μ A/strip, applicando i seguenti voltaggi: 500 V per 1h, 1000 V per 1h, 10000 V per 3h e 55min, per raggiungere un totale di 27kV/h. Dopo la focalizzazione, le strip sono equilibrate con la soluzione E (50 mM Tris, pH 6.8, 6 M urea, 2% SDS, 30% glycerol, bromophenol blue) e DTT (2%) in un primo passaggio, e iodoacetamide (2.5%) in un secondo passaggio. L'analisi SDS-PAGE è condotta utilizzando *Ettan DALTwelve* (GE Healthcare, Uppsala, Svezia) su gel di poliacrilammide al 10% (18 cm x 24 cm x 1 mm) nel seguente tampone di corsa: 25mM Tris (pH 8.3), 1.92M glicina and 1% w/v SDS. Parametri elettrici: 5 W/gel per 30 min, poi 15 W/gel, fino a fine corsa. Gli *spot* sono colorati al Coomassie colloidale.

2.5 Analisi DIGE

Per l'analisi differenziale su gel, le proteine estratte da fragola a tre diversi stadi di maturazione (I, T, R) sono marcate con le cianine fluorescenti, specifiche per l'analisi DIGE: Cy2, Cy3, Cy5.

50 μ g di ciascun estratto (pH 8.5) sono mescolati con 400 pmoli dell'opportuna cianina e la reazione è lasciata incubare in ghiaccio, per almeno 30 minuti. La reazione viene bloccata per addizione di 1 μ g di lisina. Coppie di campioni marcati con Cy3 e Cy5 sono unite allo standard interno, marcato con Cy2, e portate ad un volume di 450 μ l per aggiunta del *buffer*: 8 M urea, 2% CHAPS, 20 mM DTT, and 0.5% IPG buffer. I campioni così mescolati sono applicati alla IPG strip, pH 3-11, 24 cm, mediante reidratazione passiva. L'equilibratura delle strip e l' SDS-PAGE sono condotte come descritto in precedenza per l'analisi 2-D.

Per il confronto tra genotipi, le proteine estratte dai frutti rossi di QE, M ed U sono marcate come descritto sopra.

Al fine di ottenere dati statisticamente significativi, quattro repliche biologiche sono analizzate nel confronto tra stadi di maturazione, mentre 3 repliche biologiche sono analizzate nel confronto tra genotipi (Tabella 2.1, a e b).

2.6 Acquisizione delle immagini ed analisi dei dati

I gel 2-D sono acquisiti mediante l'uso di scanner commerciali e analizzati con il software *Image Master 2D Platinum 6.0* (GE Healthcare, Uppsala, Svezia). Le immagini corrispondenti ai tre diversi stadi di maturazione (I, T e R) sono paragonate a due a due, utilizzando lo stadio T come riferimento. Soltanto gli *spot* contemporaneamente presenti nei tre diversi stadi di maturazione sono prelevati dal gel e sottoposti ad analisi di spettrometria di massa tandem, per l'identificazione proteica.

Le immagini relative ai gel DIGE sono acquisite mediante l'uso dell'apposito scanner *Ettan DIGE Imager* (GE Healthcare, Uppsala, Svezia). Le scansioni sono acquisite ad una risoluzione di 100 μm . Le immagini, ritagliate e processate grazie all'uso del software *ImageQuantTL* (GE Healthcare, Uppsala, Svezia), sono analizzate per mezzo del software DeCyder (versione 6.5; GE Healthcare, Uppsala, Svezia). Nello specifico, il modulo DIA (*differential in-gel analysis*) ha permesso la *detection* degli spot, la quantificazione del loro volume e la relativa normalizzazione del valore rispetto allo standard interno. Il modulo BVA (*biological variation analysis*) ha permesso, invece, la correlazione degli spot proteici tra i diversi gel. Le proteine espresse in maniera differenziale sono selezionate sulla base dello *Student's*

t-test ($p \leq 0.05$); quelle che hanno mostrato una variazione nel rapporto d'espressione di almeno 1.5 volte (*fold change* \leq o \geq 1.5), sono considerate differenzialmente accumulate. Gli spot proteici d'interesse sono prelevati manualmente dal gel preparativo, ottenuto in seguito ad analisi di 800 μ g di proteine.

2.7 Procedure di digestione proteica

Gli spot proteici prelevati dal gel 2-D e le bande proteiche prelevate dal gel 1-D sono tagliati grossolanamente mediante l'utilizzo di bisturi, ridotti per reazione chimica con DTT, alchilati alla cisteina per reazione con la iodoacetamide e digeriti per trattamento con tripsina (*overnight*, 37°C). I peptidi triptici sono estratti dal gel per trattamento con 25mM ammonio bicarbonato/acetonitrile 1:1. Le risultanti miscele peptidiche sono concentrate in evaporatori a centrifuga e purificate grazie all'uso di *Cleanup C18 Pipette Tips* (Agilent Technologies, Santa Clara CA, USA), prima dell'analisi nLC-ESI-IT-MS/MS.

Per l'analisi 2-D nLC-ESI-IT-MS/MS, gli estratti proteici (100 μ g) sono solubilizzati in 100 mM ammonio bicarbonato/trifluoroetanolo 1:1, ridotti con DTT 10 mM (1 ora a 60 °C); successivamente, essi sono incubati con iodoacetamide (IAA) per 1 ora al buio, a temperatura ambiente. Il prodotto di reazione è diluito, per aggiunta di ammonio bicarbonato 25 mM, fino al 5% di trifluoroetanolo. Successivamente, una soluzione di CaCl₂ è aggiunta fino alla concentrazione finale di 1 mM. La digestione proteica è eseguita per trattamento con tripsina (rapporto enzima/substrato 1:30), a 37°C, *overnight*. Il digerito è acidificato per trattamento con acido trifluoroacetico (concentrazione finale 0.2%) e purificato mediante l'uso di *Peptide Cleanup C18 Spin Tubes* (Agilent Technologies, Santa Clara CA, USA).

2.8 Analisi nLC-ESI-IT-MS/MS

Le miscele peptidiche, provenienti dagli *spot* 2-D e/o dalle bande 1-D, sono sottoposte ad analisi nLC-ESI-IT-MS/MS, utilizzando la strumentazione HPLC-Chip/MS (Agilent Technologies, Santa Clara CA, USA). Essa è costituita da un nano HPLC (Agilent, serie 1200) con pompa analitica, autocampionatore, pompa capillare, una sorgente ionica ESI ed trappola ionica (LC/MSD Trap XCT Ultra- Agilent Technologies, Santa Clara CA, USA). Le miscele peptidiche vengono separate su una colonna HPLC-Chip (40 nl Trap, 75 μm x 43 mm, 5 μm C-18SB ZX; Agilent Technologies, Santa Clara CA, USA), usando le seguenti fasi mobili: A) 97% H₂O/3% ACN/0.1% FA; B) 97% ACN/3% H₂O/0.1% FA. Il flusso operativo della pompa capillare è pari a 4 $\mu\text{l}/\text{min}$, mentre quello della pompa analitica è pari a 300 nl/min. L'eluzione in gradiente avviene rispettando i seguenti protocolli: 1) 2 min, 5% B; 40 min, 40% B; 48 min, 80% B; 50 min, 80% B per gli spot da gel 2-D; 2) 2 min, 5% B; 60 min, 50% B; 70 min, 80% B, per le bande 1-D.

Gli spettri sono acquisiti in un *range* di massa compreso tra 300 e 2000 m/z. I principali parametri operativi della trappola ionica sono riportati: flusso del drying gas, 4 L/min a 300°C; voltaggio del capillare 1820 V; skimmer, 1:30; capillary exit, 75V; trap drive, 85V. Il tempo di accumulazione massimo all'interno della trappola è settato a 150 ms, mentre il target ionico massimo è settato a 500.000. L'acquisizione è controllata da una procedura di scansione dipendente dai dati, che frammenta fino a 5 ioni, tra quelli più abbondanti. L'esclusione dinamica dei composti più abbondanti è attivata dopo la registrazione di 4 spettri relativi allo stesso composto in due minuti. La frammentazione degli ioni doppia carica è impostata come favorita, mentre la scansione MS/MS è inserita in un *range* di m/z compreso tra 200 a 1800.

2.9 Analisi 2-D nLC-ESI-IT-MS/MS

Le miscele peptidiche sono separate *on-line* attraverso cromatografia liquida bidimensionale. Nello specifico, lo scambio cationico forte (SCX- Zorbax bioSCX Series II, 0.30 x35 mm, 3.5 μ m), per l'analisi in prima dimensione, è associato alla cromatografia liquida a fase inversa, per l'analisi in seconda dimensione. Lo strumentazione utilizzata è il sistema HPLC-Chip/MS (Agilent Technologies, Santa Clara CA, USA). L'eluizione dalla colonna SCX avviene per iniezione di sali a concentrazione crescente. Il metodo utilizzato consiste nell'iniezione di 8 μ l di 11 diverse concentrazioni di formiato d'ammonio: 30, 60, 90, 150, 200, 300, 400, 500, 700, 1000, 2000 mM. Dopo ciascuno *step*, il campione passa dalla colonna SCX a quella a fase inversa, da cui viene eluito per idrofobicità relativa, variando il gradiente del solvente B: 2 min, 5% B; 60 min, 50% B; 70 min, 80% B. Gli spettri sono acquisiti così come descritto nel paragrafo precedente.

2.10 Identificazione proteica

Gli spettri MS/MS, registrati per gli spot 2-D, le bande 1-D o le analisi 2-D nLC-ESI-IT-MS/MS, sono sottoposti all'identificazione proteica. Le ricerche sono eseguite per confronto con le banche dati, proteiche e nucleotidiche, delle *Rosaceae*, disponibili, rispettivamente, sul sito web di NCBI e di Tigr. Il software utilizzato per il sequenziamento peptidico è SpectrumMill (Rev A.03.02.06; Agilent Technologies, Santa Clara CA, USA). Le identificazioni sono auto-validate, utilizzando i seguenti parametri: score minimo per l'attribuzione di un'identità pari ad 11; punteggio minimo per la validazione degli spettri MS/MS pari a 6, se lo ione ha carica singola o doppia, pari ad 8 se ha una carica superiore; il parametro SPI è fissato almeno pari al

60% per gli ioni 2+, mentre almeno pari al 70% per gli ioni diversamente carichi. Al fine di minimizzare il numero di falsi positivi, tutti gli spettri MS/MS sono confrontati contro il *database* invertito. Solo gli spettri con uno score su tale database di almeno due volte più piccolo rispetto allo score sul database originale sono presi in considerazione per l'auto-validazione. Inoltre, la differenza minima tra i punteggi della prima e della seconda interpretazione dello stesso spettro deve essere almeno pari a 2. Dopo l'auto-validazione, solo il sottogruppo degli spettri validi è processato ammettendo modifiche specifiche, come ad esempio la metionina ossidata.

3 RISULTATI

3.1 Mappa 2-D del frutto di fragola

La mappa di riferimento del proteoma espresso dal frutto di fragola durante la maturazione è stata realizzata analizzando, mediante elettroforesi 2-D, i campioni proteici estratti dal frutto della cv QE a tre diversi stadi di maturazione: I, T e R. Le analisi delle immagini, effettuate mediante l'uso di Image Master 2D Platinum (GE Healthcare, Uppsala, Svezia), hanno fornito la *detection* di 568 proteine per lo stadio I, 622 per lo stadio T, 520 per lo stadio R.

Le immagini acquisite sono state confrontate, a due a due, con lo stadio di riferimento, lo stadio T, per identificare gli *spot* contemporaneamente presenti nei 3 gel, e quindi nei tre stadi di maturazione. Sono stati così identificati 366 *spot* costantemente presenti nel corso della maturazione. Al fine di ottenere una mappa di riferimento delle proteine espresse nel frutto durante la maturazione, 274 *spot* tra i 366 complessivamente individuati, sono stati manualmente prelevati dal gel, digeriti ed analizzati via spettrometria di massa tandem. Gli spettri MS/MS sono stati confrontati con i *database* pubblicamente disponibili, utilizzando come software per il sequenziamento peptidico *in silico* Spectrum Mill (Agilent Technologies, Santa Clara CA, USA). Sono stati identificati complessivamente 158 *spot* proteici, corrispondenti a 181 proteine totali e 111 proteine distinte. La lista delle proteine identificate è riportata nella Tabella 3.1, mentre l'immagine della mappa annotata è riportata nella Figura 3.1.

Tabella 3.1 Lista degli spot identificati sulla mappa 2-D

	Spot	Numero di accesso	Organismo	Copertura di sequenza /n° Peptidi	Spectrum Mill score	Proteina	Mr (kDA)		pI	
							Sper	Teor	Sper	Teor
Allergeni	2264	<u>90185680</u>	Fragaria x ananassa	9 (1)	24.24	Major strawberry allergen Fra a 1-B	18012	17816	6.02	5.97
	2270	<u>90185678</u>	Fragaria x ananassa	41 (7)	122.97	Major strawberry allergen Fra a 1-B	17497	17816	5.59	5.97
	2267	<u>90185686</u>	Fragaria x ananassa	41 (7)	115.53	Major strawberry allergen Fra a 1-C	17744	17745	5.97	5.97
Metabolismo aminoacidi	1954	<u>57895993</u>	Fragaria vesca	5 (1)	21.85	S-adenosyl-L-methionine synthetase	42440	43208	5.56	5.42
	1953	<u>57895993</u>	Fragaria vesca	22 (5)	93.45	S-adenosyl-L-methionine synthetase	42440	43208	5.53	5.42
	1952	<u>57895993</u>	Fragaria vesca	22 (5)	100.41	S-adenosyl-L-methionine synthetase	42478	43208	5.5	5.42
	1937	<u>57895993</u>	Fragaria vesca	22 (5)	104.5	S-adenosyl-L-methionine synthetase	43404	43208	5.5	5.42
	1956	<u>89550929</u>	Fragaria vesca	13 (3)	52.35	S-adenosylmethionine synthetase	42327	42877	5.46	5.37
	1942	<u>89550929</u>	Fragaria vesca	3 (2)	25.64	S-adenosylmethionine synthetase	42818	42877	5.56	5.37
	2082	<u>89553672</u>	Fragaria vesca	11 (2)	34.47	cysteine synthase	30822	34461	5.42	5.53
Organizzazione, comunicazione cellulare e trasduzione del segnale	2110	<u>55375985</u>	Malus x domestica	16 (4)	68.98	14-3-3 family protein	31191	29683	4.84	4.75
	1966	<u>33563040</u>	Fragaria x ananassa	21 (3)	49.91	actin	42140	16134	5.39	5.91
	1961	<u>33563040</u>	Fragaria x ananassa	53 (5)	74.57	actin	42140	16134	5.29	5.91
	1958	<u>33563040</u>	Fragaria x ananassa	53 (6)	93.76	actin	42028	16134	5.35	5.91
	1957	<u>116812230</u>	Rosa hybrid cultivar	20 (3)	46.42	actin	42290	31112	5.15	5.16
	1923	<u>15029368</u>	Rosa hybrid cultivar	15 (6)	98.26	alpha tubulin subunit	46612	49717	5.21	4.96
	1921	<u>15029368</u>	Rosa hybrid cultivar	8 (5)	90.49	alpha tubulin subunit	47089	49717	5.09	4.96
	1904	<u>78210838</u>	Pyrus pyrifolia	12 (5)	74.80	beta tubulin like protein	48882	43595	4.82	6.17
	2302	<u>9992891</u>	Prunus avium	30 (3)	39.05	calmodulin	17353	16825	3.33	4.11
Metabolismo energetico e del carbonio	2195	<u>89542907</u>	Fragaria vesca	6 (1)	16.80	aconitase C-terminal domain-containing protein	30307	26792	4.61	6.33
	2022	<u>89542385</u>	Fragaria vesca	14 (3)	50.47	Alpha-1,4-glucan-protein synthase	37043	41573	5.58	5.73
	1831	<u>89543693</u>	Fragaria vesca	3 (1)	18.50	apgm (Phospoglyceromutase)	58920	60791	5.35	5.44
	2174	<u>89547145</u>	Fragaria vesca	15 (3)	47.21	putative ATP synthase	33397	27599	5.63	6.8
	2242	<u>77980570</u>	Fragaria vesca	5 (2)	38.81	ATP synthase D chain, mitochondrial, putative	19804	21801	5.34	5.22
	2223	<u>51047003</u>	Fragaria x ananassa	30 (7)	86.99	ATP synthase delta chain, mitochondrial precursor	26092	27213	9.4	9.56
	1894	<u>7578491</u>	Kerria japonica	4 (2)	30.65	ATP synthase, B subunit	49719	53339	5.36	5.57
	1940	<u>24636275</u>	Fragaria x ananassa	9 (5)	85.64	Citrate synthase, mitochondrial precursor	43256	52302	6.07	8.66
	2040	<u>10645188</u>	Fragaria x ananassa	34 (15)	250.49	cytosolic aldolase	34930	38519	6.14	6.93
	2036	<u>78216493</u>	Malus x domestica	11 (3)	52.73	cytosolic malate dehydrogenase	35834	35597	5.53	6.01
	1871	<u>89557589</u>	Fragaria vesca	11 (3)	41.28	D-3-phosphoglycerate dehydrogenase, putative	54309	62130	5.48	8.41
	1895	<u>63192024</u>	Prunus armeniaca	17 (3)	41.45	enolase	50399	15855	5.55	5.13

Tabella 3.1 Lista degli spot identificati sulla mappa 2-D

	2070	57896259	Fragaria vesca	26 (9)	138.97	fructose-bisphosphate aldolase, putative	33361	42333	6	8.19
	2060	16226117	Fragaria x ananassa	33 (4)	44.08	glyceraldehyde 3-phosphate dehydrogenase	33532	14108	7.7	7.86
	2058	16226117	Fragaria x ananassa	33 (4)	45.45	glyceraldehyde 3-phosphate dehydrogenase	33790	14108	6.47	7.86
	2053	89545819	Fragaria vesca	9 (2)	32.95	malate dehydrogenase precursor	34311	43188	5.38	8.11
	2043	24638017	Fragaria x ananassa	14 (7)	122.54	Malate dehydrogenase, mitochondrial precursor	35199	35518	6	8.72
	2029	24638017	Fragaria x ananassa	5(2)	29.27	Malate dehydrogenase, mitochondrial precursor	35742	35518	5.6	8.72
	2103	57896118	Fragaria vesca	14 (3)	46.46	oxygen evolving complex photosystem II	31066	35196	5.26	5.63
	1822	51241136	Malus x domestica	4 (1)	18.68	Phosphoglucomutase	60338	63475	5.5	6.01
	1976	89543537	Fragaria vesca	10 (2)	40.59	Phosphoglycerate kinase	41768	50050	5.45	6.64
	2228	7443216	Nicotiana tabacum	3 (1)	16.94	photosystem II complex protein 2 prec.	24575	28655	5.67	6.84
Metabolismo energetico e del carbonio	1942	77982483	Fragaria vesca	21 (5)	84.58	putative dihydrolipoamide succinyltransferase	42818	49881	5.56	9.07
	1937	77982483	Fragaria vesca	5 (1)	20.69	putative dihydrolipoamide succinyltransferase	43404	49881	5.5	9.07
	2075	89554481	Fragaria vesca	23 (5)	75.45	putative fructokinase 2	32770	34971	4.71	5.20
	2083	89554481	Fragaria vesca	23 (5)	66.96	putative fructokinase 2	31781	34971	4.94	5.20
	1846	17225598	Fragaria x ananassa	1 (1)	16.39	pyruvate decarboxylase	57732	65313	5.57	6.03
	2059	115603869	Fragaria x ananassa	19 (2)	39.99	Pyruvate dehydrogenase E1	34050	40027	5.04	5.56
	1907	1352782	Fragaria x ananassa	4 (2)	29.44	Ribulose bisphosphate carboxylase large chain	48386	51668	5.6	6.04
	1784	1352782	Fragaria x ananassa	7 (3)	50.20	Ribulose bisphosphate carboxylase large chain	76303	51668	5.57	6.04
	2175	78069814	Rosa woodsii	23 (1)	16.75	Triose phosphate isomerase	33497	27089	6.05	5.54
	2176	89556994	Fragaria vesca	35 (10)	197.23	Triosephosphate isomerase	33531	33532	6	7.65
	2171	13431949	Fragaria x ananassa	18 (5)	81.10	Triosephosphate isomerase	33329	33532	5.33	7.65
	2170	13431949	Fragaria x ananassa	9 (2)	26.04	Triosephosphate isomerase	33296	33532	5.38	7.65
	2168	89556994	Fragaria vesca	15 (3)	49.32	Triosephosphate isomerase	33263	33532	5.48	7.65
Varie	1935	89548661	Fragaria vesca	6 (1)	18.25	gamma-aminobutyrate transaminase	43997	57251	5.91	6.72
	2127	89542206	Fragaria vesca	16 (4)	60.68	Hypothetical protein	31952	29572	6	5.89
	2311	51048118	Fragaria x ananassa	87 (2)	13.63	Hypothetical protein	17323	42629	3.39	9.61
	2085	89547309	Fragaria vesca	24 (7)	107.85	Hypothetical protein	31458	16074	4.73	4.24
	2063	89547309	Fragaria vesca	32 (9)	124.30	Hypothetical protein	33790	16074	4.68	4.24
	2075	89547309	Fragaria vesca	30 (8)	125.45	Hypothetical protein	32770	16074	4.71	4.24
	2056	89547309	Fragaria vesca	35 (9)	135.52	Hypothetical protein	34224	16074	4.76	4.24
	2083	89550858	Fragaria vesca	46 (10)	177.12	Plasma membrane intrinsic polypeptide	31781	23336	4.94	4.95
	2308	34851124	Prunus avium	22 (5)	93.23	putative glycine-rich RNA-binding protein	17328	17328	5.49	7.84
	2307	34851124	Prunus avium	26 (5)	71.27	putative glycine-rich RNA-binding protein	17339	17328	5.35	7.84
	2297	51048034	Fragaria x ananassa	10 (2)	31.79	putative glycine-rich RNA-binding protein	17366	17328	5.48	7.84
	2300	89540933	Fragaria vesca	18 (3)	46.42	putative glycine-rich RNA-binding protein	17366	17328	5.35	7.84

Tabella 3.1 Lista degli spot identificati sulla mappa 2-D

	2202	<u>89545646</u>	Fragaria vesca	16 (3)	56.52	quinone reductase family protein	28868	21794	5.59	6.08
	2074	<u>2465008</u>	Fragaria vesca	20 (4)	66.56	ripening-induced protein	32770	35807	5.06	6.29
	2025	<u>2465008</u>	Fragaria x ananassa	35 (11)	193.25	ripening-induced protein	36480	35807	5.49	6.29
	2031	<u>2465008</u>	Fragaria vesca	10 (3)	23.33	ripening-induced protein	36109	35807	5.14	6.29
	2029	<u>2465008</u>	Fragaria vesca	11 (3)	41.23	ripening-induced protein	35742	35807	5.6	6.29
	1998	<u>2465008</u>	Fragaria vesca	15 (4)	66.2	ripening-induced protein	39410	35807	5.51	6.29
Varie	1997	<u>2465008</u>	Fragaria x ananassa	15 (4)	39.14	ripening-induced protein	39410	35807	5.47	6.29
	2021	<u>2465008</u>	Fragaria vesca	21 (7)	120.26	ripening-induced protein	37233	35807	5.45	6.29
	2013	<u>2465008</u>	Fragaria x ananassa	3 (1)	20.49	ripening-induced protein	37423	35807	5.4	6.29
	2254	<u>71819138</u>	Malus x domestica	19 (3)	50.49	methyltransferase-like protein	19124	18064	5.4	5.35
	1992	<u>89558212</u>	Fragaria vesca	28 (9)	155.89	succinyl CoA ligase beta subunit-like protein	41180	45310	5.27	5.7
	2260	<u>89548322</u>	Fragaria vesca	8 (1)	17.10	temperature-induced lipocalin	18629	21462	5.27	5.6
	2092	<u>10187173</u>	Fragaria x ananassa	30 (7)	100.92	unnamed protein product	30509	30509	5.33	5.33
	2016	<u>89548637</u>	Fragaria vesca	48 (11)	197.14	Adenosine kinase	37519	37519	5.14	5.14
Metabolism dei nucleotidi	2015	<u>89548637</u>	Fragaria vesca	11 (2)	35.82	Adenosine kinase	37519	37519	5.01	5.14
	2010	<u>89548637</u>	Fragaria vesca	47 (12)	192.31	Adenosine kinase	38318	37519	4.92	5.14
	2321	<u>89554153</u>	Fragaria vesca	21 (6)	90.67	NDPK III	17284	25677	6.12	9.52
	1849	<u>89556189</u>	Fragaria vesca	21 (5)	85.15	chaperonin 60 alpha subunit	57341	61430	4.56	5.23
Metabolismo, modificazione e folding proteico	2329	<u>89555280</u>	Fragaria vesca	18 (2)	28.08	copper chaperone	17248	8936	5.16	6.23
	2191	<u>57896209</u>	Fragaria vesca	13 (3)	54.42	cp10-like protein	30765	26779	5.52	7.77
	1848	<u>89554968</u>	Fragaria vesca	14 (5)	61.29	Protein disulfide-isomerase precursor (PDI)	57732	55547	4.88	4.95
	1840	<u>89554968</u>	Fragaria vesca	4 (1)	19.47	Protein disulfide-isomerase precursor (PDI)	57732	55547	4.72	4.95
	2289	<u>89551412</u>	Fragaria vesca	27 (2)	29.20	60S acidic ribosomal protein P3	17406	12212	3.86	4.55
	1760	<u>89544737</u>	Fragaria vesca	6 (1)	16.06	elongation factor EF-2	157313	55032	5.59	5.87
	2278	<u>85376261</u>	Rosa chinensis	22 (4)	45.61	eukaryotic translation initiation factor eIF5A	17497	17131	5.36	5.3
	2268	<u>85376261</u>	Rosa chinensis	15 (3)	43.74	eukaryotic translation initiation factor eIF5A	17631	17131	5.51	5.3
	1938	<u>1946218</u>	Malus x domestica	3 (1)	20.19	kn1-like protein	43256	45180	5.53	6.02
Sintesi/accumulo proteico	2177	<u>89549395</u>	Fragaria vesca	18 (5)	87.45	mitochondrial prohibitin 1	33464	30565	9.56	8.8
	2233	<u>23032118</u>	Prunus dulcis	17 (2)	31.49	prunin	22974	63024	8.66	6.59
	2226	<u>23032118</u>	Prunus dulcis	17 (2)	30.19	prunin	24760	63024	8.6	6.59
	2041	<u>14272197</u>	Fragaria x ananassa	11 (1)	18.61	prunin	35199	63024	5.43	6.59
	2053	<u>14272197</u>	Fragaria x ananassa	15 (3)	64.54	prunin	34311	63024	5.38	6.59
	2052	<u>14272197</u>	Fragaria x ananassa	13 (2)	41.55	Prunin 1 precursor	34575	35247	5.35	5.2
	2276	<u>2677830</u>	Prunus armeniaca	22 (3)	50.75	ribosomal protein L12	17564	17890	9.67	9.02
	2108	<u>20154774</u>	Rosa	13 (3)	27.14	translation initiation factor-related	31160	25154	4.71	4.69

Tabella 3.1 Lista degli spot identificati sulla mappa 2-D

Proteolisi	2192	<u>89547579</u>	Fragaria vesca	18 (4)	60.78	Peptidase T1A, proteasome beta-subunit	30650	28791	6.02	5.5
	2150	<u>89541375</u>	Fragaria vesca	14 (3)	47.92	Proteasome alpha-subunit	32732	27253	5.43	5.68
	2152	<u>89541375</u>	Fragaria vesca	32 (7)	106.42	Proteasome alpha-subunit	32831	27253	5.47	5.68
	2180	<u>77981986</u>	Fragaria vesca	5 (1)	14.06	Proteasome subunit alpha type 5	33033	25969	4.52	4.7
	2197	<u>89551799</u>	Fragaria vesca	13 (2)	17.43	Proteasome subunit alpha type 5	29857	25969	5.34	4.7
	2167	<u>77981986</u>	Fragaria vesca	27 (7)	114.96	Proteasome subunit alpha type 5	33263	25969	4.52	4.7
	2143	<u>89547162</u>	Fragaria vesca	18 (4)	55.12	proteasome, alpha subunit-like protein	32535	27079	7.32	6.93
	2146	<u>89547162</u>	Fragaria vesca	9 (3)	46.24	proteasome, alpha subunit-like protein	32568	32535	8.53	7.32
Metabolismo secondario /Biosintesi di componenti cellulari	2079	<u>53988164</u>	Fragaria x ananassa	8 (2)	26.03	aldo/keto reductase	31700	36115	5.57	5.78
	2078	<u>53988164</u>	Fragaria x ananassa	18 (4)	50.54	aldo/keto reductase	31139	36115	5.53	5.78
	1949	<u>51872683</u>	Fragaria x ananassa	15 (4)	63.99	anthocyanidin synthase	42818	42819	5.48	5.48
	1946	<u>51872683</u>	Fragaria x ananassa	14 (4)	59.72	anthocyanidin synthase	42818	42819	5.44	5.48
	2089	<u>51047818</u>	Fragaria x ananassa	22 (5)	72.94	Anx1	30901	35841	6.06	6.38
	2088	<u>89544075</u>	Fragaria vesca	28 (8)	131.27	Anx1	31059	36062	6.05	6.19
	2084	<u>77981641</u>	Fragaria vesca	3 (1)	14.42	Anx1	31619	36062	5.84	6.19
	1984	<u>109635989</u>	Rosa luciae	29 (2)	34.23	AXS1	41510	43636	5.52	5.49
	2207	<u>71979900</u>	Fragaria x ananassa	14 (3)	59.63	chalcone isomerase	27703	25464	5.54	5.19
	2149	<u>71979900</u>	Fragaria x ananassa	10 (2)	32.21	chalcone isomerase	32765	25464	5.01	5.19
	2157	<u>71979900</u>	Fragaria x ananassa	23 (5)	89.54	chalcone isomerase	33063	25464	5.08	5.19
	2179	<u>71979900</u>	Fragaria x ananassa	6 (1)	16.00	chalcone isomerase	33157	25464	5.3	5.19
	1971	<u>71979904</u>	Fragaria x ananassa	16 (8)	125.51	chalcone synthase	41917	42631	6.04	6.18
	1969	<u>71979904</u>	Fragaria x ananassa	12 (5)	72.49	chalcone synthase	42028	42631	6.02	6.18
	1962	<u>71979904</u>	Fragaria x ananassa	22 (9)	132.45	chalcone synthase	42215	42631	6	6.18
	1973	<u>71979904</u>	Fragaria x ananassa	18 (7)	125.32	chalcone synthase	41842	42631	6	6.18
	1983	<u>13919603</u>	Fragaria vesca	37 (2)	34.81	dihydroflavonol 4-reductase	41583	6091	5.93	8.15
	1973	<u>13919603</u>	Fragaria x ananassa	37 (2)	34.25	dihydroflavonol 4-reductase	41842	6091	6	8.15
	2001	<u>51493451</u>	Fragaria x ananassa	7 (3)	25.77	flavanone 3-hydroxylase	40250	41107	5.58	5.61
	2003	<u>51493451</u>	Fragaria x ananassa	6 (2)	22.97	flavanone 3-hydroxylase	40533	41107	5.55	5.61
	2017	<u>51493451</u>	Fragaria x ananassa	27 (13)	189.36	flavanone 3-hydroxylase	37519	41107	5.54	5.61
	2018	<u>51493451</u>	Fragaria x ananassa	25 (10)	128.31	flavanone 3-hydroxylase	38050	41107	5.56	5.61
	1998	<u>51493451</u>	Fragaria x ananassa	32 (13)	164.24	flavanone 3-hydroxylase	39410	41107	5.51	5.61
	1997	<u>51493451</u>	Fragaria x ananassa	24 (8)	97.92	flavanone 3-hydroxylase	39410	41107	5.47	5.61
	1994	<u>6760443</u>	Fragaria x ananassa	16 (5)	74.32	O-methyltransferase	35379	39819	6.09	5.7
	1995	<u>6760443</u>	Fragaria x ananassa	12 (4)	61.95	O-methyltransferase	40533	39819	5.42	5.7
2003	<u>6760443</u>	Fragaria x ananassa	5 (2)	29.2	O-methyltransferase	40533	39819	5.55	5.7	

Tabella 3.1 Lista degli spot identificati sulla mappa 2-D

Metabolismo secondario /Biosintesi di componenti cellulari	2017	6760443	Fragaria x ananassa	15 (5)	68.43	O-methyltransferase	37519	39819	5.54	5.7
	2018	6760443	Fragaria x ananassa	14 (5)	77.22	O-methyltransferase	38050	39819	5.56	5.7
	1998	6760443	Fragaria x ananassa	44 (14)	221.43	O-methyltransferase	39410	39819	5.51	5.7
	1997	6760443	Fragaria x ananassa	25 (9)	136.3	O-methyltransferase	39410	39819	5.47	5.7
	2021	6760443	Fragaria x ananassa	12 (4)	53.99	O-methyltransferase	37233	39819	5.45	5.7
	2013	6760443	Fragaria x ananassa	6 (2)	24.54	O-methyltransferase	37423	39819	5.4	5.7
	2086	15808674	Fragaria x ananassa	9 (2)	33.82	putative quinone oxidoreductase	30822	36198	5.36	6.61
	2042	15808674	Fragaria x ananassa	50 (15)	270.60	putative quinone oxidoreductase	35109	36198	5.47	6.61
	2150	57895710	Fragaria vesca	10 (3)	39.38	ascorbate peroxidase	32732	27949	5.43	5.88
2103	57896454	Fragaria vesca	8 (2)	39.41	ATGLX1; lactoylglutathione lyase	31066	31923	5.26	5.19	
1801	57011884	Fragaria vesca	14 (2)	30.34	binding (TPR)	63707	63707	6	6	
1797	57011884	Fragaria vesca	6 (1)	13.58	binding (TPR)	66421	63707	5.67	6	
1796	57011884	Fragaria vesca	14 (2)	40.10	binding (TPR)	67351	63707	5.58	6	
2161	5257546	Fragaria x ananassa	30 (5)	108.07	cytosolic ascorbate peroxidase	33196	27282	5.55	5.69	
2160	2738949	Fragaria x ananassa	48 (8)	116.93	cytosolic ascorbate peroxidase	33162	27278	5.5	5.69	
2183	89544420	Fragaria vesca	5 (1)	21.94	dienelactone hydrolase family protein	32420	25858	5.13	5.17	
2182	89544420	Fragaria vesca	11 (3)	54.89	dienelactone hydrolase family protein	32542	25858	5.34	5.17	
2179	89544420	Fragaria vesca	5 (1)	21.42	dienelactone hydrolase family protein	33157	25858	5.3	5.17	
2218	89550364	Fragaria vesca	18 (4)	51.59	glutathione S-transferase	26987	23680	5.6	6.13	
1804	57896185	Fragaria vesca	3 (1)	18.71	Heat shock 70 kDa protein	62210	72536	5.37	5.95	
Risposta allo stress/Detossifica zione	2235	54306589	Fragaria x ananassa	12 (1)	20.17	heat shock protein	22380	10235	4.86	6.14
	1857	24637539	Prunus dulcis	7 (4)	36.40	heat shock protein 60	55994	57755	5.36	5.26
	1852	24637539	Prunus dulcis	17 (8)	119.33	heat shock protein 60	55804	57755	5.33	5.26
	1807	6969976	Malus x domestica	4 (2)	35.17	high molecular weight heat shock protein	63276	71204	6.1	5.17
	1788	6969976	Malus x domestica	36 (22)	430.85	high molecular weight heat shock protein	71204	71204	5.11	5.17
	2096	3243234	Pyrus communis	12 (5)	75.35	isoflavone reductase related protein	30540	33824	6.01	6.02
	2093	57896334	Fragaria vesca	7 (2)	23.63	isoflavone reductase related protein	30540	33824	5.58	6.02
	2281	2911276	Fragaria x ananassa	8 (1)	14.21	LMW heat shock protein	17419	17419	6.17	6.17
	2275	2911276	Fragaria x ananassa	26 (6)	66.83	LMW heat shock protein	17497	17419	5.45	6.17
	2262	2911276	Fragaria x ananassa	8 (1)	17.01	LMW heat shock protein	18147	17419	5.47	6.17
	2270	2911276	Fragaria x ananassa	48 (8)	101.63	LMW heat shock protein	17497	17419	5.59	6.17
	2290	51046962	Fragaria x ananassa	13 (2)	37.14	peptidylprolyl isomerase (cyclophilin)	17408	18303	8.84	8.68
	2282	51046962	Fragaria x ananassa	26 (5)	91.95	peptidylprolyl isomerase (cyclophilin)	17430	18303	7.26	8.68
	2098	51049496	Fragaria x ananassa	19 (4)	57.94	putative lactoylglutathione lyase	30786	31952	5	5.11
	2078	89550785	Fragaria vesca	22 (4)	58.74	putative thioredoxin reductase	31139	37756	5.53	5.85

Tabella 3.1 Lista degli spot identificati sulla mappa 2-D

	2247	<u>89541551</u>	Fragaria vesca	36 (8)	127.32	putative thioredoxin-dependent peroxidase	20688	17338	5.47	5.15
Risposta allo	2323	<u>89556432</u>	Fragaria vesca	26 (3)	42.85	superoxide dismutase	17283	16731	5.6	6.27
stress/Detossifica	2210	<u>37999810</u>	Prunus persica	7 (2)	40.57	Superoxide dismutase [Mn]	27703	25466	6.029	8.62
zione	2082	<u>89542413</u>	Fragaria vesca	25 (3)	50.31	thiosulfate sulfurtransferase	30822	41389	5.42	6.52
	2080	<u>89542413</u>	Fragaria vesca	14 (2)	31.03	thiosulfate sulfurtransferase	31298	41389	5.48	6.52
	2148	<u>60279272</u>	Solanum tuberosum	3 (2)	29.85	pom30	32666	29331	8.97	8.85
	2132	<u>5031279</u>	Prunus armeniaca	7 (2)	24.92	porin	32242	29732	6.12	7.06
Trasporto	2237	<u>15242952</u>	Arabidopsis thaliana	2 (1)	13.29	protein transport protein-related	22803	98473	5.55	5.07
	2151	<u>89544539</u>	Fragaria vesca	5 (3)	52.86	VDAC2.1	32732	29151	8.81	9.13
	2146	<u>89544539</u>	Fragaria vesca	4 (1)	21.95	VDAC2.1	32568	29151	8.53	9.13

Gli spot sono stati identificati mediante nLC-ESI-IT-MS/MS e raggruppati in classi funzionali.

Per ciascuna identificazione, sono riportati: il peso molecolare ed il punto isoletttrico- teorico e sperimentale

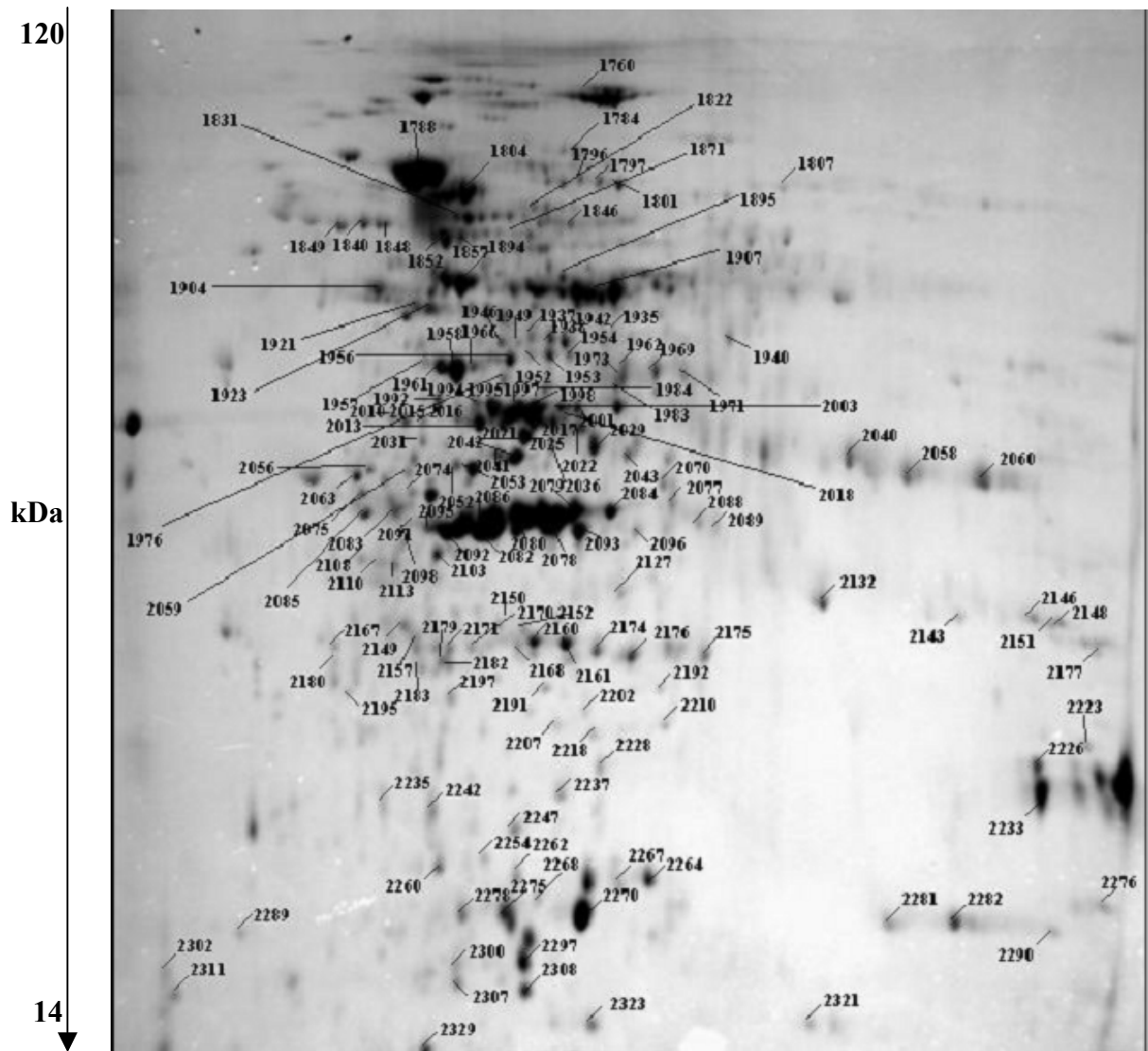


Figura 3.1 Mappa proteica 2-D del frutto di fragola. La mappa di riferimento è stata costruita sovrapponendo le immagini 2-D di tre diversi gel, ognuno corrispondente ad un diverso stadio di maturazione (I, T, R). In essa sono mostrate le posizioni dei 158 spot identificati, annotati come riportato nella tabella 2. Gli spot mappati sono stati identificati contemporaneamente in tutti e tre gli stadi analizzati, perciò essi rappresentano un *set* di proteine espresse nel frutto nel corso della maturazione.

3.2 Identificazione delle proteine mediante approcci *shotgun*

L'uso combinato di differenti metodologie investigative è generalmente consigliato nel corso delle analisi di proteomica sistematica, allo scopo di aumentare il numero di macromolecole proteiche identificate. Per implementare i dati sul proteoma espresso dal frutto di fragola durante la maturazione, gli estratti proteici, esaminati via 2-D, sono stati ulteriormente analizzati, utilizzando due differenti approcci di proteomica *shotgun*.

Nel primo caso, i campioni proteici, estratti dal frutto di fragola (genotipo QE) a tre diversi stadi di maturazione (I, T, R) sono stati solubilizzati in un *buffer* contenente SDS ed analizzati mediante SDS-PAGE. Ciascuna *lane* del gel è stata tagliata in fette di uguale spessore, a loro volta digerite ed analizzate mediante nL-ESI-IT-MS/MS, per l'identificazione delle componenti proteiche (*vedi Materiali e metodi, paragrafo 2.3- 2.7-2.9*).

Nel secondo caso, gli stessi estratti proteici sono stati analizzati via 2-D nLC-ESI-IT-MS/MS, metodologia analitica che consente di analizzare miscele peptidiche complesse mediante cromatografia liquida bidimensionale, accoppiata *on-line* alla spettrometria di massa tandem (*vedi Materiali e metodi, paragrafo 2.7-2.9*).

Gli spettri MS/MS, collezionati per entrambi gli approcci di proteomica *shotgun* (circa 20.000 per stadio di maturazione nel primo caso, circa 80.000 per stadio di maturazione nel secondo caso), sono stati sottoposti ad identificazione proteica mediante sequenziamento peptidico *in silico* (*vedi Introduzione, paragrafo 1.5*), utilizzando il software SpectrumMill (Agilent Technologies, Santa Clara CA, USA) (*vedi Materiali e metodi, paragrafo 2.10*).

La prima strategia ha fornito l'identificazione di 494 proteine per lo stadio I, 555 per lo stadio T e 281 per lo stadio R. Esse corrispondono complessivamente a 870 proteine distinte (Tabella 3.2). Il secondo approccio ha fornito, invece, l'identificazione di 516, 525 e 366 proteine, rispettivamente per gli stadi I, T ed R, identificando complessivamente 906 proteine differenti (Tabella 3.2). Le proteine identificate, simultaneamente, da entrambe le strategie, per ciascuno stadio di maturazione sono: 182 per lo stadio I, 197 per lo stadio T e 118 per lo stadio R. Inoltre, sono state identificate 41 proteine uniche, contemporaneamente presenti in tutti e tre gli stadi, in entrambi gli approcci analitici (Tabella 3.3). L'integrazione dei *dataset*, ottenuti per ciascun metodo, ha permesso, invece, l'identificazione complessiva di 828 proteine per lo stadio I, 883 per quello T e 529 per lo stadio R. In totale, sono state identificate 1370 differenti proteine, espresse nel frutto di fragola nel corso della maturazione (Tabella 3.2); esse includono tutte le proteine già riportate nelle mappa 2-D. Il 36% di queste identificazioni è stato eseguito per confronto con *database* proteici, reperibili sul sito web di NCBI; il resto è stato eseguito per confronto con la banche dati di sequenze nucleotidiche di organismi appartenenti alla famiglia delle *Rosaceae*, depositate sul sito www.tigr.org (vedi *Materiali e metodi, paragrafo 2.10*).

	1D SDS-PAGE nLC-ESI-IT-MS/MS				2D nLC-ESI-IT-MS/MS				Combinati			
	I	T	R	Totale	I	T	R	Totale	I	T	R	Totale
Peptidi totali†	1287	1525	971	3783	1692	2058	1632	5382	2979	3583	2603	9165
Peptidi distinti‡	886	1128	643	1828	932	969	777	1753	2094	2278	1726	2889
Proteine distinte§	494	555	281	870	516	525	366	906	828	883	529	1370

Tabella 3.2 Analisi numerica di peptidi e proteine rilevate ed identificate, mediante approcci di proteomica *shotgun*, nel proteoma espresso del frutto di fragola durante la maturazione. Le proteine, estratte dai frutti a tre diversi stadi di maturazione, I, T ed R, sono state analizzate mediante 1D-SDS-PAGE nLC-ESI-IT-MS/MS e 2D nLC-ESI-IT-MS/MS, come descritto in Materiali e Metodi. La sezione “Combinati” rappresenta l’integrazione dei set di dati generati da entrambi gli approcci, per ciascuno stadio di maturazione.

† Il numero totale di spettri di massa tandem auto-validati ed interpretati dal software SpectrumMill è elencato come peptidi totali.

‡ Il numero totale di peptidi unici, identificati per ciascun campione, in ciascun metodo è elencato come peptidi distinti. Nella sottosezione “Totale” e nella sezione “Combinati”, il numero totale di peptidi unici non è additivo perché molti peptidi sono identificati in campioni multipli e/o in entrambi i metodi.

§ Il numero totale di proteine uniche, identificate per ciascun campione, in ciascun metodo, è elencato come proteine distinte. Nella sottosezione “Totale” e nella sezione “Combinati”, il numero totale di peptidi unici non è additivo perché molti peptidi sono identificati in campioni multipli e/o in entrambi i metodi.

Proteina	Organismo	Numero d'accesso	Peptidi / Copertura di sequenza %	Spectrum Mill score
Cytoplasmic aconitate hydratase	Fragaria vesca	DY673485	2(12)	35.73
Translational elongation factor 1 subunit b	Pisum sativum	38232568	2(7)	40.88
Structural constituent of ribosome	Arabidopsis thaliana	15231316	5(43)	59.27
Adenosine kinase 1	Arabidopsis thaliana	15232763	4(17)	61.76
Enolase	Fragaria x ananassa	CO816781	3(40)	65.57
Actin-depolymerizing factor	Fragaria vesca	TA2568_57918	3(51)	67.4
Rhamnose biosynthesis 3	Arabidopsis thaliana	15231926	5(10)	73.21
Cytosolic phosphoglycerate kinase	Pisum sativum	9230771	5(19)	89.74
Prunin precursor	Fragaria x ananassa	AI795142	5(31)	90.82
Glutamate--ammonia ligase	Fragaria x ananassa	CO817764	6(53)	93.3
Tubulin beta-7 chain	Arabidopsis thaliana	15227559	6(19)	93.59
17.9 kDa class II heat shock protein	Fragaria vesca	DY667157	5(57)	99.64
60S ribosomal protein L8-3	Fragaria vesca	TA2913_57918	7(36)	102.08
Glutamate decarboxylase 1	Fragaria vesca	TA2130_57918	7(35)	106.52
Serine hydroxymethyltransferase 4	Arabidopsis thaliana	15236375	6(16)	106.76
Unknown protein	Fragaria vesca	DY676138	5(41)	107.77
Domain protein At1g63220	Fragaria vesca	TA2685_57918	8(44)	115.88
Cytosolic malate dehydrogenase	Malus x domestica	78216493	7(35)	120.67
Monodehydroascorbate reductase	Fragaria vesca	TA2236_57918	9(30)	123.5
Peptidyl-prolyl cis-trans isomerase	Fragaria x ananassa	TA512_3747	7(48)	125.93
Actin	Rosa hybrid cultivar	116812230	7(42)	128.54
Putative epoxide hydrolase	Fragaria vesca	TA2405_57918	8(44)	128.59
Glyceraldehyde 3-phosphate dehydrogenase	Fragaria x ananassa	16226117	9(62)	128.97
Triosephosphate isomerase, cytosolic	Fragaria x ananassa	CO817313	8(58)	137.4
Cytosolic ascorbate peroxidase	Fragaria x ananassa	5257552	8(61)	137.69
NADP-dependent isocitrate dehydrogenase	Fragaria vesca	TA2708_57918	10(22)	148.19
Elongation factor 1 alpha subunit	Malus x domestica	4138179	12(30)	165.8
F1-ATPase alpha subunit	Rubus sp. JPM-2004	57115599	11(38)	174.25
Clathrin heavy chain, putative	Arabidopsis thaliana	30681617	11(11)	185.4
Annexin P38	Fragaria vesca	TA3280_57918	13(42)	194.14
Cell division cycle 48	Arabidopsis thaliana	15232776	15(22)	211.77
Beta tubulin like protein	Pyrus pyrifolia	78210838	14(44)	217.83
14-3-3 family protein	Malus x domestica	55375985	15(66)	242.71
Heat shock protein 81-1	Arabidopsis thaliana	15237214	14(19)	252.53
Alpha-tubulin	Prunus dulcis	20413	15(55)	256.46
ATP synthase subunit beta	Zea mays	114420	15(39)	268.55
Phospholipase D alpha	Fragaria x ananassa	58891689	18(29)	281.65
Cytosolic aldolase	Fragaria x ananassa	10645188	19(71)	289.36
Elongation factor EF-2	Arabidopsis thaliana	6056373	18(29)	298.55
O-methyltransferase	Fragaria x ananassa	6760443	21(76)	319.32
High molecular weight heat shock protein	Malus x domestica	6969976	22(47)	400.3

Tabella 3.3. Proteine contemporaneamente presenti in tutti e tre gli stadi di maturazione, identificate con entrambi gli approcci analitici.

3.2.1 Distribuzione funzionale delle proteine identificate

Le proteine identificate, raggruppate per funzione putativa, sono state classificate sulla base dei dati disponibili in letteratura e/o utilizzando le informazioni contenute nel *database* Swiss-Prot/TrEMBL (Figura 3.2).

La categoria più rappresentata è quella delle proteine, alle quali non è stato ancora possibile associare una funzione certa (Proteine “non classificate”, Figura 3.2). Ben 243 proteine sono state collocate in questo gruppo. Esse corrispondono al 17.7% del totale. Tra di esse, 73 proteine sono state identificate per confronto con sequenze EST di fragola, a cui non è stato possibile associare, mediante BLAST, alcun’omologia con le sequenze depositate in NCBI. Per tale motivo, esse potrebbero essere considerate proteine esclusive di fragola.

Le proteine sono state classificate come miscellanee, quando, invece, l’informazione a disposizione era complessa, ipotetica o basata su scarsa omologia di sequenza. Le proteine raggruppate in questa classe sono 175, corrispondenti al 12.8% del totale. Tra di esse, ve ne sono alcune di particolare interesse, poiché particolarmente abbondanti e/o potenzialmente coinvolte nel processo di maturazione.

Un numero relativamente alto di proteine identificate è coinvolto nel metabolismo energetico e del carbonio (163), nella sintesi proteica (150), nel trasporto (99), nella proteolisi (87), nella risposta allo stress e la detossificazione (87), nel metabolismo degli aminoacidi (69), nel metabolismo secondario (64), nell’organizzazione e comunicazione cellulare/trasduzione del segnale (51) e nel metabolismo dei nucleotidi e acidi nucleici (36). Un numero minore di proteine è coinvolto in altre attività cellulari, ad esempio il metabolismo proteico (34), il metabolismo lipidico (24), la regolazione della trascrizione (23), l’omeostasi redox e il trasferimento

elettronico (20), la fotosintesi e respirazione (19), il metabolismo della parete (18). Sono stati individuati anche 8 allergeni (Figura 3.2).

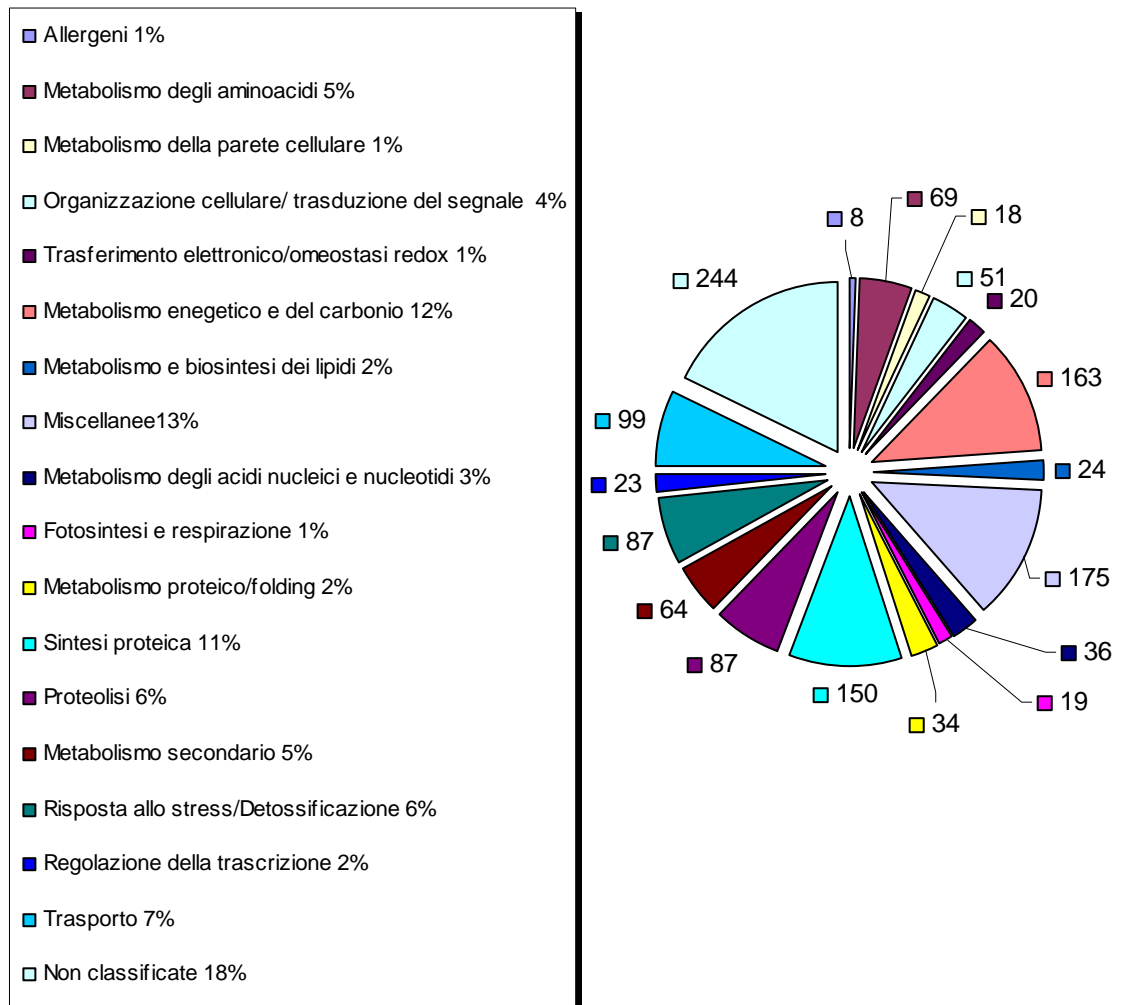


Figura 3.2 Distribuzione funzionale di tutte le proteine identificate. Le proteine, estratte dai frutti di fragola allo stadio I, T ed R, sono state analizzate mediante approcci di proteomica *shotgun* per l'analisi proteomica su larga scala. L'integrazione di dataset ottenuti da entrambi i metodi ha permesso l'identificazione di 1370 proteine espresse. Tali proteine sono state classificate in gruppi funzionali sulla base dell'informazione disponibile in letteratura e/o in base all'informazione disponibile per esse nel database Swiss-Prot-TrEMBL. Il numero di proteine identificate per classe funzionale è stato riportato all'interno della figura, mentre le loro relative percentuali sono riportate vicino la legenda.

3.3 Variazioni del proteoma durante la maturazione

L'analisi proteomica differenziale del frutto di fragola durante il processo di maturazione è stata ottenuta mediante 2-D DIGE, accoppiata alla spettrometria di massa tandem per l'identificazione proteica degli *spot* differenzialmente accumulati. Il disegno sperimentale dell'analisi DIGE è riportato nella Tabella 2.1a di *Materiali e Metodi*. La figura 3.3, di seguito riportata, rappresenta il gel numero 1 del disegno sperimentale descritto.

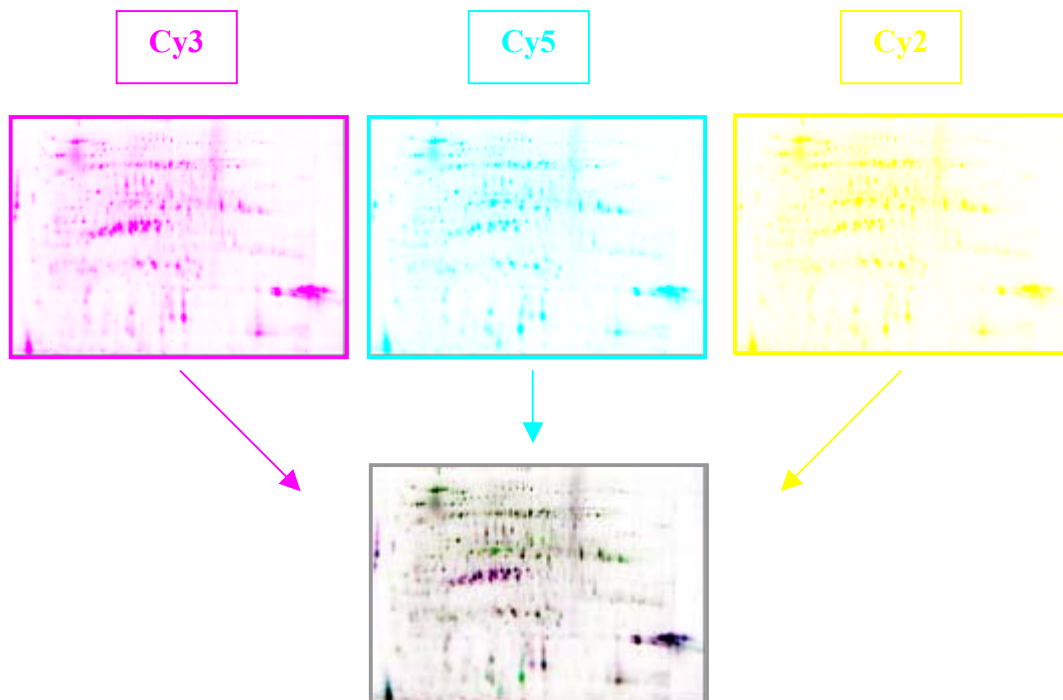


Figura 3.3 Immagine relativa al gel n°1 dell'analisi DIGE (Tabella 2.1a), per l'identificazione di proteine differenzialmente espresse nel corso della maturazione. Il campione marcato con Cy2 è la replica biologia n°2 delle proteine estratte dal frutto allo stadio turning, mentre il campione marcato con Cy5 è la replica n°1 delle proteine estratte dal frutto allo stadio rosso. Lo standard interno è marcato con Cy2. Esso comprende uguali quantità di tutti i campioni analizzati.

L'analisi comparativa ha confrontato gli stadi I, T ed R allo scopo di: 1) esaminare la natura dinamica del network proteico in relazione allo sviluppo del frutto; 2)

identificare le proteine variabili nel frutto; 3) legare le variazioni proteiche al ruolo fisiologico svolto durante la maturazione del frutto.

Le mappe DIGE ottenute dagli stadi I, T ed R hanno rilevato approssimativamente 2500 spot risolti per gel. Tra di essi, il 50% ha mostrato un coefficiente di correlazione inferiore al 20%, in ciascuno stadio di maturazione. Gli spot che hanno mostrato, tra le coppe di campioni poste a confronto, un cambiamento nell'espressione maggiore di 1.5 volte, sono stati considerati differenzialmente espressi se, contemporaneamente, hanno passato il filtro dell'analisi statistica dello *Student's t-test* con un *p-value* inferiore od uguale ad 0.05. Inoltre, il test ANOVA è stato eseguito su ciascuno degli spot accumulati in maniera differenziale, al fine di validarne l'andamento durante la maturazione (Tabella 3.4). La maggioranza degli *spot* risolti non ha mostrato un accumulo differenziale durante la maturazione; solo 76 *spot*, circa il 3% del totale, sono risultati complessivamente variabili (*Student's t test* $p\text{-value} \leq 0.05$) (Figura 3.4). Tra i 76 *spot* differenzialmente accumulati (Figura 3.4), 74 sono stati manualmente prelevati dal gel preparativo (Figura 3.5), colorato al Comassie colloidale, ed opportunamente trattati per l'analisi di spettrometria di massa.

Sono stati identificati, complessivamente, 68 *spot* proteici. Per i rimanenti 6 *spot*, le identificazioni sono fallite, anche se, nella maggior parte dei casi, la quantità di proteina in essi contenuta era sufficiente per l'analisi di massa e gli spettri ottenuti erano d'ottima qualità. In questi casi, la mancanza d'identificazione può essere attribuita alla limitata disponibilità di sequenze proteiche e/o nucleotidiche per l'organismo oggetto di studio e alla limitata omologia con altre specie.



Figura 3.4 Mappa elettroforetica DIGE, in cui i 76 spot differientemente accumulati sono mostrati con un cerchio nero.

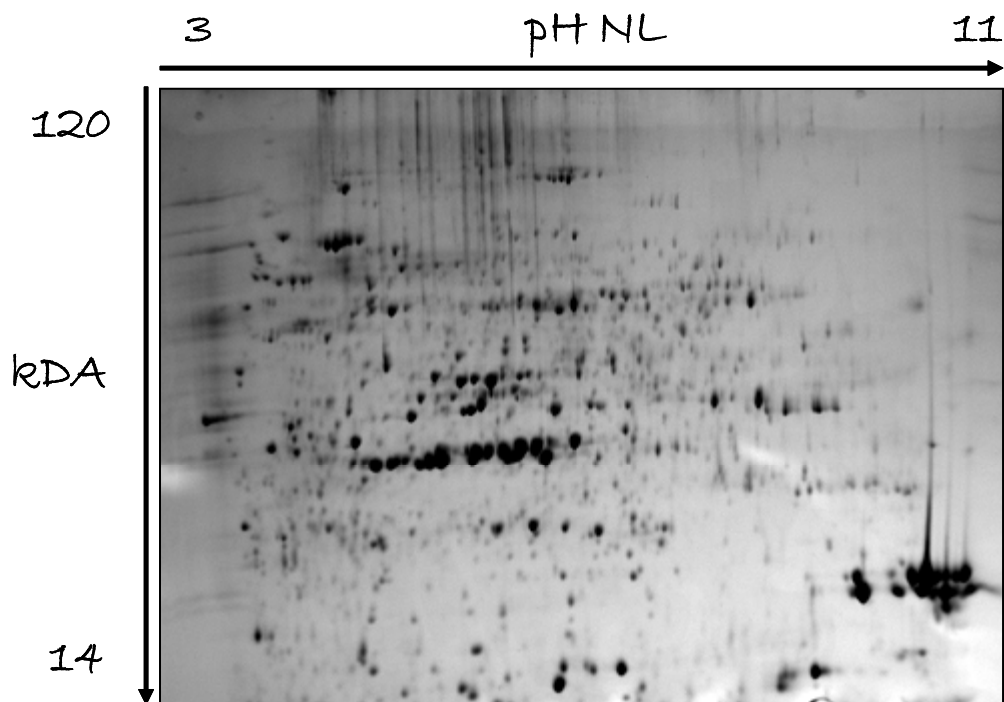


Figura 3.5 Gel preparativo (da 800 µg di proteine, colorazione Comassie colloidale) da cui sono stati manualmente prelevati gli spot differenzialmente accumulati.

3.3.1 Proteine regolate in R vs I

Il confronto R vs. I ha fornito la *detection* di 43 *spot up-regolati* in R. Tra questi, 38 *spot* sono stati manualmente prelevati dal gel preparativo e sottoposti ad analisi di spettrometria di massa. L'analisi MS/MS ha permesso di identificare 34 differenti proteine (Tabella 3.4a). Alcuni *spot* distinti hanno fornito l'identificazione di proteine uguali, suggerendo l'ipotesi che i diversi *spot* corrispondano, in realtà, a diverse modifiche post-traduzionali della stessa proteina o a membri di una medesima famiglia multigenica. Le proteine identificate in più di uno *spot* sono: la chalcone isomerasi (2 *spot*), la chalcone sintasi (4 *spot*), la flavanone 3-idrossilasi (2 *spot*), la Hsp 60 (2 *spot*), la proteina correlata all'isoflavone reduttasi (2 *spot*), la proteina "major storage" (2 *spot*), l'O-metiltransferasi (4 *spot*), la quinone ossidoreduttasi (2 *spot*) e la proteina indotta nella maturazione (3 *spot*) (Tabella

3.4a). Inoltre, 9 *spot* hanno fornito identificazioni proteiche multiple, ciascuno di essi conteneva, cioè, 2 o 3 differenti proteine. Tale fenomeno è causato essenzialmente dalla co-migrazione di proteine, che hanno massa molecolare e punto isoelettrico molto simili tra loro.

Il confronto R vs. I ha identificato, inoltre, 15 *spot down-regolati* in R. Per 12 di essi, l'analisi di spettrometria di massa ha fornito identificazioni proteiche singole, mentre, per i restanti 3 *spot*, essa ha fornito identificazioni doppie (Tabella 3.4b). Inoltre, 2 proteine (l'aldo/cheto reduttasi ed il precursore della disolfuro isomerasi) sono state identificate in più di uno *spot* (Tabella 3.4b).

3.3.2 Proteine regolate in R vs T

Il confronto R vs. T ha evidenziato, complessivamente, 26 *spot* proteici regolati in maniera differenziale: 20 *spot up-regolati* e 6 *spot down-regolati* in R. Le analisi di spettrometria di massa tandem hanno fornito, rispettivamente, le identificazioni di 18 proteine sovra-espresse e 6 proteine sotto-espresse (Tabella 3.4c,d). Tra le proteine accumulate in maniera variabile durante la maturazione, la chalcone sintasi, la flavanone 3-idrossilasi, la proteina “major storage”, l'O-metiltransferasi e la “ripening induced protein” sono rappresentate, ognuna, in più di uno *spot*, sostenendo l'ipotesi che i diversi *spot* corrispondano, probabilmente, a diverse modifiche post-traduzionali della stessa proteina o a membri di una medesima famiglia multigenica. (Tabella 3.4c,d).

3.3.3 Proteine regolate in T vs I

Per investigare, fin dalle fasi iniziali, il processo maturativo, lo stadio T è stato confrontato con quello I. Sono stati individuati, complessivamente, 28 *spot* regolati

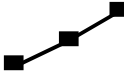



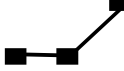




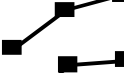




in maniera differenziale; per 25 di essi, l'identificazione della proteina corrispondente è stata ottenuta mediante analisi di spettrometria di massa. Nello specifico, sono stati identificati, in T vs. I, 22 *spot* sovra-espresi e 3 *spot* sotto-espresi (Tabella 3.4e,f).






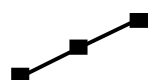
Anche in questo caso, sono stati trovati alcuni *spot up-regolati*, probabilmente corrispondenti a diverse isoforme di una stessa proteina. Nello specifico, 6 differenti proteine sono state trovate in almeno due *spot* distinti. Esse sono: la chalcone isomerasi (2 *spot*), la chalcone sintasi (3 *spot*), la HMW Hsp (2 *spot*), la proteina correlata all'isoflavone reductasi (2 *spot*), la LMW Hsp (3 *spot*) e la RuBisCO attivasi (2 *spot*). Inoltre, gli *spot* 3350, 2040 e 3736 hanno fornito identificazione proteica multipla, portando alla *detection*, all'interno dello stesso *spot*, di due differenti proteine. Esse sono la proteina HSP 22 e la proteina ribosomale 60S L21 per lo *spot* 3350, l'aldo/cheto reductasi e l'annessina per il 2040, la Perossiredoxina e la proteina 423, per lo *spot* 3736 (Tabella 3.4e,f).

Tabella 3.4 Proteine differenzialmente espresse durante la maturazione


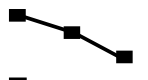

a) Proteine up-regulate in R vs. I

Spot	Proteina	Organismo	Numero d'accesso	Peptidi / coverage	Spectrum Mill score	Fold change	Andamento durante la maturazione I>T>R (ANOVA $p < 0.05$)
589	Heat shock protein Hsp70	Fragaria x ananassa	CO816704	2 (18)	39.84	1.84	
656	putative quinone oxidoreductase O-methyltransferase	Fragaria x ananassa	15808674	10 (37)	151.63	1.60	
		Fragaria x ananassa	6760443	2 (10)	26.65		
796	ripening-induced protein	Fragaria vesca	2465008	2 (7)	24.79	1.55	
836	heat shock protein 60	Prunus dulcis	24637539	3 (6)	36.77	1.98	
842	heat shock protein 60	Prunus dulcis	24637539	16 (29)	239.88	1.67	
1184	GDP dissociation inhibitor Putative RNA binding protein Eukaryotic initiation factor 4A-8	Fragaria vesca	TA2618_57918	18 (42)	295.02	1.65	
		Fragaria vesca	TA3500_57918	5 (19)	92.62		
		Fragaria vesca	DY667187	3 (12)	55.76		
1333	anthocyanidin synthase	Fragaria x ananassa	51872683	1 (3)	18.14	2.19	
1400	chalcone synthase	Fragaria x ananassa	71979904	8 (30)	120.67	3.37	
1425	chalcone synthase	Fragaria x ananassa	71979904	11 (33)	162.33	3.14	
1448	chalcone synthase	Fragaria x ananassa	71979904	5 (16)	82.3	2.95	
1462	chalcone synthase	Fragaria x ananassa	71979904	5 (18)	68.65	4.63	
1504	unnamed protein product	Fragaria x ananassa	10187153	8 (25)	84.53	1.93	
1536	Alcohol dehydrogenase	Fragaria x ananassa	113436	4 (14)	51.55	2.33	
1539	cytosolic aldolase	Fragaria x ananassa	10645188	2 (11)	27.68	1.50	

1700	flavanone 3-hydroxylase	Fragaria x ananassa	51493451	7 (23)	110.48	3.61	
1701	O-methyltransferase ripening-induced protein major storage protein	Fragaria x ananassa	6760443	11 (36)	144.81	3.51	
		Fragaria vesca	2465008	6 (21)	73.49		
		Fragaria x ananassa	54306591	3 (15)	52.12		
1706	ripening-induced protein O-methyltransferase major storage protein	Fragaria vesca	2465008	15 (53)	236.19	2.28	
		Fragaria x ananassa	6760443	9 (32)	143.88		
		Fragaria x ananassa	54306591	3 (27)	49.58		
1707	flavanone 3-hydroxylase O-methyltransferase	Fragaria x ananassa	51493451	12 (39)	183.55	3.97	
		Fragaria x ananassa	6760443	10 (32)	163.03		
1722	Glutamine synthetase GS1 Glutamate--ammonia ligase	Fragaria x ananassa	CO817064	4 (15)	55.25	1.60	
		Fragaria x ananassa	CO817764	3 (14)	45.49		
1725	quinone oxidoreductase	Fragaria x ananassa	29468088	5 (23)	75.39	1.60	
1759	quinone oxidoreductase	Frgaria x ananassa	29468088	20 (62)	315.87	5.60	
1766	putative quinone oxidoreductase	Fragaria x ananassa	15808674	29 (76)	490.07	7.58	
1812	Prunin precursor	Fragaria x ananassa	AI795142	3 (15)	65.43	1.77	
1824	unknown	Fragaria x ananassa	CO379584	2 (10)	27.27	2.87	
1828	Putative epoxide hydrolase	fragaria vesca	CX661489	8 (32)	121.94	2.04	
2041	isoflavone reductase related protein	Pyrus communis	3243234	5 (10)	85.76	2.24	
2108	isoflavone reductase related protein	Pyrus communis	3243234	10 (4)	65.6	2.44	
2156	Annexin-like protein RJ4	Fragaria x ananassa	21264397	22 (48)	334.05	1.77	

2617	cytosolic ascorbate peroxidase In2-1 protein	Zea mays Fragaria x ananassa	600116 CO817712	2 (6) 2 (8)	26.41 33.33	2.31	
2687	chalcone isomerase	Fragaria x ananassa	71979900	7 (39)	129.62	1.85	
2730	chalcone isomerase	Fragaria x ananassa	71979900	10 (48)	153.57	2.23	
3028	expansin	Fragaria x ananassa	6646885	2 (12)	33.74	2.02	
3062	ARD-like protein	Fragaria vesca	DY674064	5 (25)	84.81	1.61	
3164	unknown ATP synthase delta chain	Fragaria x ananassa Fragaria x ananassa	CO380639 TA608_3747	2 (25) 2 (11)	34.89 27.37	1.53	
3222	methionine sulfoxide reductase	Fragaria x ananassa	31455408	6 (28)	102.94	3.22	
3371	TCTP protein	Fragaria x ananassa	1922278	6 (25)	77.28	1.58	
3635	Protein At3g04780 Temperature-induced lipocalin	Fragaria vesca Fragaria vesca	TA3330_57918 TA2100_57918	3 (24) 2 (13)	58.67 25.84	1.54	
3749	eIF5A	Rosa chinensis	121281798	6 (39)	103.23	1.71	

b) Proteine down-regulate in R vs. I

Spot	Proteina	Organismo	Numero d'accesso	Peptidi / coverage	Spectrum Mill score	Fold change	Andamento durante la maturazione I>T>R (ANOVA $p < 0.05$)
831	Protein disulfide-isomerase prec.	Fragaria vesca	TA2477_57918	6 (12)	92.78	-1.60	
834	Protein disulfide-isomerase prec. RuBisCO large subunit-binding prot.	Fragaria vesca Fragaria vesca	TA2477_57918 TA2694_57918	13 (29) 3 (16)	218.99 51.89	-1.55	
1970	Putative aldolase	Fragaria vesca	TA2554_57918	2 (9)	25.81	-2.01	

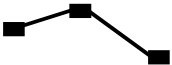

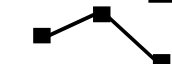
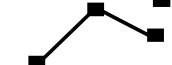
2040	aldo/keto reductase	Fragaria x ananassa	53988164	1 (3)	22.18	-1.67	
	Annexin	Fragaria x ananassa	CO817235	5 (23)	58.49		
2060	unknown	Malus x domestica	DR994225	1 (66)	15.72	-1.64	
	Triosephosphate isomerase	Fragaria x ananassa	CO816817	1 (94)	13.23		
2066	pyridoxin biosynthesis PDX1-like	Fragaria vesca	DY670982	2 (10)	40.52	-1.99	
2115	aldo/keto reductase	Fragaria x ananassa	53988164	2 (6)	31.14	-1.53	
2212	F10A5.19 (similar to histone H3.2)	Arabidopsis thaliana	9369379	1 (27)	22.22	-1.94	
2215	unknown	Fragaria x ananassa	51046757	1 (20)	13.47	-1.62	
	aldo/keto reductase	Fragaria x ananassa	53988164	6 (20)	83.77		
3050	triosephosphate isomerase	Fragaria x ananassa	7650502	2 (8)	30.01	-1.94	
3637	CBS1	Fragaria x ananassa	CO817074	6 (13)	98.6	-1.56	

c) Proteine up-regulate in R vs. T





Spot	Proteina	Organismo	Numero d'accesso	Peptidi / coverage	Spectrum Mill score	Fold change	Andamento durante la maturazione I>T>R (ANOVA $p < 0.05$)
589	Heat shock protein Hsp70	Fragaria x ananassa	CO816704	2 (18)	39.84	1.60	
1400	chalcone synthase	Fragaria x ananassa	71979904	8 (30)	120.67	1.85	
1425	chalcone synthase	Fragaria x ananassa	71979904	11 (33)	162.33	1.98	
1448	chalcone synthase	Fragaria x ananassa	71979904	5 (16)	82.3	1.58	
1462	chalcone synthase	Fragaria x ananassa	71979904	5 (18)	68.65	1.76	
1536	Alcohol dehydrogenase	Fragaria x ananassa	113436	4 (14)	51.55	1.65	
1700	flavanone 3-hydroxylase	Fragaria x ananassa	5943451	7 (23)	110.48	2.01	


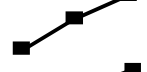










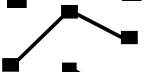
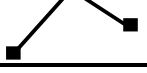


1701	O-methyltransferase ripening-induced protein major storage protein	Fragaria x ananassa Fragaria vesca Fragaria x ananassa	6760443 2465008 54306591	11 (36) 6 (21) 3 (15)	144.81 73.49 52.12	1.65	
1706	ripening-induced protein O-methyltransferase major storage protein	Fragaria vesca Fragaria x ananassa Fragaria x ananassa	2465008 6760443 54306591	15 (53) 9 (32) 3 (27)	236.19 143.88 49.58	1.61	
1707	flavanone 3-hydroxylase O-methyltransferase	Fragaria x ananassa Fragaria x ananassa	51493451 6760443	12 (39) 10 (32)	183.55 163.03	1.90	
1722	Glutamine synthetase Glutamate--ammonia ligase	Fragaria x ananassa Fragaria x ananassa	CO817064 CO817764	4 (15) 3 (14)	55.25 45.49	1.62	
1759	quinone oxidoreductase	Frgaria x ananassa	29468088	20 (62)	315.87	3.81	
1766	putative quinone oxidoreductase	Fragaria x ananassa	15808674	29 (76)	490.07	3.18	
1967	Chloroplast latex aldolase-like prot.	Fragaria x ananassa	TA616_3747	6 (30)	82.21	2.01	
2156	Annexin-like protein RJ4	Fragaria x ananassa	21264397	22 (48)	334.05	1.80	
2589	In2-1 protein	Fragaria x ananassa	CO817712	6 (24)	91.99	2.33	
2969	F3F9.11 (similar to glut. S-transf.)	Fragaria x ananassa	TA1997_57918	7 (29)	102.5	2.24	
3222	methionine sulfoxide reductase	Fragaria x ananassa	31455408	6 (28)	102.94	2.10	
3264	TCTP protein	Fragaria x ananassa	1922278	8 (30)	102.01	1.62	
3507	Major strawberry allergen Fra a 1-B	Fragaria x ananassa	90185680	5 (41)	63.76	2.30	

d) Proteine down-regulate in R vs. T

Spot	Proteina	Organismo	Numero d'accesso	Peptidi / coverage	Spectrum Mill score	Fold change	Andamento durante la maturazione I>T>R (ANOVA $p < 0.05$)
1476	Alcohol dehydrogenase	Fragaria x ananassa	113436	4 (15)	53.77	-1.94	
2369	GLX2-2 (GLYOXALASE 2-2)	Fragaria vesca	NM111922	6 (27)	107.91	-1.72	
3350	Heat shock protein 22	Fragaria vesca	TA1968_57918	7 (21)	101.36	-2.19	
	60S ribosomal protein L21	Fragaria vesca	TA2283_57918	6 (23)	85.51		
3636	Major strawberry allergen Fra a 1-E	Fragaria x ananassa	90185690	13 (74)	248.37	-1.59	
3791	LMW heat shock protein	Fragaria x ananassa	2911276	12 (49)	204.72	-1.97	

e) Proteine up-regulate in T vs. I

Spot	Proteina	Organismo	Numero d'accesso	Peptidi / coverage	Spectrum Mill score	Fold change	Andamento durante la maturazione I>T>R (ANOVA $p < 0.05$)
321	unknown	Fragaria x ananassa	CO816798	3 (10)	52.93	1.64	
486	high molecular weight HSP	Malus x domestica	6969976	17 (29)	307.68	1.59	
541	Prunin precursor	Fragaria x ananassa	AI795142	2 (21)	27.9	2.07	
836	heat shock protein 60	Prunus dulcis	24637539	3 (6)	36.77	1.96	
1372	RuBisCO activase	Malus x domestica	3914605	9 (21)	142.05	1.90	
1382	RuBisCO activase	Malus x domestica	3914605	10 (21)	159.38	1.56	

1400	chalcone synthase	Fragaria x ananassa	71979904	8 (30)	120.67	1.83	
1448	chalcone synthase	Fragaria x ananassa	71979904	5 (16)	82.3	1.86	
1462	chalcone synthase	Fragaria x ananassa	71979904	5 (18)	68.65	2.62	
1654	high molecular weight HSP	Malus x domestica	6969976	2 (4)	27.91	1.81	
1700	flavanone 3-hydroxylase	Fragaria x ananassa	51493451	7 (23)	110.48	1.79	
1766	putative quinone oxidoreductase	Fragaria x ananassa	15808674	29 (76)	490.07	2.38	
1824	unknown	Fragaria x ananassa	CO379584	2 (10)	27.27	2.17	
1828	Putative epoxide hydrolase	fragaria vesca	CX661489	8 (32)	121.94	1.94	
2041	isoflavone reductase related protein	Pyrus communis	3243234	5 (10)	85.76	1.64	
2108	isoflavone reductase related protein	Pyrus communis	3243234	10 (4)	65.6	1.85	
2687	chalcone isomerase	Fragaria x ananassa	71979900	7 (39)	129.62	1.51	
2730	chalcone isomerase	Fragaria x ananassa	71979900	10 (48)	153.57	1.90	
3350	Heat shock protein 22 60S ribosomal protein L21	Fragaria vesca Fragaria vesca	TA1968_57918 TA2283_57918	7 (21) 6 (23)	101.36 85.51	2.03	
3742	LMW heat shock protein	Fragaria x ananassa	2911276	5 (13)	62.23	1.85	
3791	LMW heat shock protein	Fragaria x ananassa	2911276	12 (49)	204.72	3.59	
3897	LMW heat shock protein	Fragaria x ananassa	2911276	12 (48)	189.86	2.10	

f) Proteine down-regulate in T vs. I

Spot	Proteina	Organismo	Numero d'accesso	Peptidi / coverage	Spectrum Mill score	Fold change	Andamento durante la maturazione I>T>R (ANOVA $p < 0.05$)
2040	aldo/keto reductase	Fragaria x ananassa	53988164	1 (3)	22.18	-1.57	
	Annexin	Fragaria x ananassa	CO817235	5 (23)	58.49		
2212	F10A5.19 (similar to histone H3.2)	Arabidopsis thaliana	9369379	1 (27)	22.22	-1.80	
3736	Peroxiredoxin-like protein	Fragaria x ananassa	CO817516	6 (27)	113.21	-1.53	
	MLP-like protein 423	Fragaria x ananassa	CO818041	7 (40)	109.8		

3.4 Variazioni del proteoma nel confronto tra genotipi

Al fine di correlare le variazioni del proteoma con l'espressione di tratti qualitativi nel frutto, l'analisi DIGE è stata applicata allo studio delle proteine differenzialmente espresse nel genotipo d'élite, QE, rispetto ai genotipi parentali, M ed U. Lo studio è stato condotto confrontando frutti allo stadio rosso, poiché, a questo livello, l'accumulo differenziale di proteine dovrebbe corrispondere a maggiori differenze nei principali tratti qualitativi, quali aroma e consistenza. Le mappe DIGE ottenute dai 3 genotipi hanno rilevato, approssimativamente, 2000 spot risolti per gel. Il 25% degli spot ha mostrato un coefficiente di correlazione inferiore al 20%. Sono stati identificati, complessivamente, 21 spot proteici accumulati in maniera differenziale tra QE ed i suoi parentali (Tabella 3.5). Nello specifico, il confronto M vs. QE ha fornito 11 spot up-regolati e 6 spot down-regolati in M, mentre il confronto U vs. QE ha evidenziato 11 spot up-regolati e 3 spot down-regolati in U (Tabella 3.5).

Tabella 3.5. Proteine differenzialmente regolate nel confronto tra genotipi di fragola

a) Proteine regolate in USB 35 vs. QE

Spot	Proteina	Organismo	Numero accesso	Peptidi / coverage %	Spectrum Mill score	Fold change
84	porin	Fragaria x ananassa	51049027	13 (2)	38.79	1.93
937	put. quinone oxidoreductase (FaQR)	Fragaria x ananassa	15808674	66 (18)	297.06	2.36
992	glyceraldehyde 3-phosphate dehydr.	Fragaria x ananassa	16226117	46 (5)	74.53	1.76
993	glyceraldehyde 3-phosphate dehydr.	Fragaria x ananassa	16226117	50 (4)	63.74	2.3
1007	glyceraldehyde 3-phosphate dehydr.	Fragaria x ananassa	16226117	45 (4)	53.01	2.25
1016	hypothetical protein	Fragaria vesca	89554457	16 (1)	14.17	2.06
1155	aldo/keto reductase	Fragaria x ananassa	53988164	48 (11)	131.08	2.19
1196	isoflavone reductase related protein	Pyrus communis	3243234	3 (2)	39.21	-2.63
1540	Allergen Fra a 1	Fragaria x ananassa	90185682	7 (1)	21.24	-2.14
1541	fra a 1 allergen	Fragaria x ananassa	74197562	7 (1)	20.52	-1.8
1562	LMW heat shock protein	Fragaria x ananassa	2911276	18 (2)	43.43	2.57
1570	LMW heat shock protein	Fragaria x ananassa	2911276	18 (2)	38.13	1.93
1592	Prunin 1 precursor	Fragaria x ananassa	14272197	11 (1)	19.24	1.53
1595	put. glycine-rich RNA-binding prot.	Fragaria x ananassa	51048034	11 (2)	35.6	2.44

b) Proteine regolate in Miss vs. QE

Spot	Proteina	Organismo	Numero accesso	Peptidi / coverage %	Spectrum Mill score	Fold change
578	D-3-phosphoglycerate dehydr.	Prunus persica	89490348	5 (1)	17.26	-1.87
666	F1-ATPase alpha subunit	Rubus sp. JPM-2004	57115599	5 (2)	36.46	-1.56
808	cytosolic phosphoglucomutase	Fragaria vesca	89543152	31(1)	14.75	-1.7
937	put. quinone oxidoreductase (FaQR)	Fragaria x ananassa	15808674	66 (18)	297.06	1.94
985	O-methyltransferase	Fragaria x ananassa	6760443	3 (1)	14.81	2.17
992	glyceraldehyde 3-phosphate dehydr.	Fragaria x ananassa	16226117	46 (5)	74.53	5.86
993	glyceraldehyde 3-phosphate dehydr.	Fragaria x ananassa	16226117	50 (4)	63.74	8.12
1016	hypothetical protein	Fragaria vesca	89554457	16 (1)	14.17	6.53
1050	cytosolic malate dehydrogenase	Malus x domestica	78216493	10 (2)	37.26	-1.52
1155	aldo/keto reductase	Fragaria x ananassa	53988164	48 (11)	131.08	2.01
1196	isoflavone reductase related protein	Pyrus communis	3243234	3 (2)	39.21	1.71
1215	isoflavone reductase related protein	Fragaria vesca	57896334	18 (4)	63.55	-8.22
1216	isoflavone reductase related protein	Fragaria vesca	57896334	39 (7)	128.66	-6.46
1562	LMW heat shock protein	Fragaria x ananassa	2911276	18 (2)	43.43	3.65
1570	LMW heat shock protein	Fragaria x ananassa	2911276	18 (2)	38.13	2.37
1592	Prunin 1 precursor	Fragaria x ananassa	14272197	11 (1)	19.24	3.58
1595	put. glycine-rich RNA-binding prot.	Fragaria x ananassa	51048034	11 (2)	35.6	2.49

4 DISCUSSIONE

4.1 La mappa elettroforetica bidimensionale del frutto di fragola

La mappa elettroforetica bidimensionale, mostrata in Figura 3.1, rappresenta la fase iniziale dell'analisi descrittiva del proteoma espresso nel frutto di fragola. Un set di proteine sempre espresse durante la maturazione è stato identificato mediante *matching* delle immagini dei gel 2-D, relativi a 3 diversi stadi di sviluppo del frutto (immaturo, turning e rosso) ed analisi di spettrometria di massa tandem degli spot che mostrano, nei gel sovrapposti, uguali coordinate.

La classe di proteine più rappresentata sulla mappa (il 19% delle identificazioni totali) è legata al metabolismo secondario (Tabella 3.1). In particolare, sono stati identificati i primi 5 enzimi dalla *pathway* biosintetica dei flavonoidi. Si tratta della chalcone sintasi (CHS), la chalcone isomerasi (CHI), la flavanone 3-idrossilasi (F3H), la diidroflavonol 4-reduttasi (DFR) e l'antocianidina sintasi (ANS). Ad oggi, sono stati descritti più di 6000 differenti flavonoidi (Harbone and Williams, 2000). In natura, essi svolgono una vasta gamma di funzioni: sono impegnati nella pigmentazione di frutti e dei fiori per attrarre gli impollinatori e favorire la dispersione dei semi; proteggono contro la luce ultravioletta; proteggono le piante dall'attacco di microrganismi patogeni; sono impegnate nella germinazione del polline; agiscono come molecole segnale nell'interazione pianta-microorganismo (Dixon and Paiva, 1995; Dooner et al., 1991). CHS è l'enzima che inizia la via biosintetica dei flavonoidi, favorendo la condensazione di tre unità di malonil-CoA con il p-cumaroil-CoA a dare il 4,2', 4', 6' tetraidrossichalcone (Tanaka et al., 1998). CHS rappresenta un *target* attraente per l'ingegneria genetica, poiché esistono numerosi esempi di co-soppressione e *down-regolazione* del relativo gene, che permettono di bloccare totalmente tale via metabolica (Jorgensen et al, 1996). Tuttavia, le piante non accumulano chalconi. Dopo la loro

formazione, essi sono rapidamente isomerizzati da CHI in flavanoni. Mutanti di CHI favoriscono l'accumulo di chalconi, evidenziato dal prevalente colore giallo dei fiori (Forkmann and Heller, 1999). I flavanoni vengono, successivamente, idrossilati in posizione C-3, così come è stato dimostrato in numerose organismi, tra cui Petunia, pomodoro e mais. La reazione è condotta da F3H, membro della famiglia delle diossigenasi dipendenti da 2-oxoglutarato, ampiamente conservata in specie tra loro divergenti (Britsch et al., 1993). Il prodotto della reazione è il diidrokaempferolo, che può essere idrossilato in diverse posizioni, portando, da un lato, alla diidroquercetina, responsabile della produzione di cianidine; dall'altro, alla diidromiricetina, responsabile della produzione di antocianine (Forkmann, 1991; Toda et al., 2002). DFR catalizza, invece, la riduzione stereo-specifica dei diidroflavonoli in leucoantocianidine. Queste ultime sono i precursori immediati della sintesi di antocianidine, mediata da ANS. Inoltre, sono anche i precursori della sintesi di catechine e proantocianidine (PA), coinvolte nella resistenza delle piante ai patogeni, nonché composti ad elevato valore nutraceutico per l'alimentazione umana.

Ciascun enzima chiave, associato alla biosintesi dei flavonoidi, è stato rinvenuto, nella mappa 2-D elaborata nel corso del presente lavoro, in corrispondenza di almeno due *spot* distinti. L'identificazione di una stessa proteina, in diversi *spot*, è generalmente indicativa dell'esistenza di varie isoforme per quella data proteina (Sappl et al., 2004). Tali isoforme possono originare da modifiche post-traduzionali, varianti di *splicing*, degradazione proteica o variazioni alleliche. La presenza di *spot* multipli è un fenomeno molto frequente nelle piante poliploidi, come dimostrano i dati disponibili in letteratura per mela ed orzo (Son et al., 1999; Holm et al., 2001; Laugesen et al., 2003). La correlazione tra *spot* multipli ed isoforme proteiche è stata dimostrata anche in fragola, (Hierno et al., 2006), in uno studio comparativo tra varietà bianche e rosse. Per quanto

interessante, non è tuttavia possibile sapere se le isoforme, identificate nel corso del presente lavoro, corrispondano effettivamente a prodotti di differenti geni o a modifiche post-traduzionali dello stesso prodotto genico. Ciononostante, l'abbondanza di identificazioni di proteine coinvolte nella biosintesi dei flavonoidi riflette il ruolo fondamentale di questa *pathway* durante la maturazione (Almeida et al., 2007).

Il processo di maturazione dei frutti è stato descritto come un processo ossidativo controllato, per cui l'accumulo di H₂O₂ e ROS è bilanciato dall'attività di sistemi antiossidanti, i cui principali attori sono la superossido dismutasi (SOD) e i componenti de ciclo ascorbato-glutatione (Jimenez et al., 2002). Molto rappresentate nella mappa sono, infatti, le proteine associate allo stress. Esse rappresentano il 17% delle identificazioni totali (Tabella 3.1). L'ascorbato perossidasi è l'enzima responsabile della rimozione di H₂O₂ da mitocondri e cloroplasti vegetali (Chew et al., 2003). SOD è il maggiore componente del circuito antiossidante nella cellula ed è stato riportato un aumento della sua attività durante la maturazione del frutto (Jimenez et al., 2002). La glutatione-S-trasferasi è impegnata nella detossificazione dagli xenobiotici così come nella riduzione degli idroperossidi organici in alcoli, limitando il danno ossidativo (Gong et al., 2005), mentre la lactoilglutatione liasi è impegnata nella detossificazione dal metilgliossale, che si forma prevalentemente come bioprodotto del metabolismo dei lipidi e carboidrati (Singla-Pareek et al., 2003).

Altre proteine generalmente correlate al processo di maturazione e/o alla risposta allo stress sono le *heat shock proteins* (HSP), la cui espressione in pianta è modulata da un'ampia varietà di stress, tra cui lo stress ossidativo (Wang et al., 2004). È stato dimostrato che alcune isoforme specifiche sono differenzialmente espresse durante la maturazione (Waters, 1995; Boston et al., 1996). Evidenze suggeriscono che le HSP a basso peso molecolare proteggono la cellula dallo stress ossidativo, mediante un

meccanismo legato alla loro capacità di ridurre i livelli intracellulari di ROS in maniera dipendente dal glutatione (Arrigo et al., 1998; Sun et al., 2002).

Le proteine coinvolte, invece, nella produzione d'energia rappresentano il 17% delle proteine totali (Tabella 3.1). Questo gruppo è dominato dagli enzimi della glicolisi e del ciclo degli acidi carbossilici, che forniscono i metaboliti energetici necessari per il processo di maturazione in corso. Essi alimentano, a partire dagli zuccheri, substrati per la respirazione e la sintesi di acidi organici e pigmenti (Carrari et al., 2006). Il presente studio ha rilevato la presenza di: aldolasi, triosofosfato isomerasi, enolasi, fosfoglicerato chinasi, gliceraldeide 3-fosfato deidrogenasi, fosfogliceromutasi, fruttochinasi, citrato sintasi e malato deidrogenasi.

È stata rilevata, inoltre, la presenza di allergeni e di diverse proteine coinvolte nel metabolismo proteico, dei nucleotidi, degli aminoacidi, nella proteolisi, nel trasporto, nella sintesi proteica, nella comunicazione ed organizzazione cellulare (Tabella 3.1).

4.2 Analisi shotgun: identificazione proteica su larga scala

L'approccio *shotgun*, per l'identificazione del proteoma espresso nel frutto di fragola durante la maturazione, ha fornito un set di dati molto vasto (Tabella 3.2), che include tutte le proteine già descritte nella mappa 2-D. A causa del ruolo che il processo di maturazione svolge nel determinare la qualità complessiva del frutto, l'attenzione della presente discussione è stata focalizzata sulle proteine e sulle *pathway* associate ai cambiamenti nel sapore, colore, profumo e consistenza dei frutti.

L'aroma della fragola è fortemente influenzata dalla presenza di acidi organici e zuccheri. Il contenuto totale degli zuccheri aumenta progressivamente durante la maturazione e stimola fortemente la formazione di metaboliti secondari, quali antocianine e furanoni, mentre il contenuto in acidi, in particolare acido citrico e malico,

diminuisce (Kafkas et al., 2007). Il principale acido organico in fragola è l'acido citrico. Esso contribuisce alle principali qualità organolettiche del frutto maturo. Nel corso del presente studio, è stata identificata buona parte degli enzimi coinvolti nel ciclo dell'acido citrico. In particolare, sono stati identificati i seguenti enzimi: citrato sintasi, aconitasi, isocitrato deidrogenasi, il complesso della 2-oxoglutarato deidrogenasi (subunità E1, E2, E3), succinil-CoA sintetasi, succinato deidrogenasi, fumarasi e malato deidrogenasi. L'acido citrico contribuisce in maniera preponderante all'acidità titolabile del frutto, che decrementa durante lo sviluppo (Hancock et al., 1999). Il citrato e l'isocitrato agiscono come intermedi in una vasta serie di reazioni biochimiche. I dati pubblicati in letteratura suggeriscono che, durante il declino degli acidi, il citrato possa essere utilizzato per la produzione di aminoacidi e acidi grassi (Popova et al., 1998). Il ciclo dell'acido citrico è influenzato in maniera significativa dall'attività della NAD-isocitrato deidrogenasi, mentre la NADP-isocitrato deidrogenasi partecipa, invece, alle biosintesi che partono dal citrato (Popova et al., 1998). Entrambi gli enzimi sopraelencati sono stati identificati nel corso della presente analisi proteomica su larga scala, supportando la nozione di conversione del citrato durante lo sviluppo del frutto. Il prodotto della isocitrato deidrogenasi, il 2-oxoglutarato, oltre ad essere il principale accettore di azoto, funge da precursore per la produzione di aminoacidi, ad esempio di glutammato nella bacca di pomodoro. L'accumulo di aminoacidi liberi è parte integrante del processo di maturazione dei frutti (Boggio et al., 2000; Gallardo et al., 1995). In questo contesto, l'identificazione, nel corso dell'analisi *shotgun*, di diversi enzimi coinvolti nel metabolismo del glutammato (glutammato deidrogenasi, glutammato decarbossilasi, gamma-amminobutirrato transamminasi) risulta di notevole valore.

In maniera del tutto simile, il prodotto della malato deidrogenasi, l'ossalacetato, è il composto di partenza per la biosintesi dell'acido aspartico, catalizzata dall'aspartato aminotrasferasi, enzima identificato nel corso della presente analisi. Gli aminoacidi liberi, in particolare il glutammato e l'acido aspartico, sono composti essenziali, non-volatili, responsabili del gusto complessivo di molti alimenti di origine vegetale (Boggio et al., 2000; Fuke and Konosu, 1991).

Nelle cellule vegetali, la conversione del citrato e dell'isocitrato può essere associata al metabolismo degli acidi grassi. L'analisi proteomica condotta ha identificato l'ATP-citrato liasi, coinvolto nel metabolismo dell'acetil-CoA, l'acetil-CoA acetiltrasferasi e l'HMG-CoA sintasi, coinvolti, invece, nella sintesi di HMG-CoA. L'acetil-CoA e l'HMG-CoA sono, rispettivamente, precursori del mevalonato e della *pathway* degli isoprenoidi (Omura et al, 2007). Sono stati identificati, inoltre, l'acetil-CoA decarbossilasi, la chalcone sintasi, la chalcone isomerasi, coinvolti nei primi *step* della biosintesi dei flavonoidi (D'Amico and Perrotta, 2005).

L'aroma della fragola è determinato da un bilancio tra le quantità totali di acidi e di zuccheri presenti nel frutto. La biosintesi del saccarosio è catalizzata dall'azione combinata di UTP-glucose-1-fosfato uridiltrasferasi, saccarosio fosfato sintasi e saccarosio fosfato fosfatasi (Nguyen and Foyer, 2001). Tutti questi enzimi, tranne la saccarosio fosfatasi, sono stati identificati nel corso dell'analisi *shotgun*, che ha permesso di identificare, inoltre, anche la saccarosio sintasi, enzima fondamentale per la biosintesi di saccarosio, esosi e amidi (Smith, 1999; Winter and Huber, 2000). La saccarosio sintasi catalizza, infatti, la conversione reversibile del saccarosio in fruttosio ed UDP-glucosio, metabolica necessario per la biogenesi della parete cellulare, coinvolto anche nella sintesi dei carboidrati presenti nei glicolipidi e nelle glicoproteine (Koch, 2004). Saccarosio, fruttosio e glucosio sono i composti solubili più abbondanti

in fragola. Questi zuccheri sono usati, innanzitutto, come fonte d'energia per sostenere le variazioni metaboliche durante la maturazione, ma anche come precursori dei composti aromatici, in particolare i furanoni (Wein et al., 2001). Tra le sostanze aromatiche presenti in fragola, il furaneolo (HDMF, 2,5-dimetil-4-idrossi-2H-furan-3-one) è il più importante, a causa della sua alta concentrazione e bassa soglia di odore (Schwab and Roscher, 1997). Le prime indicazioni della sintesi enzimatica dell'HDMF in fragola furono fornite da uno studio che dimostrava la correlazione tra maturazione e concentrazione di HDMF (Sanz et al., 1995). Esperimenti con precursori radiomarcati dimostrarono, inoltre, che il D-fruttosio 1,6-difosfato è un efficiente precursore per la biosintesi di HDMF (Roscher et al., 1998; Sshwab, 1998). Sebbene la via metabolica che porta alla sua formazione dal D-fruttosio 1,6-difosfato rimane ancora oggi sconosciuta, tutti gli studi condotti hanno dimostrato che il furaneolo deriva dal metabolismo degli zuccheri (Wein et al, 2001). Il precursore naturale dell'HDMF, HMMF (4-idrossi-5-metil-2-metilene-3(2H)-furanone), è convertito in HDMF da un'ossidoreduttasi, isolata recentemente nel frutto di fragola (Raab et al.,2006). Quest'enzima, chiamato FaQR (*Fragaria x ananassa* quinone ossidoreduttasi) è l'ultimo enzima della *pathway* dall'HDMF. Nel corso del presente lavoro, esso è stato identificato con 23 peptidi ed una copertura di sequenza pari al 78%, rappresentando una tra le proteine meglio caratterizzate dall'approccio *shotgun*. Il prodotto dell'attività enzimatica di FaQR, HDMF, si trasforma in DMMF (2,5-dimetil-4-metossi-3(2H)-furanone), se metilato dall'enzima FaOMT (*Fragaria x ananassa* O-metiltrasferasi) durante la maturazione. Anche quest'ultimo enzima è stato identificato, nel corso della nostra indagine proteomica, tra le proteine meglio caratterizzate, poiché identificato con 21 peptidi ed una copertura di sequenza molto alta (76%). L'alto livello di confidenza per l'identificazione di FaQR e FaOMT riflette l'abbondanza di queste proteine ed il

loro ruolo chiave in relazione alla biosintesi delle sostanze aromatiche durante la maturazione.

L'identificazione proteica su larga scala ha rilevato, inoltre, 41 proteine contemporaneamente presenti in tutti e tre gli stadi di maturazione, in entrambe le metodologie di analisi *shotgun* (Tabella 3.3). Tali proteine sono coinvolte essenzialmente nell'organizzazione e comunicazione cellulare (actina, alfa-tubulina, etc) oppure nel metabolismo energetico. Come atteso, la maggior parte di esse corrisponde a proteine *housekeeping*, che svolgono funzioni associate al metabolismo basale.

4.3 Proteine coinvolte nel processo di maturazione

Le proteine identificate come differenzialmente espresse durante la maturazione (Tabella 3.4) sono coinvolte in numerosi processi biologici, come il metabolismo del carbonio e la produzione d'energia, il metabolismo secondario e la biosintesi dei componenti cellulari, l'organizzazione cellulare, la comunicazione e la traduzione del segnale, il metabolismo proteico, la risposta allo stress e la trascrizione genica. Molte delle proteine differenzialmente espresse mostrano un regolare incremento nel volume degli spot man mano che si passa dallo stadio bianco a quello rosso maturo (Tabella 3.4), indicando il loro coinvolgimento specifico nel processo di maturazione. Tuttavia, la presenza, in alcuni casi, di diverse proteine all'interno dello stesso *spot* (Tabella 3.4) rende difficile stabilire quale di queste proteine sia effettivamente accumulata in maniera differenziale.

La maturazione della fragola è un processo di sviluppo geneticamente programmato che determina le caratteristiche qualitative e nutraceutiche del frutto. Essa è caratterizzata da cambiamenti nel colore, i pigmenti esistenti, infatti, vengono degradati per lasciare

spazio alla sintesi di nuove molecole, spesso più intensamente colorate; dalla produzione di composti aromatici che influenzano la palatabilità della fragola; da una progressiva perdita di consistenza del frutto, come risultato di una serie di processi degradativi a carico della parete cellulare, generalmente indicati con il termine *softening*. Lo stadio finale della maturazione si sovrappone, di fatto, con la senescenza del frutto (Perkins-Veazie, 1995); pertanto, la maturazione può essere definita come un processo continuo di sviluppo in cui differenti fasi fisiologiche tendono a sovrapporsi (Moing et al., 2001).

Prima dell'inizio dei processi di ammorbidimento della polpa, si assiste ad una fase di crescita del frutto, mediata dall'espansione cellulare, che richiede l'elongazione della parete cellulare e l'accumulo di soluti all'interno dei vacuoli. I risultati dell'analisi DIGE dimostrano l'espressione differenziale della proteina RJ14, accumulata in eccesso nello stadio R rispetto ad I e T (Tabella 3.4a,c). La proteina RJ14 appartiene alla famiglia delle annexine, molecole coinvolte in una moltitudine di processi fisiologici, tra cui l'espansione cellulare. L'accumulo differenziale, evidenziato dall'analisi proteomica, concorda con i dati reperibili in letteratura, che dimostrano, tramite studi d'espressione genica, un incremento nell'espressione del gene per l'annexina, durante la maturazione del frutto in fragola e peperone (Wilkinson et al., 1995; Proust et al., 1996). E' stato postulato che le annexine possano influenzare i flussi ionici attraverso le membrane, così come l'ancoraggio del citoscheletro alla membrana, attraverso un legame, Ca^{2+} -dipendente, con i fosfolipidi (Konopka-Postupolska, 2007). Le annexine sembrano essere coinvolte anche nei processi di esocitosi degli enzimi degradativi della parete. Esse agiscono sequestrando gli ioni Ca^{2+} rilasciati durante la degradazione della parete stessa. Esse sono proteine Ca^{2+} -dipendenti, poiché lo ione calcio è implicato in una vasta gamma di processi biologici, le annexine potrebbero servire anche come

molecole segnale per regolare i diversi processi fisiologici che si sovrappongono durante la maturazione. Ciò spiegherebbe l'identificazione di una seconda annexina, *down-regolata* in R vs. I e T (Tabella 3.4b,f). Le annexine appartengono ad una grande famiglia di proteine multifunzionali, pertanto è plausibile che il loro pattern di espressione possa essere regolato in maniera indipendente.

Il *softening* del frutto è un fenomeno complesso, che contribuisce a determinare la qualità della fragola matura. Tale processo richiede l'espressione e l'azione coordinata di diversi enzimi e molecole proteiche, capaci di promuovere eventi degradativi e/o di rilassamento (*loosening*) della struttura che compone la parete cellulare (Schwab and Raab, 2004). Dati presenti in letteratura dimostrano il coinvolgimento delle endopoligalatturonasi (endoPG), implicate nella degradazione della frazione pectica delle pareti cellulari. L'attività delle endoPG aumenta durante la maturazione dei frutti a polpa molle in seguito ad un incremento dell'attività trascrizionale del relativo gene (Pressey et al., 1978; Lester et al., 1994). L'azione delle endoPG appare necessaria, sia pur non sufficiente, da sola, per i processi d'ammorbidimento del frutto (Hadfield et al., 1998). Ciononostante, l'analisi proteomica da noi condotta non ha evidenziato tali enzimi tra i differenzialmente espressi, sostenendo l'ipotesi di Schwab e Raab (2004) che l'attività enzimatica delle endoPG sia sotto i limiti di detection nel frutto di fragola in maturazione. Altre proteine, tra le quali le espansine, possono essere implicate nel processo di *softening*. Queste proteine, infatti, rompendo i legami idrogeno tra i componenti polisaccaridici, determinano il rilassamento della struttura delle pareti, facilitando l'accesso degli enzimi idrolitici ai loro substrati (Li et al., 2003). Diversi geni codificanti per le espansine sono stati recentemente identificati, ed i loro profili d'espressione caratterizzati. Tali studi hanno mostrato che tre diverse isoforme di espansine si esprimono nei tessuti del frutto di pesca durante lo sviluppo e la

maturazione; di queste, soltanto la PpExp3, sembrerebbe coinvolta in maniera specifica nell'intenerimento della polpa del frutto (Hayama et al., 2003). Nonostante l'elevato numero di proteine coinvolte nel processo di *softening*, una singola proteina, appartenente alla famiglia delle espansine, è stata identificata come differenzialmente accumulata nello stadio R rispetto allo stadio I (Tabella 3.4a). Questo potrebbe essere legato alle limitazioni intrinseche sia delle tecniche di estrazione che di separazione bidimensionale delle proteine, probabilmente poco adatte a risolvere proteine scarsamente rappresentate e/o idrofobiche, come quelle correlate ai processi di *softening*.

Per contro, un elevato numero di spot proteici, corrispondenti ad enzimi della via biosintetica dei flavonoidi, è stato identificato come differenzialmente espresso nel corso dell'analisi DIGE da noi condotta (Tabella 3.4). In particolare, è stata osservata la presenza degli enzimi: chalcone sintasi (CHS), chalcone isomerasi (CHI), flavanone-3-idrossilasi (FHT) e antocianidina sintasi (ANS) coinvolti nella trasformazione di fenilalanina in chalconi, flavanoni, diidroflavonoli e antocianidine. La comune sovra-espressione in R (Tabella 3.4) riflette il graduale accumulo di flavonoidi durante la maturazione, così come dimostrato da dati di letteratura prodotti nel nostro laboratorio (Almeida et al, 2007). Parecchi di questi enzimi sono stati ritrovati in *spot* multipli, suggerendo l'abbondanza di queste proteine, rispetto al contenuto proteico totale, ed il ruolo basilare della *pathway* biosintetica dei flavonoidi nel determinare la qualità del frutto maturo.

Nel corso della maturazione della fragola, si assiste anche alla produzione di diverse sostanze aromatiche, che determinano cambiamenti nel gusto. Diversi studi hanno dimostrato che l'aroma del frutto è il risultato della percezione coordinata di numerose sostanze (Pyysalo et al., 1979; Larsen et al., 1992) che, nell'insieme, contribuiscono a

determinare il sapore del frutto stesso. Tra tutte le sostanze aromatiche che caratterizzano la fragola, l'elemento principale è rappresentato dal furaneolo (HDMF), la cui *pathway* biosintetica non è ancora completamente nota (Raab et al., 2006). Due enzimi coinvolti in questa via biosintetica, la quinone-ossidoreduttasi (FaQR) e la O-metiltransferasi (FaOMT), sono stati identificati come differenzialmente accumulati nel corso delle analisi DIGE, mostrando un regolare incremento durante la maturazione. In particolare, FaQR ha mostrato il più alto valore d'espressione differenziale (7.58 volte, Tabella 3.4a,c,e), tra gli spot proteici differenzialmente espressi.

Il processo di maturazione è stato anche descritto come un fenomeno ossidativo controllato, per cui l'accumulo di ROS e H₂O₂ è bilanciato dall'attività dei sistemi antiossidanti cellulari, rappresentati da catalasi, superossido dismutasi e componenti del ciclo dell'ascorbato-glutatione. (Jimenez et al., 2002). Il principale enzima responsabile della rimozione del perossido d'idrogeno nel citoplasma, nel mitocondrio e nel cloroplasto delle piante superiori, è l'ascorbato perossidasi (APX). Le analisi condotte sul proteoma di fragola hanno dimostrato un incremento nell'accumulo di APX, passando dallo stadio bianco a quello rosso (Tabella 3.4a). Anche la metionina-solfossido riduttasi è risultata sovra-espressa in R rispetto ad I e T (Tabella 3.4a,c); il suo accumulo differenziale allo stadio maturo riflette la documentata sovra-espressione del corrispondente trascritto genico, durante la maturazione in pomodoro e in fragola (Lincoln et al., 1987; Lopez et al., 2006), concordando anche con l'ipotesi generale che lo stress ossidativo è di fatto responsabile dell'ossidazione delle proteine. La metionina-solfossido riduttasi, nota anche come fruit-ripening protein E4, appartiene ad una classe di molecole coinvolte nei processi difensivi contro i ROS. Di fatto, le proteine ossidate, contenenti residui di metionina-solfossido, possono essere ridotte proprio grazie alla sua azione (Romero et al., 2004).

Nelle piante, anche il ciclo cellulare è controllato dallo stato redox: lo stress ossidativo, infatti, può determinare l'arresto del ciclo cellulare, mediante l'accumulo dei ROS prodotti dalla fotoreduzione dell'ossigeno, durante la fotosintesi (Reichheld et al., 1999). A sostegno di questa tesi, la nostra indagine proteomica ha individuato altri due spot, accumulati in eccesso in R rispetto a I e T, legati allo stress ossidativo e identificati come proteine In2-1 (Tab. 2a,c).

L'accumulo in R del fattore d'iniziazione eucariotico eIF-5A è, invece, probabilmente correlato alla senescenza del frutto, in quanto esso è stato precedentemente descritto come fattore chiave per la traduzione di mRNA richiesti sia per la divisione cellulare che per la morte cellulare (Thompson et al, 2004). In pomodoro, un coinvolgimento di eIF-5a nei processi di senescenza e di morte cellulare programmata è stato di fatto accertato (Wang et al., 2001).

Un'altra proteina di interesse che ha mostrato la tendenza ad accumularsi in R è la proteina TCTP (translationally controlled tumor protein). Studi sul pattern di espressione, spaziale e temporale, del gene corrispondente, condotti in fragola mediante analisi *northern blot*, ne avevano evidenziato un graduale aumento dell'espressione durante la maturazione, con il picco più alto proprio nel frutto rosso maturo (Lopez and Franco, 2006). Anche la proteina TCTP potrebbe essere collegata alla senescenza dei frutti, poiché regolata, nella sua espressione, dal fattore eIF-5E, membro della stessa famiglia del fattore eIF-5A (Gachet et al, 1997).

L'analisi DIGE, condotta sul proteoma del frutto di fragola a diversi stadi di maturazione, ha evidenziato, inoltre, l'accumulo differenziale di parecchie proteine appartenenti alla famiglia delle Hsp (Tabella 3.4). Sebbene tali proteine siano state originariamente descritte come molecole indotte da shock da calore, diversi tipi di stress e/o stimoli possono determinarne l'accumulo durante lo sviluppo della pianta e la

maturazione del frutto (Duck et al, 1989). Esse sono, inoltre, ben conosciute per il loro ruolo nella maturazione di proteine complesse, nella degradazione di peptidi danneggiati o *misfolded* e nella regolazione dell'attività di molte proteine coinvolte nella traduzione del segnale. Una Hsp a basso peso molecolare, identificata, nella nostra indagine proteomica, in più spot proteici sulla mappa DIGE, ha mostrato un accumulo per eccesso in T rispetto ad I, mentre un decremento in R rispetto a T (Tabella 3.4d,e). Questo risultato sembra essere in accordo con uno studio condotto in pomodoro, che riporta il suo effetto sul cambiamento del colore del frutto durante la maturazione. Questa proteina è, infatti, indotta nelle bacche di pomodoro nel corso della transizione dei cloroplasti a cromoplasti ed è espressa allo stadio turning (Neta-Sharir et al., 2005). Tutti questi dati supportano l'ipotesi che le Hsp a basso peso molecolare, oltre al loro effetto protettivo contro vari tipi di stress, possano giocare un ruolo fondamentale nello sviluppo della pianta, anche in condizioni fisiologiche.

Tra gli spot variabilmente accumulati nel corso della maturazione in fragola, sono stati identificati anche alcuni enzimi chiave delle vie glicolisi/gluconeogenesi, cioè: l'aldolasi, e la triosofosfato isomerasi che sono sotto-espressi in frutto rosso. Generalmente i livelli degli enzimi glicolitici decrementano nello stadio finale della maturazione, quando l'induzione dell'accumulo degli esosi inibisce la respirazione e dà il via all'afflusso di malato dal vacuolo e alla sua ossidazione. Infatti, mentre l'amido è la principale forma di riserva durante lo sviluppo del frutto, la maturazione è caratterizzata da una diminuzione nel contenuto in amido, da un richiamo di saccarosio e un accumulo di esosi. Infine, sono stati identificati anche alcuni enzimi coinvolti in trasporto di elettroni/produzione di energia e nell'apparato fotosintetico, e cioè la subunità da 27 kDa dell'ATP sintasi mitocondriali (sovra-espressa in frutto rosso), la catena grande della ribulosio bisfosfato carbossilasi/ossigenasi (sotto-espressa in frutto

rosso) e la ribulosio bisfosfato carbossilasi/ossigenasi attivasi (sovra-espresso nel frutto T). Analisi densitometriche condotte in pomodoro hanno dimostrato che i livelli di proteine appartenenti all'apparato fotosintetico presentano il loro massimo durante gli stadi iniziali della maturazione e quindi diminuiscono negli stadi successivi. Tali risultati sono interpretabili alla luce dell'attesa progressiva degradazione della clorofilla e dell'apparato fotosintetico associata alla transizione dei cloroplasti a cromoplasti (Rocco et al., 2006).

4.4 Proteine correlate ai genotipi

L'analisi comparativa tra la varietà d'élite QE ed i relativi genotipi parentali, M e U, ha evidenziato un numero relativamente limitato di proteine accumulate in maniera differenziale. Ciò non deve sorprendere, poiché la diversità biologica tra differenti genotipi è generalmente bassa. Infatti, le variazioni del proteoma, così com'è stato descritto (Alm et al., 2007), possono essere correlate a differenze nelle condizioni di crescita, piuttosto che a differenze intrinseche tra le diverse varietà di fragola. La variabilità all'interno di un genotipo è il risultato della cosiddetta plasticità fenotipica, rappresentata dall'adattamento dell'individuo alle differenze nell'ambiente circostante (Alm et al., 2007). Entrambe le linee parentali hanno mostrato, tra le proteine regolate, alcune macromolecole associate al metabolismo energetico e del carbonio ed al controllo dei principali caratteri qualitativi del frutto, come l'aroma e la consistenza (Tabella 3.5).

La regolazione differenziale dell'allergene Fra a1, due volte più abbondante nella varietà di riferimento QE vs U, offre una spiegazione plausibile del perché alcuni individui, affetti da intolleranza, possono tollerare meglio alcune varietà di fragola.

L'allergene Fra a1 è omologo al principale allergene del polline di betulla, la proteina Bet v 1.

I composti volatili presenti in fragola sono stati ampiamente studiati. Ne sono stati classificati oltre 360 (Honkanen and Hirvi, 1990; Latrasse, 1991). Anche il rapporto tra contenuto in sostanze volatili/aromatiche e varietà di fragola è stato oggetto di numerosi studi (Douillard and Guichard, 1989 e 1990; Larsen and Poll, 1992; Larsen et al., 1992), che hanno dimostrato come l'abbondanza relativa di particolari volatili, possa essere diagnostica di una particolare cultivar. L'analisi comparativa ha dimostrato, inoltre, che l'enzima FaQR è *up-regolato*, in entrambe le linee parentali, rispetto a QE (Tabella 3.5). Questi risultati concordano con quelli ottenuti da studi trascrittomici e metabolomici, precedentemente condotti nei nostri laboratori (Carbone et al, 2006), in cui sono stati investigati i livelli di trascritti, le attività enzimatiche e l'emissione di sostanze volatili. Queste indagini hanno stabilito che l'emissione di esteri aromatici in QE è più bassa rispetto ai genotipi parentali, caratterizzati da un'espressione *up-regolata* dei geni associati alla biosintesi dell'aroma (Carbone et al, 2006).

5 CONCLUSIONI

Durante le attività sperimentali inerenti al Dottorato di Ricerca in Biologia Vegetale, XXI ciclo, le metodologie analitiche della proteomica sono state sviluppate ed applicate allo studio del proteoma espresso nel frutto di fragola. In particolare, differenti approcci proteomici sono stati utilizzati per caratterizzare le proteine espresse in maniera variabile nel frutto, durante la maturazione ed in diversi genotipi.

I profili d'espressione proteica rappresentano oggi un aspetto molto importante della ricerca biologica, poiché i cambiamenti nella fisiologia di un organismo o di una cellula si accompagnano a variazioni nel *pattern* d'espressione proteica. Ciononostante, lo studio delle proteine espresse in fragola durante la maturazione è scarsamente rappresentato nella letteratura. Gli unici lavori pubblicati in questo campo riguardano, infatti, aspetti diversi dalla maturazione (identificazione di allergeni). Il lavoro svolto nel corso del Dottorato di Ricerca rappresenta, perciò, il primo tentativo di caratterizzazione del proteoma espresso nel frutto di fragola, durante la maturazione ed in diversi genotipi.

L'analisi proteomica su larga scala è stata perseguita attraverso l'utilizzo di diversi approcci proteomici, che hanno permesso:

- 1) di creare di una mappa elettroforetica bidimensionale, di riferimento per il proteoma del frutto;
- 2) di identificare e descrivere un vasto set di proteine espresse nel frutto;
- 3) di identificare un gruppo di proteine espresse in maniera differenziale durante la maturazione;
- 4) di identificare un set di proteine espresse in maniera differenziale nel confronto tra genotipi.

Le proteine costitutive, e/o accumulate in maniera differenziale durante la maturazione e nella comparazione tra genotipi, sono state discusse in base alla funzione biologica svolta e all'importanza posseduta nel determinare i principali caratteri qualitativi del frutto maturo.

Lo *screening* proteomico riportato nella presente Tesi di Dottorato apporta una maggiore conoscenza del metabolismo del frutto e della sua regolazione durante la maturazione. Solo attraverso una comprensione approfondita di questi eventi, è, infatti, possibile sviluppare nuove tecniche biomolecolari per il miglioramento genetico della qualità della fragola mediante specifici programmi di *breeding* o mediante sistemi di ingegneria metabolica.

6 BIBLIOGRAFIA

1. **Aharoni, A.**, Giri, A. P., Verstappen, F.W., Berteaux, C. M., Sevenier, R., Sun, Z., Jongsma, M. A., Schwab, W., and Bouwmeester, H. J. **2004**. Gain and loss of fruit flavor compounds produced by wild and cultivated strawberry species. *Plant Cell* 16: 3110–3131.
2. **Aharoni, A.**, Keizer, L. C., Van Den Broeck, H. C., Blanco-Portales, R., Munoz-Blanco, J., Bois, G., Smit, P., De Vos, R. C., and O'Connell, A. P. **2002**. Novel insight into vascular, stress, and auxin-dependent and -independent gene expression programs in strawberry, a non-climacteric fruit. *Plant Physiol.* 129:1019–1031.
3. **Alm R.**, Ekefjård A, Krogh M, Häkkinen J, Emanuelsson C. **2007** Proteomic Variation Is as Large within as between Strawberry Varieties. *J Proteome Res*;6:3011-20.
4. **Almeida JRM**, D'Amico E, Preuss A, Carbone F, de Vos CHR, Deiml B, Mourgues F, Perrotta G, Fischer TC, Bovy AG, Martens S, Rosati C. **2007** Characterization of major enzymes and genes involved in flavonoid and proanthocyanidin biosynthesis during fruit development in strawberry (*Fragaria × ananassa*). *Arch of Biochem and Biophys* 465:61-71.
5. **Alonso JM**, Chamarro J, Granell A. **1995**. Evidence for the involvement of ethylene in the expression of specific RNAs during maturation of the orange, a nonclimacteric fruit. *Plant Mol Biol* 29:385–390.
6. **Amme, S.**, Matros, A., Schlesier, B., Mock, H. P. **2006** Proteome analysis of cold stress response in *Arabidopsis thaliana* using DIGE-technology. *J. Exp. Bot.* 57, 1537–1546.
7. **Arabidopsis Genome Initiative** **2000**. Analysis of the genome sequence of the flowering plant *Arabidopsis*. *Nature* 408:796–815.
8. **Arrigo AP.** **1998** Small stress proteins: Chaperones that act as regulators of intracellular redox state and programmed cell death. *Biol Chem* 379:19-26.
9. **Baerenfaller, K.**, Grossmann, J, Grobei, MA:, Hull, R., Hirsch-Hoffmann, M., Yalovsky, S., Zimmermann, P., Grossniklaus, U., Gruissem, W., Baginsky S. **2008** *Arabidopsis thaliana* Gene Models and Proteome Dynamics ; 320 *Science* 938: 941.
10. **Balogh, A.**, Koncz, T., Tisza, V., Kiss, E., and Heszky, L. **2005**. The effect of 1-MCP on the expression of several ripening-related genes in strawberries. *Hortscience* 40: 2088–2090.
11. **Banks, JF.** and Whitehouse, CM. **1996** Electrospray ionisation mass spectrometry. *Methods Enzymol.* 270: 486-518.
12. **Beavis RC.**, and Chait BT. **1996** Matrix-assisted laser-desorption ionization mass spectrometry *Methods Enzymol.* 270: 519-551
13. **Blackstock, WP.**, and Weir, MP., **1999**. *Trends in Biotechnol.* 17, 121-127.
14. **Boggio SB.**, Palatnik JF, Heldt HW, Valle EM. **2000** Changes in amino acid composition and nitrogen metabolizing enzymes in ripening fruits of *Lycopersicon esculentum* Mill. *Plant Sci*;159:125-33.
15. **Boston, RS.**, Viitanen, P.V., Vierling, E. **1996** Molecular chaperones and protein folding in plants *Plant Mol Biol* 32:1919-222.
16. **Britsch, L.**, Dedio, J., Saedler, H., Forkmann, G. **1993** Molecular characterization of flavanone 3 β -hydroxylases. Consensus sequence,

- comparison with related enzymes and the role of conserved histidine residues *Eur. J. Biochem.* 217: 745-754
17. **Cagney, G., Emili, A. 2002** De novo peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abundance tagging *Nat. Biotechnol.* 20:163–170.
 18. **Carbone, F.,** Mourgues, F., Biasioli, F., Gasperi, F., Märk, TD., Rosati C., Perrotta G. **2006** Development of molecular and biochemical tools to investigate fruit quality traits in strawberry elite genotypes. *Mol Breeding* 18:127–42.
 19. **Carpentier, SC.,** Witters, E., Laukens, K., Van Onckelen, H., Swennen, R., Panis B. **2007a.** Banana (*Musa spp.*) as a model to study the meristem proteome: Acclimation to osmotic stress. *Proteomics* 7:92–105.
 20. **Carpentier, SC.,** Panis, B., Vertommen, A., Swennen, R., Sergeant, K., Renaut, J., Laukens, K., Witters, E., Samyn, B., Devreese, B. **2008** Proteome analysis of non-model Plants: A challenging but powerful approach. *Mass Spectrometry Reviews* 27:354-377.
 21. **Carrari, F., Fernie AR. 2006** Metabolic regulation underlying tomato fruit development *Journal of Experimental Botany* 57:1883-1897.
 22. **Chen CH., 2008.** Review of a current role of mass spectrometry for proteome research. *analytica chimica acta* 624:16–36.
 23. **Chen, X.,** Smith, LM., Bradbury, EM. **2000** Site-specific mass tagging with stable isotopes in proteins for accurate and efficient protein identification. *Anal. Chem.* 72: 1134-1143.
 24. **Chew, O.,** Whelan, J., Millar, H. J. **2003** Molecular Definition of the Ascorbate-Glutathione Cycle in Arabidopsis Mitochondria Reveals Dual Targeting of Antioxidant Defenses in Plants *Biol Chem* 278:46869-46877.
 25. **Civello, PM.,** Powell, AL., Sabehat, A., and Bennett, AB. **1999.** An expansin gene expressed in ripening strawberry fruit. *Plant Physiol.* 121: 1273–1280.
 26. **Clements RL.** 1970. Protein patterns in fruit. In: Hulme AC, editor. *The biochemistry of fruit and their products.* London and New York: Academic press; p.159-77.
 27. **Consortium, I.H.G.S., 2004.** Finishing the euchromatic sequence of the human genome. *Nature* 431, 931–945.
 28. **D'Amico E and Perrotta G. 2005** Genomics of berry fruits antioxidant components. *BioFactors*;23:179-87.
 29. **Daves GD. 1979.** Mass spectrometry of involatile and thermally unstable molecules. *Accts. Chem. Res.* 12: 359-365.
 30. **Davis KM. and Folta, TM 2006** Strawberry Genes and Genomics. *Strawberry Genes and Genomics', Critical Reviews in Plant Sciences,* 25:5, 399 – 415.
 31. **Dixon, RA. and Pavia, NL. 1995** Stress-induced phenylpropanoid metabolism. *Plant Cell* 7:1085-1097.
 32. **Dooner, HK.,** Robbins, TP., Jorgensen, RA. **1991.** Genetic and developmental control of anthocyanin biosynthesis. *Annu. Rev. Genet.* 25: 173-199.
 33. **Douillard, C. and Guichard, E. 1989.** Comparison by multidimensional analysis of concentrations of volatile compounds in fourteen frozen strawberry varieties. *Sci Aliments* 9:53–76.

34. **Douillard C. and Guichard E. 1990.** The aroma of strawberry (*Fragaria ananassa*): Characterisation of some cultivars and influence of freezing. *J Sci Food Agric* 50:517–531.
35. **Duck, N., McCormick, S., Winter, J. 1989** Heat shock protein Hsp70 cognate gene expression in vegetative and reproductive organs of *Lycopersicon esculentum*. *Proc Natl Acad Sci USA* 86:3674–8.
36. **Evans, C.R. and Jorgenson, J.W. 2004** Multidimensional LC-LC and LC-CE for high-resolution separations of biological molecules *Analytical and bioanalytical chemistry* 378:1952 -1961
37. **Fait, A., Hanhineva, K., Beleggia, R., Dai, N., Rogachev, I., Nikiforova, VJ., Fernie, AR., , and Aharoni, A.. 2008** Reconfiguration of the Achene and Receptacle Metabolic Networks during Strawberry Fruit Development. *Plant Physiology*, 148:730–750.
38. **Faurobert, M., Mihr, C., Bertin, N., Pawlowski, T., Negroni, L. 2007** Major Proteome Variations Associated with Cherry Tomato Pericarp Development and Ripening. *Plant Physiol* 143:1327–46.
39. **Ferrandiz C., Liljegren SJ, Yanofsky MF. 2000.** Negative regulation of the SHATTERPROOF genes by FRUITFULL during Arabidopsis fruit development. *Science* 289:436–38.
40. **Forkmann, G. 1991** Flavonoids as flower pigments: the formation. of the natural spectrum and its extension by genetic engineering *Plant Breed.* 106:1-26.
41. **Forkmann,G. and Heller, W. 1999** In: *Comprehensive Natural Products Chemistry*, Elsevier 713-748.
42. **Fuke, S. and Konosu, S. 1991.** Taste-active components in some foods: A review of Japanese research. *Physiol Behav* 49:863-8.
43. **Gachet, Y., Lee, M, Sawitzki, M., Tournier, S., Poulton, T. 1997.** Intracellular colocalisation of the translationally controlled protein P23 with cytoskeletal structures. *Biochem Soc Trans* 25: 269.
44. **Gallardo, F., Galvez, S., Gadal, P., Canovas, FM. 1995** Changes in NADP⁺-linked isocitrate dehydrogenase during tomato fruit ripening. Characterization of the predominant cytosolic enzyme from green and ripe pericarp. *Planta* 196:148-54.
45. **Gatlin, CL., Kleemann, GR., Hays, LG., Link, AJ., Yates, JR III. 1998** Protein identification at low femtomole level from silver-stained gels using a new fritless electrospray interface for liquid chromatography-microspray and nanospray mass spectrometry. *Anal >Biochem* 263:93-101.
46. **Giribaldi, M., Perugini, I., Sauvage, FX., Schubert, A., 2007.** Analysis of protein changes during grape berry ripening by 2-DE and MALDI-TOF. *Proteomics* 2007, 7, 3154–3170.
47. **Given, N. K., M. A. Venis and D. Grierson 1988.** Hormonal regulation of ripening in the strawberry, a non-climacteric fruit. *Planta* 174: 402–406.
48. **Goff, SA., Ricke, D., Lan, TH., Presting, G., Wang, RL., Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchinson D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong JP, Miguel T, Paszkowski U, Zhang SP, Colbert M, Sun WL, Chen LL, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu YS, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM,**

- Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A, Briggs S. **2002**. Adraft sequence of the rice genome (*Oryza sativa* L. ssp japonica). *Science* 296:92–100.
49. **Gong, H.**, Jiao, Y., Hu, W., Pua, E. **2005**. Expression of glutathione-S-transferase and its role in plant growth and development in vivo and shoot morphogenesis in vitro. *Plant Mol Biol* 57:53-66.
50. **Görg, A.**, Obermaier, C., Boguth, G., Harder, A., Scheibe, B., Wildgruber, R., Weiss, W **2000**. The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 21: 1037-1053
51. **Gu, S.**, Pan, S., Bradbury, E. M., Chen, X., **2002**. Use of deuterium-labeled lysine for efficient protein identification and peptide de novo sequencing. *Anal. Chem.* 74, 5774–5785.
52. **Guarino, C.**, Arena, S., De Simone, L., D’Ambrosio, C., Santoro, S, Rocco, M., Scaloni, A., and Marra M. **2007**. Proteomic analysis of the major soluble components in Annurca apple flesh. *Mol. Nutr. Food Res.* 51: 255 – 262
53. **Gygi, SP.**, Rist, B., Gerber, SA., Turecek, F., Gelb, M., Aebersold, R. **1999**. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* 17: 994-999.
54. **Gygi, S.P.**, Corthals, G.L., Zhang, Y., Rochon, Y., Aebersold, R. **2000** Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proceedings of the National Academy of Sciences of the United States of America* 97 (17), pp. 9390-9395
55. **Hadfield KA. and Bennett AB** **1998**. Polygalacturonases: many genes in search of a function. *Plant Physiol.* 117: 337-343.
56. **Hale, JE.**, Butler, JP., Knierman, MD., Becker, G.W. **2000**. Increased sensitivity of tryptic peptide detection by MALDI-TOF mass spectrometry is achieved by conversion of lysine to homoarginine *Anal. Biochem.* 287, 110–117.
57. **Hancock JF.** 1999. Strawberries. *Crop production science in horticulture.* Cambridge University Press.
58. **Harbone, JB. and Williams, CA.** **2000** Advances in flavonoid research since 1992. *Phytochemistry* 55:481-504.
59. **Harpster, MH.**, Brummell, DA., and Dunsmuir, P. **1998**. Expression analysis of a ripening-specific, auxin-repressed endo-1, 4-beta-glucanase gene in strawberry. *Plant Physiol.* 118: 1307–1316.
60. **Harrison, AG.** **1992**. Chemical ionization mass spectrometry. CRC Press, Boca Raton, Florida
61. **Harrison, AG. and Cotter RJ.** **1990**. Methods of ionization. *Method Enzymol.* 193: 237-263.
62. **Hayama, H.**, Ito, A., Moriguchi, T., Kashimura, Y. **2003**. Identification of a new expansin gene closely associated with peach fruit softening. *Postharvest Biol. and Technol.* 29: 1-10.
63. **Hays, L.**, Schieltz D, Ulaszek R, Wei J, Wolters D, Yates JR III. **2002**. Proteomic survey of metabolic pathways in rice. *Proc Natl Acad Sci USA* 99:11969–11974.
64. **Higashi, Y.**, Hirai, MY., Fujiwara, T., Naito, S., Noji, M., Saito, K. **2006**. Proteomic and transcriptomic analysis of Arabidopsis seeds: Molecular evidence for successive processing of seed proteins and its implication in the stress response to sulfur nutrition. *Plant J* 48:557–571.

65. **Hjernø, K.**, Alm R, Canbäck B, Matthiesen R, Trajkovski K, Bjork L, Roepstorff P, Emanuelsson C. **2006** Down-regulation of the strawberry Bet v-homologous allergen in concert with the flavonoid biosynthesis pathway in colorless strawberry mutant *Proteomics* 6:1574-87.
66. **Holm, J.**, Baerentzen, G., Gajhede, M., Ipsen, H., Larsen, JN., Lowenstein H, Wissenbach M, Spangfort, MD. **2001** Molecular basis of allergic cross-reactivity between group 1 major allergens from birch and apple. *J Chromatogr B Biomed Sci Appl* 756:307-13.
67. **Honkanen, E. and Hirvi, T. 1990.** The flavour of berries. In: Morton ID, MacLeod AJ, editor. *Food Flavours*. Amsterdam: Elsevier Scientific Publications; p. 125-193.
68. **Issaq, HJ.**, Conrads, TP., Janini, GM., Veenstra, TD. **2002** Methods for fractionation, separation and profiling of proteins and peptides *Electrophoresis* 23 (17), pp. 3048-3061
69. **Jimenez, A.**, Creissen, G., Kular, B., Firmin, J., Robinson, S., Verhoeven M, Mullineaux P. **2002** Changes in oxidative processes and components of the antioxidant system during tomato fruit ripening. *Plant* 214:751-8.
70. **Jimenez-Bermudez, S.**, Redondo-Nevado, J., Munoz-Blanco, J., Caballero, J. L., Lopez-Aranda, J. M., Valpuesta, V., Pliego-Alfaro, F., Quesada, M. A., and Mercado, J. A. **2002.** Manipulation of strawberry fruit softening by antisense expression of a pectate lyase gene. *Plant Physiol.* 128: 751-759.
71. **Jones, AME.**, Thomas, V., Bennett, M. H., Mansfield, J., Grant, M., **2007** Modifications to the Arabidopsis defense proteome occur prior to significant transcriptional change in response to inoculation with *Pseudomonas syringae*. *Plant Physiol.* 142, 1603-1620.
72. **Jorgensen, RA.**, Cluster, PD., English, J., Que, Q., Napoli, CA. **1996** *Plant Mol.Biol.* 31:957-973
73. **Jorriñ JV.**, Maldonado, AM., Castillejo, MA **2007** *Plant Proteome Analysis: A 2006 update. Proteomics*, 7: 2947-2962.
74. **Kafkas, E.**, Kosar, M., Paydas, S., Kafkas, S., Baser KHC. **2007** Quality characteristics of strawberry genotypes at different maturation stages. *Food Chem* 100:1229-36.
75. **Karas, M.**, Gluckmann, M., Schafer J., **2000.** Ionization in matrix-assisted laser desorption /ionization. Singly charged molecular ions are the lucky survivors. *J.Mass Spectrom.* 35:1-12.
76. **Keough, T.**, Lacey, MP., Youngquist, RS. **2002** Solid-Phase Derivatization of Tryptic Peptides for Rapid Protein Identification by Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. *Rapid Commun. Mass Spectrom.* 16:1003-1015.
77. **Keough, T.**, Lacey, MP., Youngquist, RS. **2000** Derivatization Procedures to Facilitate de Novo Sequencing of Lysine-Terminated Tryptic Peptides Using Postsources Decay Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. *Rapid Commun. Mass Spectrom.* 14:348-2356.
78. **Klee, HJ. 1995.** Identification of mRNAs with enhanced expression in ripening strawberry fruit using polymerase chain reaction differential display. *Plant Mol. Biol.* 27:1097-108.
79. **Knee, M.**, Sargent, J. A., and Osborne, D. J. **1977.** Cell wall metabolism in developing strawberry fruits. *J. Exp. Bot.* 28: 377-396.

80. **Koch, K. 2004** Sucrose metabolism: Regulatory mechanisms and pivotal roles in sugar sensing and plant development. *Curr Opin Plant Biol* 7:235-46.
81. **Koller, A,** Washburn MP, Markus-Lange B, Andon NL, Deciu C, Haynes PA, Hays L, Schieltz D, Ulaszek R, Wei J, Wolters D, Yates JR III. **2002** Proteomic survey of metabolic pathways in rice. *Proc Natl Acad Sci USA* 99:11969–11974.
82. **Komatsu, S.,** Muhammad, A., Rakwal, R. **1999.** Separation and characterization of proteins from green and etiolated shoots of rice (*Oryza sativa* L.): towards a rice proteome *Electrophoresis* 20, 630–636.
83. **Konopka-Postupolska, D. 2007** Annexins: Putative linkers in dynamic membrane-cytoskeleton interactions in plant cells *Protoplasma* 230 (3-4), pp. 203-215 .
84. **Laemmli, U. 1970** Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-5.
85. **Lander, ES.,** Linton, LM., Birren, B., Nusbaum, C., Zody, MC., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al., **2001.** Initial sequencing and analysis of the human genome. *Nature* 409, 860–921.
86. **Larsen, M.,** Poll, L., Olsen, CE. **1992** Evaluation of the aroma composition of some strawberry (*Fragaria ananassa* Duch) cultivars by use of odour threshold values. *Z. Lebensm. Unters. Forsch.* 195: 536–539.
87. **Larsen, M.,** Poll, L., Olsen, CE. **1992** Evaluation of the aroma composition of some strawberry (*Fragaria ananassa* Duch) cultivars by use of odour threshold values. *Z Lebensm Unters Forsch* 195:536–9.
88. **Larsen, M and Poll, L. 1992** Odor thresholds of some important aroma compounds in strawberries. *Z Lebensm Unters Forsch* 195:120–3.
89. **Latrasse, A. 1991** Fruits III. In: Maarse H, editor. *Volatile Compounds in Foods and Beverages.* New York: Dekker; p. 329–387.
90. **Laugesen, S. and Roepstorff, P.. 2003** Combination of Two Matrices Results in Improved Performance of MALDI MS for Peptide Mass Mapping and Protein Analysis. *J Am Soc Mass Spectrom* 4:992-1002.
91. **Lee, J.,** Garrett, W.M., Cooper, B. **2007** Shotgun proteomic analysis of *Arabidopsis thaliana* leaves *Journal of Separation Science* 30 (14), pp. 2225-2230.
92. **Lelievre, JM.,** Latche, A., Jones, B., Bouzayen, M., and Pech, J.C. **1997.** Ethylene and fruit ripening. *Physiol. Plant.* 101, 727–739.
93. **Lester, DR,** Speirs, G, Orr G, Brady CJ. **1994** Peach (*Prunus persica*) endo-PG cDNA isolation and mRNA analysis in melting and non-melting peach cultivars. *Plant Physiol.* 105: 225-231
94. **Li,Y.,** Jones L, McQueen-Mason S. **2003** Expansins and cell growth. *Current Op. Plant Biol.* 6: 603-610
95. **Lincoln, JE.,** Cordes, S., Read, E., Fischer, RL. **1987** Regulation of gene expression by ethylene during *Lycopersicon esculentum* (tomato) fruit development. *Proc Natl Acad Sci USA* 84: 2793–2797
96. Lindh, I., Hjelmqvist, L., Bergman, T., Sjovall, J., Griffiths, W.J., 2000 De novo sequencing of proteolytic peptides by a combination of C-terminal derivatization and nano-electrospray/collision-induced dissociation mass spectrometry *J. Am. Soc. Mass Spectrom.* 11, 673–686.

97. **Lopez, AP. and Franco, AR. 2006** Cloning and expression of cDNA encoding translationally controlled tumor protein from strawberry fruits. *Biol. Plantarum* 50:447-9.
98. **Lopez, AP.,** Portales, RB., Lopez-Raez, JA., Medina-Escobar N, Blanco JM, Franco AR **2006.** Characterization of a strawberry late-expressed and fruit-specific peptide methionine sulphoxide reductase. *Physiol Plant* 126: 129–139.
99. **Macherel, D.,** Benamar, A., Avelange-Macherel, M. H., Tolleter, D. **2007** Function and stress tolerance of seedmitochondria. *Physiol. Plantarum* 129, 233–241.
100. **Mann, M.,** Hendrickson, RC., Pandey, A. **2001.** Analysis of proteins and proteome by mass spectrometry. *Annu. Rev. Biochem.* 70:437-473.
101. **Mann, M., and Wilm, M. 1994** Error-tolerant identification of peptides in sequence databases by peptide sequence tags *Anal. Chem.* 66, 4390–4399.
102. **Manning, K. 1998.** Isolation of a set of ripening-related genes from strawberry: their identification and possible relationship to fruit quality traits. *Planta* 205:622–31
103. **Manning, K. 1993.** Soft fruit. In G. B. Seymour, J. E. Taylor and G. A. Tucker (eds.), *Biochemistry of fruit ripening.* Chapman & Hall, London, pp. 347–377.
104. **Martin, SE.,** Shabanowitz, J., Hunt, DF., Marto, J.A. **2000** Subfemtomole MS and MS/MS peptide sequence analysis using nano-HPLC micro-ESI fourier transform ion cyclotron resonance mass spectrometry *Analytical Chemistry* 72 (18), pp. 4266-4274
105. **Marty, I, Lelievre, JM. 1999.** Isolation and characterization of mRNAs differentially expressed during ripening of wild strawberry (*Fragaria vesca* L.) fruits. *Plant Mol. Biol.* 39:629–36.
106. **Mata, J.,** Marguerat, S, Bahler J. **2005.** Post-transcriptional control of gene expression: A genome-wide perspective. *Trends Biochem Sci* 30:506–514.
107. **Mathews, H.,** Wagoner, W., Kellogg, J., and Bestwick, R. **1995.** Genetic transformation of strawberry—Stable integration of a gene to control biosynthesis of ethylene. *In Vitro Cell. Dev. Biol.-Plant* 31: 36–43.
108. **Medina-Escobar, N.,** Cardenas, J., Moyano, E., Caballero, J. L., and Munoz-Blanco, J. **1997.** Cloning, molecular characterization and expression pattern of a strawberry ripening-specific cDNA with sequence homology to pectate lyase from higher plants. *Plant Mol. Biol.* 34: 867–877.
109. **Medina-Escobar, N.,** Cardenas, J., Munoz-Blanco, J., and Caballero, J. L. **1998.** Cloning and molecular characterization of a strawberry fruit ripening-relatedcDNA corresponding a mRNA for a low-molecular-weight heat-shock protein. *Plant Mol. Biol.* 36: 33–42.
110. **Mitulović, G.,** Stingl, C., Smoluch, M., Swart, R., Chervet, J.-P., Steinmacher, I., Gerner, C., Mechtler, K. **2004** Automated, on-line two-dimensional nano liquid chromatography tandem mass spectrometry for rapid analysis of complex protein digests. *Proteomics* 4 (9), pp. 2545-2557
111. **Moing, AC.,** Renaud, M. Gaudillere, P. Raymond, P. Roudeillac and B. Denoyes-Rothan **2001.** Biochemical changes during fruit development of four strawberry cultivars. *Journal of the American Society for Horticultural Science* 126: 394–403.

112. **Morris HR.**, Dickinson RJ., Williams DH. **1973.** Studies toward the complete sequence determination of protein by mass spectrometry. Derivatization of methionine, cysteine, arginine containing peptides. *Biochem. Biophys. Res. Commun.* 51:247-255.
113. **Moyano, E.**, Portero-Robles, I., Medina-Escobar, N., Valpuesta, V., Munoz-Blanco, J., and Caballero, J. L. **1998.** A fruit-specific putative dihydroflavonol 4-reductase gene is differentially expressed in strawberry during the ripening process. *Plant Physiol.* 117: 711-716.
114. **Munchbach, M.**, Quadroni, M., Miotto, G., James, P. **2000** Quantitation and facilitated de novo sequencing of proteins by isotopic N-terminal labeling of peptides with a fragmentation-directing moiety *Anal. Chem.* 72, 4047–4057.
115. **Nam, YW.**, Tichit, L., Leperlier, M., Cuerq, B., Marty, I., Lelièvre, J.-M. **1999** Isolation and characterization of mRNAs differentially expressed during ripening of wild strawberry (*Fragaria vesca* L.) fruits *Plant Molecular Biology* 39: 629-636
116. **Neta-Sharir, I.**, Isaacson, T., Lurie, S., Weiss D. **2005** Dual role for tomato heat shock protein 21: protecting photosystem II from oxidative stress and promoting color changes during fruit maturation. *Plant Cell* 17: 1829–1838
117. **Neuhoff, V.**, Arold, N., Taube, D., Ehrhardt, W. **1988** Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* 9:255-62.
118. **Nguyen-Quoc, B. and Foyer, CH.** **2001.** A role for 'futile cycles' involving invertase and sucrose synthase in sucrose metabolism of tomato fruit. *J Exp Bot* 52:881-9.
119. **Nitsch, JP.** **1950.** Growth and morphogenesis of the strawberry as related to auxin. *Am J Bot* 37: 211–215.
120. **Oda, Y.**, Huang, K., Cross, FR., Cowburn, D., Chait BT. **1999** Accurate quantitation of protein expression and site specific phosphorylation. *Proc. Nat. Acad. Sci. USA* 8: 6591-6596.
121. **Omura, T.**, Watanabe, S., Iijima, Y., Aoki, K., Shibata, D., Ezura, H. **2007** Molecular and genetic characterization of transgenic tomato expressing 3-hydroxy-3-methylglutaryl coenzyme A reductase. *Plant Biotech* 24:107-15.
122. **Ong, SE.**, Blagoev, B., Kratchmarova, I., Kristensen, DB., Stehen, H., Pandey, A., Mann M. **2002** Stable isotope labelling by Amino Acid in Cell Culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* 1:376-386.
123. **Parinov, S and Sundaresan, V.** **2000.** Functional genomics in Arabidopsis: Largescale insertional mutagenesis complements the genome sequencing project. *Curr. Opin. Biotechnol.* 11:157–61.
124. **Patterson and Aebersold** **2003.** Proteomics: the first decade and beyond *Nature genetics supplement* 33: 311-323
125. **Perkins-Veazie, P.** **1995.** Growth and ripening of strawberry fruit. *Hortic Rev (Am Soc Hortic Sci)* 17: 267–297.
126. **Popova, TN.**, Pinheiro de Carvalho, MAA. **1998** Citrate and isocitrate in plant metabolism. *Biochim. Biophys. Acta - Bioenergetics* 1364:307-25.
127. **Pressey, R.** **1978.** Avants JK. Differences in polygalacturonase composition of clingstone and freestone peaches. *J. Food Sci.* 43: 1415-1421.

128. **Proust, J., Houlné, G., Schantz, ML., Schantz, R. 1996** Characterization and gene expression of an annexin during fruit development in *Capsicum annum*. *FEBS Letters* 383:208-12.
129. **Pyysalo, T., Honkanen, E., Hirvi, T. 1979** Volatiles of wild strawberries, *Fragaria vesca* L. compared to those of cultivated berries, *Fragaria 3 ananassa* cv. Senga Sengana. *J. Agric. Food Chem.* 27: 19–22.
130. **Raab, T., Lopez-Raez, JA., Klein, D., Caballero, J.L., Moyano, E., Schwab, W., Munoz-Blanco, J., 2006.** FaQR, required for the biosynthesis of the strawberry flavor compound 4-hydroxy-2,5-dimethyl-3(2H)-furanone, encodes an enone oxidoreductase. *The plant Cell*, 18, 1023-1037.
131. **Rabilloud, T., Valette C, Lawrence JJ 1994** Sample application by in-gel rehydration improves the resolution of two-dimensional electrophoresis with immobilized pH gradients in the first dimension. *Electrophoresis* 15: 1552-1558
132. **Raubilloud, T. 1996.** Solubilization of proteins for electrophoretic analyses. *Electrophoresis* 17, 813-829.
133. **Reddy, ASN., Jena, PK., Mukherjee, SK., and Poovaiah, B. W. 1990.** Molecular cloning of cDNAs for auxin-induced messenger-RNAs and developmental expression of the auxin-inducible genes. *Plant Mol. Biol.* 14:643–653.
134. **Reichheld, JP., Vernoux, T., Lardon F., Van Montagu, M., Inzé D. 1999** Specific checkpoints regulate plant cell cycle progression in response to oxidative stress. *Plant J* 17:647-656.
135. **Reid G.E. and McLuckey, S.A. J. 2002** 'Top down' protein characterization via tandem mass spectrometry *Mass Spectrom.* 37 663.
136. **Roberts, JKM. 2002.** Proteomics and a future generation of plant molecular biologists *Plant Molecular Biology* 48, 143-154.
137. **Rocco, M., D'Ambrosio, C., Arena, S., Faurobert, M., Scaloni, A., Marra, M. 2006** Proteomic analysis of tomato fruits from two ecotypes during ripening *Proteomics* 6 (13), pp. 3781-3791
138. **Romero, HM., Berlett, BS., Jensen, PJ., Pell E, Tien M. 2004** Investigations into the Role of the Plastidial Peptide Methionine Sulfoxide Reductase in Response to Oxidative Stress in *Arabidopsis*. *Plant Physiol* 136: 3784–3794.
139. **Roscher, R., Bringmann, G., Schreier, P., Schwab W. 1998** Radiotracer studies on the formation of 2,5- dimethyl-4-hydroxy-3(2H)-furanone in detached ripening strawberry fruits. *J Agric Food Chem* 46:1488-93.
140. **Rose, JKM., Bashir, S., Giovannoni, JJ., Jahn, MM., Saravan, RS. 2004.** *The plant journal* 39, 715-733.
141. **Ruebelt, MC., Leimgruber, NK., Lipp, M., Reynolds, TL. et al., 2006** Application of two-dimensional gel electrophoresis to interrogate alterations in the proteome of genetically modified crops. 1. Assessing analytical validation. *J. Agric. Food Chem.* 54, 2154–2161.
142. **Ruebelt, MC., Lipp, M., Reynolds, TL., Astwood, JD. et al., 2006** Application of two-dimensional gel electrophoresis to interrogate alterations in the proteome of genetically modified crops. 2. Assessing natural variability. *J. Agric. Food Chem.* 54, 2162–2168.
143. **Ruebelt, MC., Lipp, M., Reynolds, TL., Schmuke, JJ. et al., 2006** Application of two-dimensional gel electrophoresis to interrogate

- alterations in the proteome of genetically modified crops. 3. Assessing unintended effects. *J. Agric. Food Chem.*, 54, 2169–2177.
144. **Salentijn EMJ.**, Aharoni A., Schaart JG., Boone MJ. and Krens FA. **2003** Differential gene expression analysis of strawberry cultivars that differ in fruit-firmness *Physiologia Plantarum* 118: 571-578.
 145. **Sanz, C.**, Richardson, DG., Pérez, AG. **1995** 2,5-Dimethyl-4- hydroxy-3(2H)-furanone and derivatives in strawberries during ripening. In: Rouseff RL, Leahy MM, editor. *Fruit Flavors: Biogenesis, Characterization, and Authentication*. ACS Symposium Series 596; American Chemical Society: Washington, DC; p.268-75.
 146. **Sappl, PG.**, Heazlewood, JL., Millar, AH. **2004** Untangling multi-gene families in plants by integrating proteomics into functional genomics. *Phytochemistry* 65:1517-30.
 147. **Schwab, W. and Roscher, R.** **1997** 4-Hydroxy-3(2H)-furanones: Natural and Maillard Products. *Recent Res DeV Phytochem* 1:643-73.
 148. **Schwab W.** Metabolome diversity: Too few genes, too many metabolites? *Phytochemistry* 1998;62:837-49.
 149. **Schwab W., and Raab, T.** **2004** Developmente changes during Strawberry fruti ripening and physico-chemical changes during postharvest storage. R. Dris and S. M. Jain (eds.), *Production Practices and Quality Assessment of Food Crops*, Kluwer Academic Publishers. Printed in the Netherlands. Vol. 3, “Quality Handling and Evaluation”, pp. 341–369.
 150. **Silverstein R. M. and Webster F. X.** **1999** Identificazione spettroscopica di composti organici. CEA - Casa Editrice Ambrosiana
 151. **Simpson, RJ.**, **2003**. Protein and Proteomics: a laboratory manual. Cold Spring Harbor Laboratory Press
 152. **Singla-Pareek, SL.**, Reddy, MK., Sopory, SK., **2003** *Proc. Natl. Acad Sci. USA* 100: 244-
 153. **Smith HB.** 1999. Sucrose synthase and the fruit of its labor. *Plant Cell* 11:2261-2.
 154. **Son, DY.**, Scheurer, S., Hoffmann, A, Haustein D, Vieths S. **1999** Pollen-related food allergy: cloning and immunological analysis of isoforms and mutants of Mal d 1, the major apple allergen, and Bet v 1, the major birch pollen allergen. *Eur J Nutr* 38:201-15.
 155. **Sonsmann, G.**, Romer, A., Schomburg, D., J. **2002** Investigation of the influence of charge derivatization on the fragmentation of multiply protonated peptides *Am. Soc. Mass Spectrom.* 13: 47–58.
 156. **Steen, H. and Mann M.**, **2004**. The ABC’s and XYZ’s of peptide sequencing. *Nature Reviews*; 7-19.
 157. **Sun, W.**, Van Montagu, M., Verbruggen, N. **2002** Small heat shock proteins and stress tolerance in plants. *Biochim Biophys. Acta - Gene Structure and Expression* 1577:1-9.
 158. **Suutarinen, J.**, Änäkääinen L. and Autio K. **1998**. Comparison of light microscopy and spatially resolved fourier transform infrared (FT-IR) microscopy in the examination of cell wall components of strawberries. *Lebensmittel-Wissenschaft und – Technologie* 31: 595–601.
 159. **Tanaka, Y.**, Tsuda, S., Kusumi, T. **1998** *Plant Cell. Physiol.* 39: 1119-1126

160. **Thompson, JE.**, Hopkins, MT., Taylor C, Wang TW. **2004** Regulation of senescence by eukaryotic translation initiation factor 5A: implications for plant growth and development. *Trends Plant Sci* 9:174–9.
161. **Toda, K.**, Yang, D., Yamanaka, N., Watanabe, S., Harada, K., Takahashi, R. **2002** *Plant Mol. Biol.* 50:187-196
162. **Tonge, RP.**, Shaw, J., Middleton, B., Rowlinson, R., Rayner, S., Young, J., Pognan, F., Hawkins, E., Currie, I., Davison, M. **2001**. Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology *Proteomics* 1 (3), pp. 377-396
163. **Trainotti, L.**, Pavanello, A., and Casadoro, G. **2005**. Different ethylene receptors show an increased expression during the ripening of strawberries: Does such an increment imply a role for ethylene in the ripening of these non-climacteric fruits? *J. Exp. Bot.* 56: 2037–2046.
164. **Trainotti, L.**, Spolaore, S., Pavanello, A., Baldan, B., and Casadoro, G. **1999**. A novel E-type endo-beta-1,4-glucanase with a putative cellulose-binding domain is highly expressed in ripening strawberry fruits. *Plant Mol. Biol.* 40:323–332.
165. **Tsugita, A.**, Kamo, M., Kawakami, T. & Ohki, Y. **1996** *Electrophoresis* 17, 855–865.
166. **Unlu, M.**, Morgan, M.E., Minden, J.S. **1997** Difference gel electrophoresis: A single gel method for detecting changes in protein extracts *Electrophoresis* 18 (11), pp. 2071-2077.
167. **Venter, JC.**, Adams, MD., Myers, EW., Li, PW., Mural, RJ., Sutton, GG., Smith, HO., Yandell, M., Evans, CA., Holt, RA., et al., **2001**. The sequence of the human genome. *Science* 291, 1304–1351.
168. **Vestal, M.L.** **1990** Liquid chromatography-mass spectrometry *Methods in Enzymology* 93 :107 -130
169. **Vrebalov, J.**, Ruezinsky, D., Padmanabhan, V., White, R, Medrano, D., Drake, R., Schuch W, Giovannoni J. **2002**. A MADS-box gene necessary for fruit ripening at the tomato ripening-inhibitor (rin) locus. *Science* 296:343–346.
170. **Wang, TW.**, Lu, L., Wang, D., Thompson, JE. **2001** Isolation and characterization of senescence-induced cDNAs encoding deoxyhypusine synthase and eukaryotic translation initiation factor 5A from tomato. *J Biol Chem* 276:17541–9.
171. **Wang, W.**, Vinocur B, Shoseyov O, Altman A. **2004** Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends Plant Sci* 9:244-52.
172. **Washburn, MP.**, Ulaszek, RR., Yates III, JR. **2003** Reproducibility of quantitative proteomic analyses of complex biological mixtures by multidimensional protein identification technology *Analytical Chemistry* 75 (19), pp. 5054-5061
173. **Wasinger, VC.**, Cordwell, XJ., Cerpapoliak, A., Yan, OX., Gooley, AA., Wilkins, MR., et al., 1995. Progress with gene-product mapping of the Mollicutes *Mycoplasma genitalium* *Electrophoresis* 16, 1090-1094.
174. **Waters, ER.** **1995** The molecular evolution of the small heat-shock proteins in plants *Genetics* 141: 785-795.
175. **Wein, M.**, Lewinsohn, E., Schwab, W. **2001**. Metabolic fate of isotopes during the biological transformation of carbohydrates to 2,5-dimethyl-4-

- hydroxy-3(2H)-furanone in strawberry fruits. *J Agric Food Chem* 49:2427-32.
176. **Wilkins, MR.,** Sanchez, JC., Gooley, AA., Appel, RD., Umphery-Smith, I., Hochstrasser, DF., Williams, KL. **1995** Progress with proteome projects : why all proteins expressed by a genome should be identified and how to do it. *Biotech. Gen. Eng. Rev.* 13:19-50.
177. **Wilkinson, JQ.,** Lanahan, MB., Conner, TW., and Klee, H. J. **1995.** Identification of messenger RNAs with enhanced expression in ripening strawberry fruit using polymerase chain reaction differential display. *Plant Mol. Biol.* 27:1097–1108.
178. **Wilson K., and Walker, J.** **1994** *Metodologia Biochimica Raffaello Cortina Editore*
179. **Winter, H, Huber SC.** **2000** Regulation of sucrose metabolism in higher plants: Localization and regulation of activity of key enzymes. *Crit Rev Biochem Mol Biol* 35:253-89.
180. **Wu, CC.,** MacCoss, M.J., Howell, K.E., Yates III, J.R. **2003** A method for the comprehensive proteomic analysis of membrane proteins *Nature Biotechnology* 21 (5), pp. 532-538
181. **Scheurer, S.,** Hoffmann, A., Haustein, D., Vieths, S. **1999** Pollen-related food allergy: cloning and immunological analysis of isoforms and mutants of Mal d 1, the major apple allergen, and Bet v 1, the major birch pollen allergen. *Eur J Nutr* 38:201-15.
182. **Yates III, JR.** **1998** Mass spectrometry and the age of the proteome. *Journal of Mass Spectrometry* 33:1-19.
183. **Yen, HC.,** Lee, S., Tanksley SD., Lanahan, MB., Klee HJ, Giovannoni JJ. **1995.** The tomato Never-ripe locus regulates ethyleneinducible gene expression and is linked to a homolog of the Arabidopsis ETR1 gene. *Plant Physiol* 107:1343–1353.
184. **Yu, J.,** Hu SN, Wang J, Wong GKS, Li SG, Liu B, Deng YJ, Dai L, Zhou Y, Zhang XQ, Cao ML, Liu J, Sun JD, Tang JB, Chen YJ, Huang XB, Lin W, Ye C, TongW, Cong LJ, Geng JN, Han YJ, Li L, LiW, HuGQ, Huang XG, Li WJ, Li J, Liu ZW, Li L, Liu JP, Qi QH, Liu JS, Li L, Li T, Wang XG, Lu H, Wu TT, Zhu M, Ni PX, Han H, Dong W, Ren XY, Feng XL, Cui P, Li XR, Wang H, Xu X, Zhai WX, Xu Z, Zhang JS, He SJ, Zhang JG, Xu JC, Zhang KL, ZhengXW, Dong JH, ZengWY, Tao L, Ye J, Tan J, Ren XD, ChenXW, He J, Liu DF, TianW, Tian CG, Xia HG, Bao QY, Li G, Gao H, Cao T, Wang J, ZhaoWM, Li P, ChenW, Wang XD, Zhang Y, Hu JF, Wang J, Liu S, Yang J, Zhang GY, Xiong YQ, Li ZJ, Mao L, Zhou CS, Zhu Z, Chen RS, Hao BL, Zheng WM, Chen SY, Guo W, Li GJ, Liu SQ, TaoM, Wang J, Zhu LH, Yuan LP, Yang HM. **2002.** A draft sequence of the rice genome (*Oryza sativa* L. ssp *indica*). *Science* 296:79–92.
185. **Zheng, Q.,** Song, J., Doncaster, K., Rowland E, Byers DM. **2007.** Qualitative and quantitative evaluation of protein extraction protocols for apple and strawberry fruit suitable for two-dimensional electrophoresis and mass spectrometry analysis. *J Agric Food Chem* 55:1663–1673.

Bianco L., Lopez L., Scalone A.G., Di Carli M., Desiderio A., Benvenuto E., and Perrotta G. Strawberry proteome characterization and its regulation during fruit ripening and in different genotypes. **Novembre 2008** *Journal of Proteomics In Press*.

Bianco L., Lopez L., Scalone A.G., Di Carli M., Desiderio A., Benvenuto E., and Perrotta G. Description of major proteins expressed in strawberry fruit during ripening. *ItPA* (The Italian Proteomic Association) 3rd Annual National Conference, Selva di Fasano (Brindisi), Italy, 11-14 **Giugno 2008**. Comunicazione orale O-Plant5.

Bianco L., Lopez L., Scalone A.G., Di Carli M., Desiderio A., Benvenuto E., and Perrotta G. Description of major proteins expressed in strawberry fruit during ripening. 2nd EU-Summer School in Proteomic Basics - 13-19 **Luglio 2008**- Bressanone Italy. Poster Abstract

Bianco L., Di Carli M., Scalone A.G., Benvenuto E., Desiderio A., and Perrotta G. Large-scale protein characterization of strawberry fruit. 51^o Annual Congress Società Italiana di Genetica Agraria Riva del Garda, 23-26 **Settembre 2007**, Poster Abstract D-79.

Carbone F., Rosati C., Bianco L., Di Carli M., Desiderio A., Benvenuto E., Perrotta G. Characterization of strawberry genotypes through a combined transcriptomics and proteomics approach. Proc I Congresso nazionale ItPA (The Italian Proteomic Association), Pisa, Italia, 2-4 **Luglio 2006**, comunicazione orale OP23.

2nd EU-Summer School in Proteomic Basics - 13-19 Luglio 2008-Brixen (Bressanone)

Ettan DIGE Training course – 18-21 Settembre 2007- GE Healthcare - Monaco di Baviera



**Relazione finale del Collegio dei docenti
sull'attività svolta dal Dott.ssa Linda BIANCO**

Nel corso del dottorato la dott.ssa Linda Bianco ha condotto l'attività di ricerca sul tema prescelto relativo all'Analisi proteomica del frutto di fragola durante la maturazione.

Nel periodo di formazione, la dottoranda ha acquisito una buona autonomia sviluppando capacità propositiva in parte derivante da una puntuale consultazione della letteratura scientifica del settore ed in parte da una approfondita competenza tecnica sull'utilizzo delle tecnologie applicate al lavoro sperimentale.

Il lavoro svolto dalla dottoranda ha riguardato lo studio dell'espressione del proteoma nel frutto di fragola durante la maturazione utilizzando differenti approcci sperimentali ad elevata processività come: 2D-LC/MS-MS, 1DE-LC/MS-MS e la tecnologia DIGE, per analisi differenziali accurate.

Questo esteso *screening* proteomico ha fornito numerose informazioni sulle dinamiche molecolari coinvolte nello sviluppo e maturazione del frutto e rappresenta, ad oggi, il primo lavoro *high-throughput* sulla fragola, dal punto di vista proteico.

La dott.ssa Bianco ha svolto il lavoro sperimentale presso l'ENEA CR Trisaia e alcuni risultati della sua ricerca sono stati sottoposti per la pubblicazione su rivista scientifica internazionale e presentati a convegni scientifici.

Sulla base della relazione orale della Dottoranda e del parere fornito dal tutor, Dott. Gaetano Perrotta, il Collegio dei docenti esprime pieno apprezzamento per i risultati conseguiti nel triennio dalla Dott.ssa Bianco.

A nome del Collegio dei Docenti
Il coordinatore del Corso di Dottorato


Prof. Aldo Musacchio