



UNIVERSITÀ DEGLI STUDI DI TRIESTE
SCUOLA DI DOTTORATO IN NANOTECNOLOGIE
XX CICLO

**Studio della risposta compatibile
di *Nicotiana benthamiana*
al *Turnip Vein-Clearing Virus***

Supervisore:

Dr. Fiorella FLORIAN

Dottorando:

René DREOS

Tutore:

Chiar.mo Prof. Alberto PALLAVICINI

Coordinatore:

Chiar.mo Prof. Maurizio FERMEGLIA

Indice

1	Introduzione	1
1.1	Meccanismi di resistenza all'infezione virale	2
1.1.1	Geni della resistenza	3
1.1.2	Attenuazione dell'RNA virale	6
1.2	Il <i>Turnip Vein-Clearing Virus</i> (TVCV)	10
1.2.1	Interazioni con l'ospite	13
1.3	Tecniche molecolari per lo studio delle infezioni virali	14
1.3.1	<i>Agrobacterium</i> come veicolo di DNA esogeno	15
1.3.2	<i>Microarray</i>	18
1.3.3	PCR quantitativa	23
2	Scopo della tesi	25
3	Materiali e metodi	26
3.1	Crescita delle piante ed inoculazione	26
3.2	Estrazione dell'RNA e delle proteine	27
3.2.1	Purificazione, quantizzazione e verifica qualitativa dell'RNA	29
3.3	Verifica dell'infezione	29
3.3.1	SDS-PAGE	29
3.3.2	PCR quantitativa	30
3.4	Sintesi del cDNA e marcatura del bersaglio	32
3.5	Costruzione del <i>microarray</i>	33
3.5.1	Annotazione delle sequenze	34

3.6	Esperimenti di ibridazione su <i>microarray</i> di cDNA	34
3.6.1	Disegno sperimentale	34
3.6.2	Preibridazione	35
3.6.3	Ibridazione	35
3.6.4	Lavaggi	36
3.6.5	Scansione dei <i>microarray</i>	37
3.7	Analisi dei dati	37
3.7.1	Analisi dei dati di <i>microarray</i>	37
3.7.2	Sviluppo del metodo <i>LOGIT</i> di analisi dei dati di qPCR	43
4	Risultati e Discussione	47
4.1	Crescita delle piante ed inoculazione	47
4.2	Verifica dell'infezione virale	48
4.2.1	Verifica <i>in vivo</i>	48
4.2.2	Verifiche <i>in vitro</i>	49
4.3	Verifica della qualità dell'RNA	51
4.4	Ibridazioni ed analisi dei dati	54
4.5	Interpretazione dei risultati	57
5	Conclusioni	64

Sommario

In agricoltura, la monocoltura su vaste aree geografiche può portare ad una rapida diffusione di malattie. D'altro canto, l'uso di composti chimici per controllarne la diffusione può causare seri problemi di inquinamento e aumentare i costi di produzione. Uno studio dettagliato delle interazioni pianta-patogeno può contribuire a fornire soluzioni sostenibili per il controllo delle malattie che colpiscono le specie coltivate. Per queste ragioni lo scopo principale di questa tesi è stato lo studio delle prime fasi della risposta di *N. benthamiana*, una pianta appartenente alla famiglia delle *Solanaceae*, all'infezione da parte del *Turnip Vein-Clearing Virus* (TVCV). Per farlo è stato messo a punto un metodo di infezione indiretto che sfrutta la capacità di *Agrobacterium* di inserire un frammento di DNA esogeno nel genoma della pianta. In questo modo la percentuale di cellule infettate ha raggiunto il 90% sul totale delle cellule del tessuto. L'analisi del profilo di espressione genica è stata effettuata mediante l'utilizzo di una nuova piattaforma *microarray* contenente circa 6000 sequenze specifiche di *N. benthamiana*. Le ibridazioni sono state eseguite durante il periodo in cui il genoma virale aumenta la propria concentrazione all'interno della cellula fino a raggiungere il valore massimo. Complessivamente sono state fatte 30 ibridazioni con il *microarray*. La determinazione della concentrazione del genoma virale è stata effettuata con l'uso della tecnica della PCR quantitativa per la quale è stato sviluppato un nuovo metodo di analisi dei dati. I risultati delle ibridazioni dimostrano che la risposta della pianta all'infezione virale è caratterizzata dalla presenza di due fasi distinte. Nella prima è presente un picco di risposta durante le prime fasi dell'esperimento, quando il genoma virale non ha ancora raggiunto la fase di crescita esponenziale. La seconda, invece, coincide con il massimo della concentrazione virale all'interno delle cellule. Analisi approfondite dei geni attivati nelle due fasi lasciano ipotizzare la presenza di due risposte differenti, la prima rivolta verso *Agrobacterium* mentre la seconda rivolta verso

l'infezione virale.

Questo lavoro è stato svolto in parte presso il laboratorio di Genetica dell'Università degli Studi di Trieste sotto la supervisione del Prof. Alberto Pallavicini ed in parte nel laboratorio del Prof. Andy Maule, Disease and Stress Biology, John Innes Centre, UK.

Capitolo 1

Introduzione

Le piante devono difendersi continuamente dall'attacco di batteri, virus, funghi, invertebrati e anche di altre piante. Dato che la loro immobilità ne preclude la possibilità di fuga, ogni cellula vegetale possiede una capacità di difesa sia costitutiva che inducibile. Ciò è in netto contrasto con il sistema immunitario dei vertebrati, in cui cellule specializzate deputate alla difesa sono rapidamente mobilitate verso il sito di infezione dove uccidono l'organismo invasore o ne limitano la diffusione. Tuttavia anche la strategia di difesa delle piante, che non prevede un sistema circolatorio, riduce al minimo le infezioni. In natura, la maggior parte delle piante è generalmente sana, per cui la presenza di malattie è di solito limitata a pochi individui ed interessa solo una piccola parte dei tessuti.

Uno studio dettagliato delle interazioni pianta-patogeno può contribuire a fornire soluzioni sostenibili per il controllo delle malattie che colpiscono le specie coltivate. Infatti, la monocoltura di specie geneticamente uniformi, coltivate su vaste aree geografiche, porta spesso ad una notevole diffusione di malattie. L'uso di composti chimici per controllarne la diffusione può causare seri problemi di inquinamento ed aumentare i costi di produzione. Inoltre, questi studi aiuterebbero a spiegare i meccanismi di trasmissione del segnale che le cellule vegetali utilizzano per far fronte ad una situazione di stress.

Per queste ragioni, lo scopo della mia attività di ricerca è stato quello di stu-

diare l'espressione genica nella fase iniziale della risposta difensiva di *Nicotiana benthamiana* al *Turnip Vein-Clearing Virus* (TVCV). A tale proposito è stato sviluppato un *microarray* a cDNA con sequenze specifiche di *N. benthamiana* provenienti da tre diverse librerie geniche. L'infezione è stata effettuata tramite l'uso di *Agrobacterium* come veicolo del genoma virale, il quale è stato appositamente modificato per essere inserito nel genoma vegetale. Questa procedura ha consentito di aumentare la percentuale di cellule infettate nella pianta. La risposta all'infezione è stata monitorata in tempi successivi (4, 4.5, 5, 5.5, 6 giorni dopo l'infezione) tramite ibridazioni con il *microarray*. L'infezione virale è stata monitorata mediante la tecnica della qPCR (*quantitative polymerase chain reaction*) per la quale è stato sviluppato un nuovo algoritmo di analisi dei dati.

1.1 Meccanismi di resistenza all'infezione virale

Il numero di specie virali che hanno come ospite una pianta è molto elevato: il database pubblico DPV [<http://www.dpvweb.net>] ha classificato 21 famiglie, divise in 86 generi a cui appartengono più di 1000 specie. D'altro canto, una specie vegetale può venire infettata solamente da un numero ristretto di virus, mentre è immune all'infezione di altri. Quindi, le relazioni che intercorrono tra piante e patogeni sono state divise in due categorie: resistenza e suscettibilità [Kang et al., 2005]. La prima comprende il caso in cui tutti i membri di una specie vegetale sono resistenti o non sono infettati da tutti i membri appartenenti ad una specie virale. In questo caso la pianta viene definita non-ospite del virus. L'impossibilità di infettare un non-ospite è normalmente dovuta alle difese di base della pianta che includono barriere fisiche alle infezioni come la parete cellulare, la cuticola, o la produzione di numerosi composti ad azione antimicrobica [Thordal-Christensen, 2003]. Se, d'altra parte, un microorganismo riesce ad infettare e replicarsi in alcuni membri appartenenti ad una specie vegetale, questa viene detta ospite o

suscettibile. A differenza della resistenza, la suscettibilità è analizzabile geneticamente, caratteristica che la rende molto più interessante dal punto di vista biologico. In questo caso il patogeno viene definito virulento e l'interazione compatibile [Soosaar et al., 2005].

Per contrastare l'infezione e limitarne la diffusione, la pianta deve essere in grado di riconoscere il patogeno e di reagire modificando il proprio metabolismo in modo da adattarsi al cambiamento. Nel sito di infezione deve attivare la trascrizione dei geni importanti per la resistenza diretta verso il virus [Tameling and Takken, 2007], mentre i tessuti circostanti devono attivare un altro *set* di geni che ne limitino la diffusione (SAR) [Heil and Bostock, 2002]. Lo scopo di questa ricerca è lo studio dettagliato delle prime fasi della risposta della pianta nel sito di infezione.

1.1.1 Geni della resistenza

La capacità di distinguere tra il *self* ed il *non-self* è l'aspetto più importante del sistema immunitario. Sia nelle piante che negli animali il riconoscimento degli organismi invasori avviene ad opera di recettori di membrana ed intracellulari. Tutte le cellule delle piante hanno dei recettori transmembrana capaci di riconoscere strutture molecolari tipiche dei batteri (PAMPs da *pathogen-associated molecular patterns*) come frammenti della parete cellulare, la cuticola o la flagellina [Ausubel, 2005; Chisholm et al., 2006]. Il riconoscimento dei PAMP da parte dei recettori transmembrana (recettori PAMP) porta all'induzione di una risposta di difesa di base denominata PTI (da *PAMP-triggered immunity*) [Jones and Dangl, 2006]. Tuttavia numerosi patogeni delle piante, tra cui i virus, producono proteine all'interno delle cellule dell'ospite. Per combattere questo tipo di infezioni, le piante hanno evoluto un meccanismo di riconoscimento intracellulare legato ad una risposta di difesa amplificata da una cascata di secondi messaggeri [Jones and Dangl, 2006]. Questo secondo tipo di resistenza viene detto ETI (da *effector-triggered immunity*) ed è mediato dai geni della resistenza (*geni R*).

Struttura proteica dei geni R

La classe delle proteine R è molto eterogenea, per questa ragione è stata divisa in differenti famiglie a seconda della loro struttura proteica. Ad esempio in *Arabidopsis thaliana* se ne contano più di 150 e in *Oryza sativa* almeno 400 [Meyers et al., 2003; Monosi et al., 2004]. Sono tra le proteine più grandi del regno vegetale le cui dimensioni variano tra gli 860 e i 1900 aminoacidi [Mchale et al., 2006]. Un aspetto comune sia dei recettori transmembrana PAMP che dei geni R è il dominio LRR (da *Leucine-Rich Repeat*) [Nurnberger and Kemmerling, 2006]. Esso è caratterizzato dalla presenza di un motivo particolarmente ricco in leucine con sequenza consenso di 14 residui [LxxLxxLxLxxC/Nxx (L, leucina; x, qualsiasi aminoacido; C, cisteina; N, asparagina)] [Kajava, 1998]. Nella maggior parte delle proteine R è presente un secondo dominio NB (da *Nucleotide Binding*) di legame con i nucleotidi. Poiché strutture molto simili si ritrovano in altre proteine che legano l'ATP o il GTP, la trasmissione del segnale potrebbe essere molto simile a quella della superfamiglia delle proteine con attività GTPasica [Traut, 1994]. In questi polipeptidi il dominio di legame con i nucleotidi fa parte di una entità più grande denominata dominio NB-ARC data la sua presenza nella proteina Apaf-1 (una proteina che induce la morte cellulare programmata in *Caenorhabditis elegans* [van der Biezen and Jones, 1998]).

Il dominio LRR è spesso caratterizzato dalla presenza di aminoacidi a sequenza variabile che conferiscono un'alta variabilità strutturale alle proteine R. Di conseguenza questa regione è importante per il riconoscimento delle proteine della virulenza codificate dal patogeno. Infatti, ogni proteina codificata da un virus può essere considerata un determinante della virulenza e un attivatore delle difese immunitarie grazie all'azione di una specifica proteina R. Nonostante siano stati clonati numerosi geni R [De Nardi et al., 2006], non sono stati fatti progressi significativi nel capire come avvenga il riconoscimento delle proteine del patogeno. Inizialmente era stato proposto un modello del tipo recettore-ligando, nel quale una proteina R è capace di riconoscere solamente una particolare proteina della virulenza [Martin et al.,

2003]. Tuttavia, questo modello molto semplice si basa su un unico esempio in cui una proteina batterica e una proteina fungina sono riconosciute direttamente dalle corrispettive proteine R [Martin et al., 2003]. In un modello più sofisticato viene proposta la presenza di un complesso multiproteico contenente una proteina R [Van Der Biezen and Jones, 1998]. Questo “modello sentinella” afferma che le proteine R (sentinelle) sono associate costitutivamente a proteine cellulari dell’ospite importanti per l’infezione del patogeno. Durante l’infezione, il patogeno modifica alcune proteine dell’ospite che vengono percepite dalle sentinelle. Ogni modifica che altera la struttura quaternaria di queste proteine ha come effetto il rilevamento del patogeno. Ciò attiva le sentinelle che iniziano una cascata di segnali intracellulari che culminano nella risposta di difesa [Van Der Biezen and Jones, 1998].

Dopo aver rilevato la presenza di un patogeno, le cellule vegetali attuano delle risposte precoci che comprendono cambiamenti nei flussi ionici, attivazione di vie metaboliche di secondi messaggeri (specialmente le chinasi), elevata alterazione del profilo di trascrizione, produzione di specie reattive dell’ossigeno (ROS, da *reactive oxygen species*) e la produzione di ossido di azoto (NO). Tali cambiamenti portano la cellula vegetale ad una morte programmata (PCD, da *Programmed Cell Death*) che in questo caso prende il nome di risposta ipersensitiva (HR, da *Hypersensitive Response*).

Diversi studi riportano un cambiamento nella concentrazione degli ioni calcio (Ca^{2+}) nel sito di infezione [Ma and Berkowitz, 2007; Romeis et al., 2000]. Cambiamenti nei flussi ionici sono considerati attivatori delle chinasi, ad esempio il legame del Ca^{2+} con proteine chinasi calcio-dipendenti attivano una cascata di fosforilazioni [Navazio et al., 2007; Romeis et al., 2000]. Questa cascata è importante per la trasduzione del segnale durante la difesa e per l’alterazione dell’espressione di un grande numero di geni. Analisi con *microarray* hanno evidenziato grossi cambiamenti nell’attività trascrizionale in cellule e tessuti infettati. Per esempio, in *A. thaliana* infettata con il *Cucumber Mosaic Virus* (CMV) si registra l’alterazione trascrizionale di 444 geni [Marathe et al., 2004].

Le molecole ROS prodotte durante la HR, oltre ad attivare la PCD, attivano anche la trascrizione di geni della resistenza e molto probabilmente contribuiscono all'ispessimento della parete cellulare favorendo i legami intramolecolari [Delledonne et al., 1998]. Anche la molecola NO è richiesta durante la HR. Questo composto induce l'espressione di geni della resistenza [Durner et al., 1998], anche se la funzione di alcuni di questi nella risposta all'infezione virale è ancora poco chiara. In *A. thaliana* l'enzima *AtNOS1* (*NITRIC OXIDE SYNTHASE 1*) catalizza la produzione di ossido di azoto indotta dall'acido abscissico, un ormone importante nei meccanismi di difesa [Guo et al., 2003]. Mutanti *atnos1* mostrano una resistenza ridotta verso infezioni batteriche [Zeidler et al., 2004].

1.1.2 Attenuazione dell'RNA virale

Anche se il meccanismo dell'attenuazione dell'RNA (dall'inglese RNA *silencing*) è diventato una delle nuove frontiere della biologia, il primo articolo che ne descrive gli effetti risale agli inizi del 1900 [Wingard, 1928]. In questo articolo, infatti, S. A. Wingard descrive che nelle piante di tabacco soltanto le foglie inizialmente infettate risultano necrotiche e presentano i sintomi del *Ringspot Virus*. Le foglie superiori invece diventano in qualche modo immuni al virus e conseguentemente sono asintomatiche e resistenti ad una seconda infezione. A quel tempo questa "guarigione" era un mistero: non esisteva alcun modo razionale per spiegare la specificità ad una seconda infezione. I dettagli dell'infezione da *Ringspot Virus* devono ancora essere chiariti, ma adesso è noto che la risposta della pianta coinvolge il meccanismo dell'attenuazione dell'RNA diretto specificatamente verso l'RNA virale [Ratcliff et al., 1997]. In aggiunta alla protezione della pianta alle infezioni virali, questo processo si è diversificato in meccanismi che proteggono il genoma dai trasposoni e regolano la trascrizione genica.

Fasi iniziali dell'attenuazione

P. M. Waterhouse in un articolo del 1998 suggerisce che la causa scatenante il meccanismo del *silencing* sia l'RNA a doppia elica (dsRNA) [Waterhouse et al., 1998]. Infatti, molti virus delle piante (58 generi su 80) sono virus a RNA che necessitano di una RNA polimerasi RNA dipendente (RdRP) per potersi replicare. La RdRP è codificata dal genoma virale e, nella prima fase della replicazione, produce copie complementari al genoma virale. Waterhouse suggerisce che questo meccanismo di replicazione generi una grande quantità di filamenti di dsRNA che scatenano l'attenuazione dell'RNA. La sequenza di RNA omologa al dsRNA è degradata e conseguentemente l'infezione virale è efficientemente attenuata. Tuttavia, sia i virus delle piante che quelli degli animali riducono l'accessibilità dei loro intermedi replicativi a doppia elica ai meccanismi di difesa dell'ospite. Ad esempio, nel virus dell'influenza umana gli intermedi di dsRNA sono rivestiti dalle proteine del nucleocapside che lo proteggono dalla degradazione [Klumpp et al., 1997]. Più recentemente Paul Ahlquist ha proposto che l'RdRP virale produca molecole di RNA che possano attivare la trascrizione di famiglie geniche codificanti RdRP presenti nel genoma della pianta [Ahlquist, 2002]. Sarebbero questi enzimi i responsabili della formazione delle molecole di dsRNA che attivano i meccanismi di attenuazione del genoma virale. Sembra inoltre che i membri di questa famiglia differiscano per il legame a RNA virali specifici. In *A. thaliana*, la mutazione di uno di questi geni (RDR6, da *RNA-Dependent RNA polymerase 6*) rende la pianta ipersensibile all'infezione da parte del *Cucumber Mosaic Virus* ma non all'infezione da parte del *Tobacco Rattle Virus* e del *Tobacco Mosaic Virus* (TMV). Invece in *Nicotiana tabacum* la soppressione dell'espressione del gene RDR1 causa ipersensibilità al TMV [Baulcombe, 2004].

Meccanismi dell'attenuazione

Anche se il meccanismo di innesco dell'attenuazione dell'RNA non è stato ancora chiarito in dettaglio, gli eventi che ne derivano sono ben studiati. Le

lunghe molecole di dsRNA formate nelle fasi iniziali della replicazione virale vengono tagliate dall'enzima DICER (una ribonucleasi di tipo III) in frammenti di 21–26 nucleotidi in un processo ATP-dipendente denominati siRNA (*small interfering RNA*) [Meister and Tuschl, 2004]. Le piante possiedono diversi omologhi della nucleasi DICER (enzimi DICER-LIKE), ognuno dei quali è responsabile della degradazione di uno specifico genoma virale. Nei mutanti di *A. thaliana dcl1*, *dcl2* and *dcl3*, l'accumulo di siRNA virale a seguito di un'infezione da parte del *Turnip Mosaic Virus* (TuMV) e *Cucumber Mosaic Virus* (CMV) è simile a quella del fenotipo selvatico [Xie et al., 2004]. Tuttavia, in piante *dcl2*, l'accumulo di siRNA a seguito dell'infezione da parte del *Turnip crinkle virus* (TCV) è ritardato, aumentandone la suscettibilità. I frammenti di RNA a doppia elica creati dall'enzima DICER vengono separati probabilmente da una elicasi DEAD-box [Gong et al., 2005], creando frammenti di RNA a singola elica. Questi sono immediatamente riconosciuti dalla proteina *ARGONAUTE* che fa parte del complesso multiproteico RISC (*RNA-induced silencing complex*). Altri componenti proteici del RISC includono membri della famiglia *ARGONAUTE*, nucleasi ed altri fattori [Hannon, 2002]. Il RISC ha la capacità di legare e di tagliare il genoma virale complementare all'siRNA [Meister and Tuschl, 2004].

È interessante notare che il meccanismo di attenuazione coinvolge anche geni dell'ospite. Infatti molti siRNA hanno come bersaglio fattori di trascrizione importanti per lo sviluppo morfologico e per la divisione cellulare [Park et al., 2002]. Mutazioni di *A. thaliana* che alterano la biogenesi o l'attività dei siRNA causano difetti pleiotropici durante lo sviluppo embrionale, vegetativo e riproduttivo.

Un meccanismo così efficace per la protezione contro le infezioni virali non poteva non essere soggetto all'evoluzione di meccanismi che ne limitino l'azione. Recentemente sono state scoperte numerose proteine codificate dai genomi virali che bloccano in diversi punti la via metabolica dell'*RNA silencing* [Chapman et al., 2004]. La funzione principale dei soppressori dell'attenuazione è la virulenza. Ad esempio il TuMV codifica la proteina HC-Pro (dall'inglese

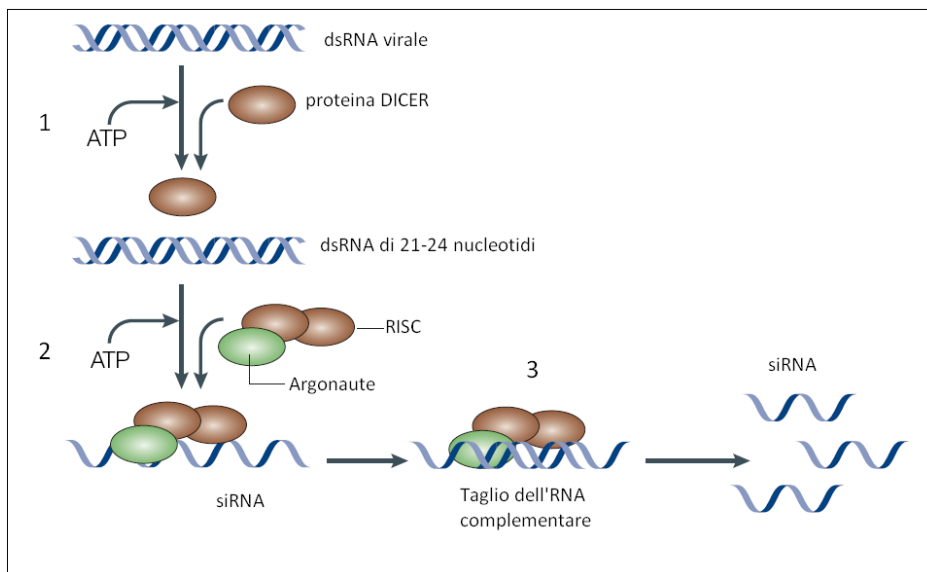


Figura 1.1: Schema del meccanismo di attenuazione dell'RNA virale. L'RNA a doppia elica (probabilmente prodotto da un enzima RdRP dell'ospite) viene riconosciuto dall'enzima DICER [1] e tagliato in frammenti di 21 – 24 nucleotidi. I frammenti di dsRNA vengono spaiati e legati dalla proteina *ARGONAUTE*, facente parte del complesso multiproteico RISC [2]. Il legame del siRNA con il genoma virale ne consente il taglio e l'attenuazione da parte del RISC.

helper-competent protein) che è un determinante della virulenza e anche un soppressore del *silencing*. La funzione dell'HC-Pro di soppressione potrebbe spiegare i casi di infezioni sinergiche che si riscontrano spesso tra i *potivirus* [Anandalakshmi et al., 1998]. Molti virus che non sono capaci di infettare una pianta, acquisiscono la capacità infettiva in sinergia con l'infezione di altri *potivirus*. Questa proteina è anche la causa di numerosi difetti nello sviluppo della pianta, inclusi alcuni simili a quelli prodotti dalla mutazione *dcl1*. Mutazioni della HC-Pro causano la perdita della capacità del virus di replicare e diffondere [Kasschau et al., 2003].

La P19 è una proteina prodotta dal *Tomato Bushy Stunt Virus* ed è il determinante della soppressione dell'attenuazione. Questa proteina lega le molecole di siRNA create dall'enzima DICER, sequestrandole al legame con il complesso multiproteico RISC e proteggendole dalla sua azione di degradazione [Ye et al., 2003]. Altri esempi di proteine virali coinvolte nei meccanismi anti-attenuazione sono la P21 del *Beet Yellow Virus* e la proteina del capsido del *Turnip Crinkle Virus* [Chapman et al., 2004; Brigneti et al., 1998].

1.2 Il *Turnip Vein-Clearing Virus* (TVCV)

Dal momento che il *Turnip Vein-Clearing Virus* (TVCV) ha numerose caratteristiche in comune con il *Tobacco Mosaic Virus* (TMV), il capostipite del genere *Tobamovirus*, esso è stato classificato nello stesso genere. Infatti come per il TMV, anche nel TVCV le particelle virali hanno una forma a bastoncino e l'unica proteina che forma il capsido ha un peso molecolare di 17 kDa. Il capsido racchiude un genoma formato da una singola molecola di RNA di senso positivo lunga circa 6300 nucleotidi. Il suo genoma è leggermente più corto di quello del TMV [Melcher, 2003]. L'organizzazione del genoma ad RNA del TVCV è simile a quella delle altre specie appartenenti al genere *Tobamovirus* (Figura 1.2a). Il genoma ha la classica struttura 5'-cap che la rende molto simile agli RNA messaggeri dell'ospite e ne consente una più efficace traduzione [Efimov et al., 2001]. L'RNA contiene l'informazione

per almeno quattro diversi polipeptidi [Dorokhov et al., 1994; Lartey et al., 1995]. La ORF (dall'inglese *Open Reading Frame*) al 5'-terminale codifica per un polipeptide di 125 kDa, che è leggermente più piccolo dell'omologo in TMV. Essa contiene dei motivi che sono riconducibili alla classe proteica delle guanine metiltransferasi e delle RNA elicasi [Harrell et al., 2002]. Frequentemente il codone di stop di questo polipeptide non viene riconosciuto portando alla formazione di una proteina più lunga contenente un dominio tipico delle RNA replicasi [Harrell et al., 2002].

La parte 3'-terminale del genoma virale codifica per due proteine, la proteina del movimento MP (*Movement Protein*) e la proteina del capsido CP (*Capsid Protein*). La proteina MP, un membro della superfamiglia delle proteine del movimento 30 K [Melcher, 2000], è richiesta per il movimento delle particelle virali dal sito di infezione alle cellule adiacenti ma anche per la diffusione dell'infezione a lunga distanza sia in *Brassica rapa* che in *Arabidopsis thaliana* [Melcher, 2003]. Una piccola ORF che si sovrappone alla parte 3'-terminale della regione che codifica la proteina del movimento si ritrova frequentemente nei virus appartenenti ai *Tobamovirus* [Morozov et al., 1993]. È stato suggerito che codifichi per un polipeptide funzionale durante l'infezione di *N. benthamiana* ma non di *A. thaliana*. Tuttavia, anche se i tentativi di darle un ruolo nelle infezioni di altre specie non hanno dato buoni risultati, è stato dimostrato che lega il fattore di trascrizione EF-1 α (*Elongation Factor 1 α*) [Fedorkin et al., 1995]. Questo gene potrebbe non essere funzionale data l'assenza del codone di inizio AUG, ma sono possibili codoni di inizio alternativi.

Il ciclo vitale del virus comincia con l'entrata in una cellula vegetale (Figura 1.2b). Qui le proteine del capsido si disassemblano lasciando libero il genoma a RNA di interagire con i ribosomi presenti nella cellula vegetale. In queste prime fasi viene prodotta la RNA replicasi che è indispensabile per produrre copie del genoma virale a senso negativo (RNA-) da usare come stampo per la replicazione. L'RNA- viene usato anche per la creazione di frammenti di RNA più piccoli, sgRNA (*subgenomic RNA*), in modo da consentire la

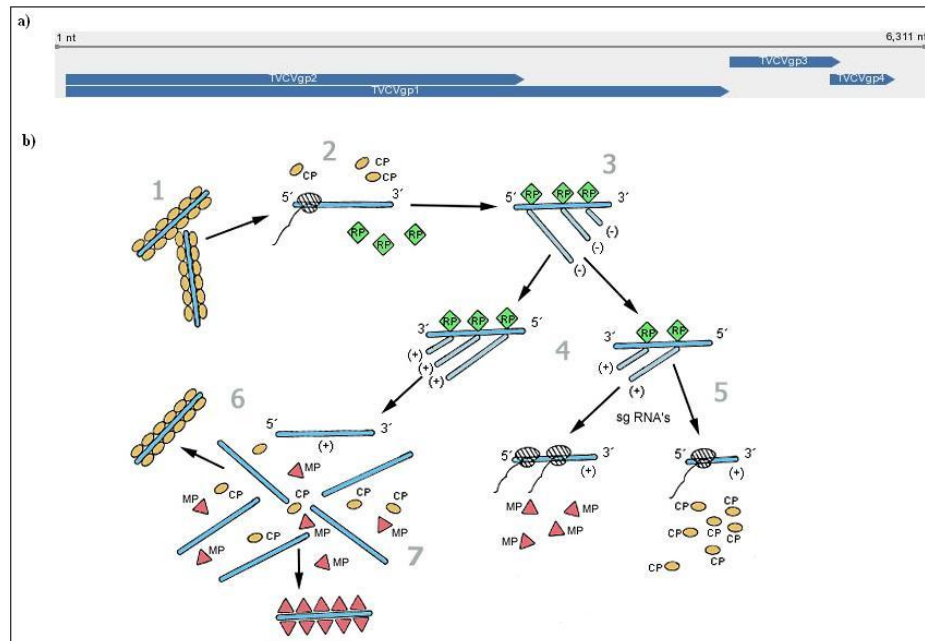


Figura 1.2: **a)** Rappresentazione schematica del genoma del *Turnip Vein-Clearing virus*. TVCVgp1: componente allungata dell'RNA replicasi; TVCVgp2: componente dell'RNA replicasi (proteina da 125 kDa); TVCVgp3: proteina del movimento; TVCVgp4: proteina del capside. Tratto da: <http://www.ncbi.nlm.nih.gov/>. **b)** Ciclo di replicazione del TVCV. Il virus entra nella cellula vegetale [1]. Quando le proteine del capside (CP) si staccano dall'RNA [2] i ribosomi dell'ospite cominciano la traduzione delle RNA replicasi (RP). Queste proteine sono utilizzate per produrre copie complementari del genoma virale (RNA-) [3]. L'RNA- viene usato come stampo sia per la fabbricazione di molecole a lunghezza completa del genoma di TVCV (RNA+), che di frammenti più corti (sgRNA) [5] che sono usati per esprimere le proteine del movimento (MP) e del capside (CP). L'RNA+ lega sia la CP per la creazione di nuove particelle virali [6] che la MP per muoversi verso un'altra cellula dove il ciclo ricomincia [7].

traduzione della proteina MP e della proteina CP. La proteina MP lega le molecole del genoma virale e le dirige verso le cellule adiacenti dove il ciclo di infezione può ricominciare. La proteina CP invece si lega alle molecole del genoma virale per creare nuove particelle virali [Buck, 1999].

1.2.1 Interazioni con l'ospite

La peculiarità del TVCV di infettare la pianta modello *A. thaliana* lo ha reso molto comune in studi di interazione pianta-patogeno. I virus devono sfruttare i meccanismi cellulari dell'ospite per potersi replicare, per tale ragione componenti virali devono interagire con molecole dell'ospite durante tutte le fasi del processo di infezione. Yamanaka in due articoli del 2000 e del 2002 mette in luce l'importanza delle due proteine TOM1 e TOM3 (da *TOBAMOVIRUS MULTIPLICATION locus*) di *A. thaliana* per la replicazione del TVCV [Yamanaka et al., 2000, 2002]. Mutanti di *A. thaliana* di questi due geni diventano resistenti all'infezione da parte del TVCV o di virus a lui strettamente correlati. In particolare, *TOM1* codifica per una proteina di membrana che, assieme alla proteina TOM2A, forma un complesso multiproteico localizzato nella membrana del tonoplasto [Tsujimoto et al., 2003]. È in questa specifica regione cellulare che l'RNA virale forma dei complessi con altre proteine durante la replicazione [Hagiwara et al., 2003]. In un articolo pubblicato nel 2007, Chen dimostra che un gene analogo a *TOM1* è necessario per la corretta replicazione del TMV in *N. benthamiana* [Chen et al., 2007].

La proteina del movimento MP, un prodotto virale delle prime fasi della replicazione, lega l'RNA virale e altre quattro proteine dell'ospite [Ivanov et al., 1994]. La proteina MP è un substrato per le chinasi dell'ospite ed infatti la sua attivazione avviene tramite fosforilazione. Ciò consente il legame con l'RNA e la successiva diffusione sistemica del virus [Karpova et al., 1999]. Il legame con l'RNA sembra non essere sequenza-specifico dal momento che è in grado di riattivare l'infezione del *Potato virus X* nel quale tale proteina era stata resa non funzionale [Morozov et al., 1997].

Uno dei meccanismi più efficaci per contrastare l'infezione virale resta comunque l'*RNA silencing*. L'efficacia di tale meccanismo lo ha reso bersaglio di diversi virus che ne limitano la sua azione. Ad esempio nel *Tomato Mosaic Virus* l'RNA replicasi è la proteina che contrasta il *silencing* [Kubota et al., 2003]. Una proteina analoga non è stata ancora caratterizzata in TVCV, anche se è stata dimostrata la sua esistenza [Vaucheret et al., 2001].

1.3 Tecniche molecolari per lo studio delle infezioni virali

Le basi scientifiche per questa ricerca derivano dall'osservazione che il cambiamento dell'espressione genica di una pianta a seguito dell'infezione da parte di un virus è sotto stretto controllo temporale e spaziale, quando questa viene misurata come cambiamento dell'espressione genica o della fisiologia dell'ospite. Nel sito di infezione la pianta attiva una risposta che contrasta attivamente il diffondersi della malattia (risposta ipersensitiva, HR). A livello sistemico invece la risposta è incentrata sulla prevenzione della malattia (risposta SAR). Quando un virus entra in una cellula vegetale il suo genoma si replica in modo molto veloce (5-10 ore) se confrontato con il tempo di giorni che impiega per diffondersi a tutte le cellule del tessuto circostante e successivamente a tutta la pianta (infezione sistemica). La scelta di monitorare la risposta all'infezione campionando foglie infettate risulta inadeguata per tre ragioni. La prima è che, in foglie infettate, la proporzione di cellule a diretto contatto col virus è estremamente bassa, con un conseguente mascheramento della risposta di difesa. La seconda ragione è che, il virus diffonde in modo lineare attraverso il tessuto, quindi le foglie campionate rappresentano una media della risposta temporale mediata dalle singole cellule. La terza è che se si campionano delle foglie con un'infezione ormai ben stabilita, si ha una predominanza di cellule con una risposta di tipo tardivo, perdendo così l'informazione delle prime fasi di risposta.

Per aumentare la percentuale di cellule infettate si è reso necessario l'utilizzo

di una tecnica di infezione indiretta che sfrutta *Agrobacterium* per trasportare il genoma virale all'interno della cellula dell'ospite [Marillonnet et al., 2005]. Non è stato possibile, invece, sincronizzare l'infezione virale, anche se l'aumento della percentuale di cellule infettate ha permesso di aumentare la percentuale di cellule sincronizzate.

1.3.1 *Agrobacterium* come veicolo di DNA esogeno

Agrobacterium è un genere a cui appartengono numerose specie di batteri che hanno la capacità di infettare un elevato numero di piante. Tuttavia, questa divisione si basa sui sintomi provocati dai batteri nei rispettivi ospiti. Ad esempio, *A. radiobacter* è una specie non virulenta, *A. tumefaciens* provoca tumori apicali, *A. rhizogenes* causa la malattia delle radici pelose e *A. rubi* causa tumori nel fusto [Otten et al., 1984]. Questa classificazione non è del tutto corretta dal momento che si basa sulla sintomatologia che non è causata dal genoma batterico ma da un anello di DNA accessorio denominato plasmide. Togliere un plasmide e sostituirlo con un altro può alterare i sintomi della malattia. Ad esempio, l'infezione di una pianta con *A. tumefaciens* C58, contenente il plasmide pTiC58 (Ti, *Tumor-inducing*), causa la formazione di tumori nella zona apicale. Se il batterio viene privato del plasmide, questo diventa non patogenico e quindi può essere classificato come *A. radiobacter*. Introducendovi un plasmide Ri (da *Root-inducing*) il batterio diventa rizogenico come *A. rhizogenes* [White and Nester, 1980]. A questo punto il termine specie diventa fuorviante ed allora è stata proposta una classificazione che divide il genere *Agrobacterium* in *biovars* basati sulle caratteristiche metaboliche e di crescita di ogni ceppo [Costantino et al., 1980]. Usando questo sistema, molti ceppi di *A. tumefaciens* e *A. rubi* appartengono al biovar I mentre *A. rhizogenes* appartiene al biovar II [Tighe et al., 2000].

Una caratteristica unica di questo genere è la possibilità di trasferire una parte del proprio corredo genetico all'organismo ospite [McCullen and Binns, 2006]. In effetti, questo è l'unico caso osservato in natura di trasferimento di

DNA da una specie ad un'altra. I biologi hanno cominciato a studiare questo microorganismo per utilizzarlo come veicolo di frammenti di DNA all'interno delle piante. Ed in effetti *Agrobacterium* è molto utilizzato per la produzione di piante transgeniche [Tzfira and Citovsky, 2006; Gelvin, 2003].

Questo microorganismo ha la possibilità di trasferire il DNA presente in una specifica regione del suo genoma. Questa particolare regione di DNA detta T-DNA (*Transferred DNA*) non è localizzata nel cromosoma del batterio ma nel plasmide. Normalmente essa contiene l'informazione per la codifica di proteine che inducono tumori nelle cellule infettate consentendo la proliferazione dell'infezione. È possibile però sostituire questa regione di DNA con una a sequenza prestabilita senza pregiudicarne il trasferimento nell'ospite. In questo modo si può trasferire stabilmente alle piante una sequenza di DNA contenente l'informazione per la codifica di una o più proteine. È allora possibile creare in modo semplice un organismo transgenico.

Durante il lavoro di dottorato ho sfruttato questa caratteristica per trasferire in *N. benthamiana* l'intero genoma di TVCV per facilitarne l'infezione e per aumentare la percentuale di cellule infettate [Marillonnet et al., 2005].

Proteine coinvolte nel processo di trasformazione

La regione *vir* presente nel plasmide Ti di *Agrobacterium* codifica per molte proteine della virulenza (Vir) usate dal battere per produrre il T-DNA e per trasportarlo all'interno della cellula ospite. Il processo di trasformazione incomincia con l'adesione del battere alla parete cellulare dell'ospite (Figura 1.3), seguita dall'induzione dei geni della regione *vir* da parte di segnali chimici prodotti dall'ospite. A questo punto una molecola di T-DNA a singolo filamento (*T-strand*) viene prodotta dall'azione combinata delle proteine batteriche VirD1 e VirD2 [Filichkin and Gelvin, 1993]. Nella cellula batterica il T-DNA esiste come complesso formato dal DNA a singola elica e dalla proteina VirD2 che si attacca covalentemente all'estremità 5' del T-DNA [Vergunst et al., 2000]. Questo complesso, assieme a numerose altre proteine Vir, viene esportato nella cellula ospite tramite l'azione di un canale che col-

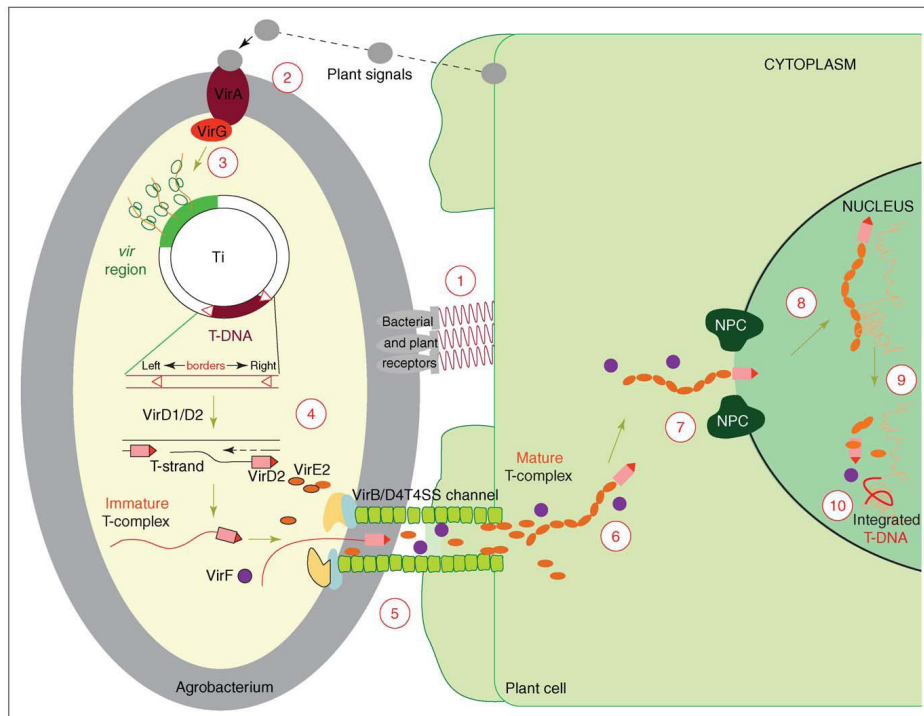


Figura 1.3: Schema del meccanismo di trasformazione mediato da *Agrobacterium*. Proteine presenti sulla membrana del batterio riconoscono e legano proteine della parete cellulare dell'ospite [1]. Messaggeri extracellulari dell'ospite [2] attivano i geni della regione *vir* (VirD1 e VirD2) presente nel plasmide Ti [3]. L'azione delle proteine VirD1 e VirD2 stacca il frammento di T-DNA dal plasmide. Nel citoplasma il T-DNA si trova legato alla proteina VirD2 [4] che lo indirizza verso il canale di collegamento con la cellula ospite [5]. Nel citoplasma dell'ospite il T-DNA viene ricoperto dalle proteine VirE2 a formare il complesso T (*T-complex*) [6] il quale viene trasportato verso il nucleo [7]. Sia la proteina VirE2 che la proteina VirD2 interagiscono con le proteine di trasporto nel nucleo dell'ospite favorendone l'entrata [8]. A questo punto il complesso T viene disassemblato [9] e il T-DNA si integra con il genoma dell'ospite [10] [Tzfira and Citovsky, 2006]

lega le due cellule formato dalle proteine VirB/D4 e da almeno una proteina dell'ospite [Christie, 2004; Hwang and Gelvin, 2004]. Si ritiene che, una volta all'interno del citoplasma della cellula ospite, il T-DNA esista come un complesso T maturo (*T-complex*), nel quale la molecola di T-DNA viene ricoperta per l'intera lunghezza da numerose proteine VirE2. Esse conferiscono al T-DNA la struttura e la protezione necessaria per poter entrare nel nucleo dell'ospite [Abu-Arish et al., 2004; Citovsky et al., 1989]. Sembra che il complesso VirE2-ssDNA utilizzi il citoscheletro della cellula ospite per raggiungere il nucleo. L'organizzazione del citoscheletro, orientato con la sua estremità negativa verso il nucleo, supporta l'idea che *Agrobacterium* utilizzi un motore proteico simile a quelli presenti negli eucarioti (*dynein-like*) per trasportare il T-DNA in prossimità dei pori nucleari [Salman et al., 2005]. Le grosse dimensioni del complesso T maturo, suggeriscono la presenza di un meccanismo attivo per l'entrata nel nucleo, probabilmente ad opera dei trasportatori nucleari. Infatti entrambe le proteine che formano il complesso T, VirD2 e VirE2, interagiscono con proteine dell'ospite per il loro trasporto nel nucleo [Ballas and Citovsky, 1997]. Una volta entrato nel nucleo il complesso T viene disassemblato e il T-DNA è libero di integrarsi nel genoma dell'ospite dove comincia la trascrizione dei propri mRNA.

1.3.2 *Microarray*

Il profilo di espressione genica è l'insieme dei geni attivati in un tessuto in un particolare momento o condizione. È possibile misurare le differenze tra diverse tipologie tissutali, tra tessuti diversamente trattati o tra un tessuto sano e un tessuto infettato da patogeno, analizzandone l'espressione genica. Il *microarray* è una tecnologia utilizzata per confrontare i livelli di espressione genica in diversi campioni, intendendo come espressione genica la quantità di DNA trascritto in RNA messaggero. Esso è caratterizzato dalla presenza di molecole di DNA a singola elica (molecole sonda) ordinatamente disposte in punti sulla superficie di un vetrino [Barrett and Kawasaki, 2003]. Ogni punto contiene una sola specie di DNA dalla sequenza nucleotidica uguale a

quella di una porzione di un gene. Queste molecole forniscono il substrato per l'ibridazione con le molecole di cDNA a singola elica provenienti dalla retrotrascrizione dell'RNA messaggero di due campioni differenti (molecole bersaglio). È quindi possibile determinare il livello di espressione dei singoli geni per confronto diretto tra l'abbondanza relativa di RNA prodotto nei due campioni. Per effettuare tale confronto, le molecole dei due campioni vengono marcate con molecole fluorescenti diverse e fatte reagire contemporaneamente con quelle legate alla superficie di silicio mediante un'ibridazione competitiva, creando un punto con intensità luminosa proporzionale alla quantità, e dunque all'attività, del gene relativo [Xiang and Chen, 2000]. L'intensità relativa registrata è proporzionale alla quantità del trascritto specifico nel campione biologico preso in esame. Uno schema generale della tecnica del *microarray* è rappresentato in Figura 1.4. Esistono due tipi di *microarray* basati sull'ibridazione specifica di sonda e bersaglio: i) gli acidi nucleici vengono pre-sintetizzati e poi depositati sul vetrino da un robot; ii) gli oligonucleotidi-sonda vengono sintetizzati direttamente sul vetrino tramite varie tecniche come quella fotolitografia (Affymetrix), quella MAS (*Maskless Array Synthesizer*, Nimblegen) e quella elettrochimica (CombiMatrix). La marcatura fluorescente delle molecole bersaglio può essere diretta o indiretta. Nella marcatura diretta il cDNA a singola elica viene marcato durante il processo di retrotrascrizione grazie all'incorporazione di un nucleotide modificato legante un fluorocromo. Tale metodica è caratterizzata da costi elevati, da una bassa resa d'incorporazione, in quanto l'enzima retrotrascrittasi fatica ad incorporare i nucleotidi visto il grosso ingombro sterico dato dal fluorocromo. Inoltre, il cDNA marcato in questo modo si degrada rapidamente. Nella marcatura indiretta, nucleotidi aminomodificati (aminoallil-dUTP o aminoexil-dUTP) vengono incorporati nel cDNA in via di sintesi, e, in un secondo momento, i fluorocromi legano chimicamente tali gruppi amminici, marcando il cDNA [Naderi et al., 2004; 't Hoen et al., 2003]. Le molecole fluorescenti più comunemente usate sono la cianina 3 (*Cy3*) e la cianina 5 (*Cy5*), le quali assorbono ed emettono la luce rispettivamente alle lunghezze

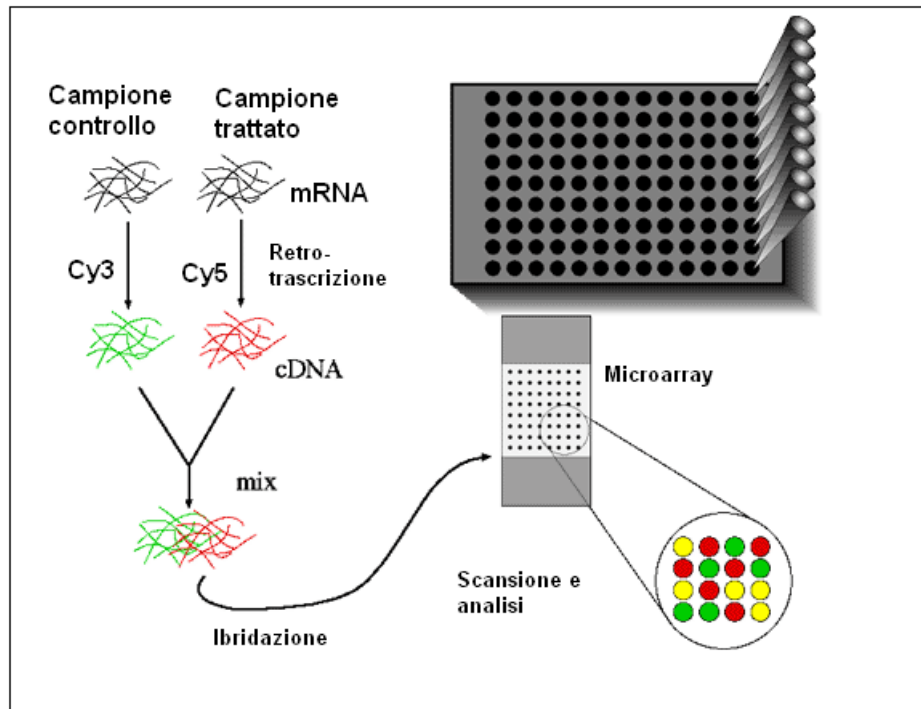


Figura 1.4: Schema riassuntivo di un esperimento di ibridazione su *microarray*. L'RNA estratto dai due campioni (controllo e trattato) viene retrotrascritto in cDNA e marcato indirettamente con due molecole fluorescenti differenti, Cy3 e Cy5. Dopo aver mescolato i due campioni si ibrida la soluzione con le molecole sonda presenti sul *microarray*. A questo punto uno scanner ad alta risoluzione acquisisce le immagini che vengono analizzate da particolari *software* per identificare i geni che alterano la loro espressione a seguito del trattamento.

d'onda di 550 nm e 570 nm, la Cy3, di 649 nm e 670 nm, la Cy5. I principali vantaggi della tecnologia del *microarray* sono:

- l'analisi veloce e precisa, data la procedura di costruzione standardizzata e robotizzata;
- le dimensioni ridotte del supporto (un vetrino da microscopia) e l'alta densità con cui le sequenze sono depositate sulla piattaforma di vetro (maggiore di 5000 elementi per cm^2), che consentono di analizzare in un unico esperimento migliaia di geni, addirittura un intero genoma o via metabolica, utilizzando pochissimo materiale di partenza;
- non necessita di prenoscenze genetiche o molecolari, perciò può essere utilizzata per studiare organismi poco noti di cui si possiede almeno una libreria genica;
- la rivelazione dell'ibridazione attraverso molecole fluorescenti è veloce ed ha una grande risoluzione spaziale;
- rispetto alla marcatura radioattiva è più sicura da maneggiare ed è più facile lo stoccaggio;
- l'automazione e la standardizzazione delle procedure ha portato alla gestione informatica dei protocolli di esecuzione degli esperimenti e dei dati all'interno dei singoli laboratori (LIMS, *laboratory information management systems*) e anche a livello pubblico (Data Base di *microarray*, GEO, ARRAYEXPRESS, e MIAME: *Minimum information about a microarray experiment*). Queste informazioni raccolte e standardizzate danno un prezioso appoggio sia per il disegno sperimentale che per l'analisi dei dati ottenuti.

Il primo esperimento di *microarray* risale al 1995, anno in cui Schena e collaboratori hanno studiato l'espressione di foglia e radice in *Arabidopsis thaliana*, utilizzando una piattaforma composta da 45 ESTs (*expressed sequence tag*)

[Schena et al., 1995]. In dieci anni la bibliografia che riguarda l'uso del *microarray* è notevolmente aumentata. Svariate applicazioni di questa tecnica sono state condotte su numerosi organismi: uomo, topo ed altri mammiferi, piante, tra cui *Arabidopsis*, riso, mais, tabacco, pomodoro, patata, petunia, fragola.

Una tra le più importanti applicazioni del *microarray* è lo studio delle interazioni pianta-patogeno e delle conseguenti risposte di difesa. Molti progetti di questo tipo sono stati condotti su *N. tabacum*, *A. thaliana*, *Z. mays* ed altre specie vegetali [Baldwin et al., 1999; Schenk et al., 2000; Scheideler et al., 2002; Chen et al., 2002]. Questi studi hanno confermato l'espressione di geni già noti, implicati nelle risposte di difesa, ma hanno anche portato alla scoperta di nuovi geni legati alla patogenesi, di vie metaboliche o di trasduzione del segnale prima soltanto ipotizzate, di interazioni sinergiche e/o antagonistiche tra differenti vie difensive. Grazie ad esperimenti di *microarray* sono state delucidate vie di difesa patogeno-specifiche e non-patogeno-specifiche, sono stati approfonditi i meccanismi alla base della reazione ipersensitiva (HR), della resistenza sistemica acquisita (SAR) e della risposta di difesa contro gli insetti erbivori.

L'errore nell'analisi dei *microarray*

Nell'ambito delle scienze sperimentali la parola *errore* non ha la tradizionale connotazione di equivoco o di sbaglio, bensì assume il significato di incertezza intesa come quell'entità di cui sono affette inevitabilmente tutte le misure. Infatti, se si potessero trattare le incertezze sperimentali alla stregua di veri e propri errori, basterebbe agire sulle fonti di questi o sforzarsi di operare nel modo più accurato possibile per eliminarli, mentre è inevitabile che esse siano presenti nel processo di misura. Anche procedendo con estrema cura nei vari passaggi che comprendono un esperimento di ibridazione, fonti di errore di varia natura si presentano ad ogni passaggio [Schuchhardt et al., 2000].

Esistono due classi di errori definiti *casuali* e *sistematici*. Entrambi colpiscono

no in modo diverso gli esperimenti di ibridazione dei *microarray* pregiudicando l'accuratezza dei risultati finali. Metodi manuali e matematici vengono utilizzati per cercare di ridurli al minimo.

Errori casuali colpiscono la preparazione delle molecole bersaglio e delle molecole sonda in modo tale da distribuire il valore di espressione misurato di un gene attorno al suo valore vero seguendo una distribuzione normale [Schuchhardt et al., 2000]. È possibile ridurre questa categoria di errori tramite la ripetizione degli esperimenti un numero sufficiente di volte. Una buona stima del valore vero è dato dalla media delle misure compiute sulla medesima grandezza. L'intervallo fiduciale nel quale ci si aspetta si trovi il valore vero è rappresentato invece dalla varianza dei dati.

Gli errori che colpiscono un esperimento di ibridazione appartengono anche al gruppo degli errori sistematici. Essi fanno variare sistematicamente il valore della grandezza misurata in modo positivo o negativo. L'errore sistematico più importante che si compie nella preparazione di un esperimento di ibridazione con *microarray* è dato dalla differenza intrinseca di luminosità che presentano le due molecole fluorescenti utilizzate per la marcatura del bersaglio. Ne deriva che il campione marcato con la molecola meno luminosa ha mediamente un valore osservato di espressione genica minore. Per ridurre dunque questo genere di errore non serve ripetere l'esperimento più volte, bisogna invece applicare ai dati alcuni processi statistici di analisi definiti globalmente *normalizzazione*.

1.3.3 PCR quantitativa

Le tecniche di verifica basate sulla PCR quantitativa consentono di identificare e misurare un particolare trascritto in modo rapido, sensibile e altamente specifico, inoltre richiedono piccolissime quantità di materiale di partenza (pochi ng di RNA) [Al-Taher et al., 2000; Rotenberg et al., 2006]. Per queste ragioni essa rappresenta attualmente la tecnica d'elezione nella verifica dei dati ottenuti da saggi di *microarray*.

La PCR quantitativa è una tecnica che consente il monitoraggio della quan-

tità di DNA stampo prodotto durante il processo di PCR. Essa riesce a combinare l'amplificazione e la rilevazione in un unico passaggio. Ciò è reso possibile grazie all'utilizzo di diversi composti chimici fluorescenti che si legano al prodotto di PCR in modo direttamente proporzionale alla sua concentrazione (SYBR-GREEN). La quantificazione del trascritto può essere assoluta o relativa. Nella quantificazione assoluta, l'amplificato a quantità ignota viene confrontato con varie diluizioni di un campione a quantità nota. Nella quantificazione relativa, la quantità del trascritto di interesse viene rapportata a quella di un gene controllo. Generalmente il controllo è un gene che mostra un livello di espressione invariato nei diversi tessuti e/o condizioni analizzate (es. un gene *housekeeping*).

Capitolo 2

Scopo della tesi

Il lavoro svolto durante il dottorato ha avuto come scopo il monitoraggio della risposta di *N. benthamiana* nelle prime fasi di infezione con il *Turnip Vein-Clearing Virus*. Si è reso quindi necessario:

1. sviluppare dei metodi molecolari efficienti per aumentare la percentuale di cellule infettate nel tessuto. Un'infezione diretta del tessuto ha una resa molto bassa, il che si traduce in piccole aree infettate immerse in una maggioranza di cellule sane non ancora venute a contatto con il virus. Questa "diluizione" dell'infezione impedisce uno studio dettagliato dei geni attivi durante la risposta precoce;
2. sviluppare una nuova piattaforma di *microarray* a cDNA per poter studiare in modo più generale le interazioni pianta-patogeno;
3. studiare i cambiamenti nel profilo di espressione genica di *N. benthamiana* nelle prime fasi della risposta quando la concentrazione virale all'interno delle cellule deve ancora raggiungere il valore massimo con l'ausilio di questa nuova tecnica di infezione e di analisi;
4. sviluppare un nuovo metodo di analisi dei dati di PCR quantitativa per valutare con precisione la concentrazione del genoma virale all'interno delle cellule.

Capitolo 3

Materiali e metodi

3.1 Crescita delle piante ed inoculazione

Trenta piante di *Nicotiana benthamiana* sono state fatte crescere in condizioni controllate ad una temperatura di 25 °C , al 70% di umidità e con un fotoperiodo di 12 ore. Dopo 3 settimane di crescita, le piante erano pronte per l'inoculo. Sono state separate in due gruppi, mediante un campionamento completamente casuale, di 15 piante ognuno. Un gruppo è stato inoculato con *Agrobacterium* con il plasmide pICH18000 contenente il genoma virale ed il gene reporter GFP (da *Green Fluorescence Protein*) (gruppo di piante infettate). Il secondo gruppo è stato inoculato con *Agrobacterium* con il plasmide pICBV24GFP contenente una parte non funzionale della GFP (gruppo di piante controllo).

La crescita batterica e l'inoculo sono stati eseguiti come segue:

- 1 μ L di cultura batterica contenente il plasmide pICH18000 o il plasmide pICBV24GFP è stato trasferito in 20 mL di terreno di coltura liquido LB (Tiptone 10 g/L, Cloruro di Sodio 10 g/L, estratto di lievito 5 g/L) con l'aggiunta degli antibiotici rifampicina (50 μ g/mL) e canamicina (50 μ g/mL)
- La coltura batterica è stata incubata in agitazione per 15 ore a 28 °C

- i batteri sono stati concentrati tramite centrifugazione a 6000 rpm per 10 minuti
- eliminato il surnatante, il pellet batterico è stato risospeso in 3 mL di una soluzione di MgCl_2 10 mM
- alla soluzione è stato aggiunto 3 μL di acetosiringone 100 mM ed è stata incubata per 1 ora a temperatura ambiente
- i batteri sono stati diluiti con una soluzione di MgCl_2 10 mM fino ad una densità ottica di 0.7 (calcolata mediante l'uso di uno spettrofotometro)
- con un puntale da 20 μL è stato praticato un foro su una foglia di *N. benthamiana*
- la soluzione batterica è stata aspirata con una siringa da 5 ml
- la siringa è stata appoggiata sulla zona della foglia bucata chiudendo l'estremità inferiore del foro con un dito
- Con una leggera pressione dello stantuffo la soluzione batterica è stata inoculata all'interno della foglia

3.2 Estrazione dell'RNA e delle proteine

Le foglie di *N. benthamiana* inoculate con *Agrobacterium* sono state campionate ad intervalli regolari di 12 ore da 4 a 6 giorni successivi all'inoculo sia per il campione di piante infettato che per il controllo. Da ogni campione biologico sono stati estratti l'RNA e le proteine.

Le fasi di estrazione dell'RNA sono riportate di seguito:

- 2 foglie inoculate per ogni pianta sono state tagliate ed inserite in una provetta da 1.5 mL per essere immediatamente congelate in azoto liquido

- il tessuto vegetale è stato frantumato con un pestello fino a formare un polvere sottile
- il campione biologico è stato diviso in due aliquote: $\frac{2}{3}$ del campione sono stati usati nell'estrazione dell'RNA, mentre $\frac{1}{3}$ del tessuto è stato usato per l'estrazione delle proteine
- 1 ml di soluzione di estrazione (TRIzol, Invitrogen, Numero di Catalogo 15596) è stato aggiunto per ogni 100 mg di tessuto fresco ed è stato incubato per 5 minuti a temperatura ambiente
- i campioni sono stati quindi centrifugati a 12000 g per 10 minuti per recuperare il surnatante e eliminare i detriti cellulari
- alla soluzione sono stati aggiunti 0.2 volumi di Cloroformio. La soluzione è stata quindi mescolata vigorosamente ed incubata per 5 minuti a temperatura ambiente
- i campioni sono stati centrifugati a 12000 g per 15 minuti per separare la fase acquosa (contenente l'RNA) dalla fase organica (contenente proteine e acidi grassi)
- alla fase acquosa sono stati aggiunti 0.5 volumi di isopropanolo
- dopo una incubazione di 10 minuti a temperatura ambiente, i campioni sono stati centrifugati a 12000 g per 15 minuti
- il pellet è stato lavato con Etanolo 75 % e centrifugato a 12000 g per 5 minuti
- dopo l'evaporazione completa dell'etanolo, il pellet è stato risospeso in 100 μL di H_2O

Protocollo di estrazione delle proteine:

- al tessuto polverizzato e separato nella fase precedente sono stati aggiunti 100 μL di tampone di estrazione proteica (SDS 2% peso/volume;

β -mercaptoetanolo 2 mM; glicerolo 4%; Tris-HCl 40 mM; blu di bromofenolo 0.01%)

- i campioni sono stati denaturati a 95 °C per 10 minuti per consentire la completa denaturazione delle proteine
- dopo una centrifugazione a 5000 g per 5 minuti è stato recuperato il surnatante
- i campioni così trattati erano pronti per essere caricati in gel di acrilamide SDS-PAGE

3.2.1 Purificazione, quantizzazione e verifica qualitativa dell'RNA

L'RNA è stato purificato mediante l'utilizzo del kit RNeasy Plant Mini Kit (Qiagen, n. cat. 74903). La quantizzazione è avvenuta mediante lettura al NanoDrop (ND 1000, Thermo Fisher Scientific) e la sua qualità è stata testata con il Bioanalyzer 2100 (Agilent Technologies).

3.3 Verifica dell'infezione

La verifica dell'infezione virale è stata condotta sia in vivo che in vitro. In vivo è stata monitorata la fluorescenza emessa dalla proteina GFP mediante esposizione delle foglie inoculate con una sorgente di raggi UV (lunghezza d'onda di 365 nm). Per le prove in vitro è stata usata la corsa in gel di acrilamide (SDS-PAGE) delle proteine estratte dai campioni biologici e la quantizzazione del genoma virale mediante amplificazione con la tecnica della PCR quantitativa (qPCR).

3.3.1 SDS-PAGE

20 μ L della soluzione di estrazione proteica sono stati corsi in SDS-PAGE (da *sodium dodecyl sulfate polyacrylamide gel electrophoresis*) con un gel di

corsa contenente acrilamide al 10% (TrisHCl pH 8.8 0.375 M; SDS 0.1%; acrilamide 10%; bisacrilamide 0.27%; ammonio persolfato (APS) 0.05%; TEMED 0.01%), e un gel di impaccamento contenete acrilamide al 4% (TrisHCl pH 6.8 0.125M; SDS 0.1%; acrilamide 4%; bisacrilamide 0.1 %; ammonio persolfato (APS) 0.05%; TEMED 0.01%). Il gel è stato fatto correre con un voltaggio costante di 250 V per circa 30 minuti (fino a che il blu di bromofenolo contenuto nel tampone di estrazione non raggiungeva il margine inferiore del gel) usando un tampone di corsa contenente TrisHCl 25 mM, glicina 200 mM e SDS 0.1%.

Sucessivamente il gel è stato immerso in una soluzione di colorazione (SimplyBlue™ SafeStain, Invitrogen, n. cat. LC6060) e mantenuto in agitazione per circa 30 minuti (fino a quando le bande proteiche diventavano ben visibili). A colorazione ultimata, il gel è stato lavato in acqua distillata ed è stato fotografato per la catalogazione.

3.3.2 PCR quantitativa

Circa 500 ng di RNA totale ottenuto dai medesimi campioni usati per le ibridazioni, è stato trattato per 15 minuti a temperatura ambiente con 1 U di enzima DNasi I (Amplification Grade, Invitrogen) in presenza del tampone 10X DNasi I (200 mM Tris-HCl pH 8.4, 20 mM MgCl₂, 500 mM KCl). Ciò ha permesso la digestione dell'eventuale DNA genomico residuo presente nel campione. Dopo inattivazione della DNasi con 0.8 μ L di EDTA 25 mM, l'RNA è stato retrotrascritto in cDNA dall'enzima MultiScribe™ Reverse Transcriptase in presenza dell'innesco oligo(dT)₁₆, secondo le istruzioni del kit "GeneAmp® Gold RNA PCR Core Kit" (Applied Biosystems). Un'identica reazione senza l'enzima (RT-) è stata condotta per verificare la presenza di DNA genomico nel campione di RNA. Una volta ottenuto il cDNA, esso è stato amplificato tramite PCR quantitativa in presenza di inneschi specifici per il genoma virale e per il *reference gene* EF1- α . I *primers* sono stati disegnati usando il software on-line Primer3 [<http://frodo.wi.mit.edu>] e, quando possibile, sono stati disegnati a cavallo di una regione intronica,

in modo tale da limitare la possibilità di una amplificazione del DNA genomico. Tutte le coppie di *primers* sono state disegnate in modo da avere temperature di fusione (T_m da *melting temperature*) simili e da poter essere utilizzate nelle medesime condizioni di amplificazione. Le sequenze degli inneschi sono state vagliate per la formazione di forcine o siti di legame secondari con lo stampo e per la formazione di dimeri tramite programmi standard di analisi del DNA. È stata condotta un'analisi di similarità BLAST (<http://www.ncbi.nlm.nih.gov/BLAST>) per escludere la possibilità di legami aspecifici dei *primer*. Le PCR quantitative sono state condotte utilizzando i reagenti e seguendo il protocollo del kit “iQTM SYBR Green Supermix” (BioRad). Per ogni campione sono state allestite tre repliche di una stessa PCR, una reazione RT- e un controllo negativo avente per stampo l'acqua. La miscela di amplificazione è stata allestita come segue:

Reagente	Concentrazione finale
2X iQ SYBR Green Supermix	1X
<i>Primer forward</i>	100 nM
<i>Primer reverse</i>	100 nM
DNA stampo	1 μ L
H ₂ O	a volume (50 μ L)

Le PCR sono state condotte nel termociclatore “DNA Engine Thermal Cycler” (MJ Research) provvisto del sistema di rivelazione “Chromo4” (BioRad), secondo il seguente programma:

Fase	Temp [°C]	Durata	N° di cicli
Denaturazione iniziale	95	5'	1
Denaturazione	95	15"	
Appaiamento	60	30"	50
Estensione	72	30"	
Estensione finale	72	10'	1
Analisi della curva di fusione	da 95 a 65	1"	30

3.4 Sintesi del cDNA e marcatura del bersaglio

La sintesi del cDNA è avvenuta in duplicato per ogni campione biologico in modo tale da consentirne la marcatura con entrambi i fluorocromi (*dye swap*). In questo modo si sono minimizzati gli errori sistematici dovuti alla differente intensità luminosa delle due molecole fluorescenti. La reazione di retro-trascrizione è stata condotta grazie alla presenza dell'innesco oligo(dT)₁₆, seguendo le istruzioni del kit "SuperScript™ Indirect cDNA Labelling System" (Invitrogen). Al cDNA di neosintesi sono stati aggiunti 15 μL di NaOH 1 N in modo tale da interrompere la reazione di sintesi e di degradare l'RNA rimasto in soluzione. La reazione è stata incubata 10 minuti at 70 °C e poi neutralizzata tramite l'aggiunta di 15 μL di HCl 1 N e 20 μL di NaAc 3 M pH 5.2. L'amminoallil-cDNA a singola elica è stato purificato dai nucleotidi non incorporati e dall'enzima tramite l'utilizzo del kit RNeasy Plant Mini Kit (Qiagen, n. cat. 74903). La marcatura indiretta dell'aminoallil-cDNA è avvenuta a temperatura ambiente per 1 ora al buio, tramite l'accoppiamento del cDNA con le molecole fluorescenti (CyScrige™ Post-Labeling Kit, Amersham Biosciences, n. cat. RPN5660). Ogni replica biologica è stata marcata alternativamente con la sonda Cy3 e con la sonda Cy5. Trascorsa l'ora di incorporazione, i fluorocromi non accoppiati sono stati rimossi tramite purificazione con il kit RNeasy Plant Mini Kit (Qiagen). Per confermare l'avvenuta marcatura, i bersagli sono stati quantificati al NanoDrop. La concentrazione dei fluorocromi è stata calcolata nel seguente modo (A = assorbanza):

$$Cy3[pmol] = \frac{(A_{550} - A_{650}) \cdot 55}{0.15} \quad (3.1)$$

$$Cy5[pmol] = \frac{(A_{650} - A_{750}) \cdot 55}{0.25} \quad (3.2)$$

Dunque, le coppie di campioni da utilizzare negli esperimenti di ibridazione competitiva sono state riunite e precipitate come segue:

- alle molecole bersaglio sono stati aggiunti 0.1 volumi di sodio acetato 3M pH 5.2, 2 μL di glicogeno (20 mg/ml) e 2 volumi di etanolo asso-

luto freddo, quindi la precipitazione del cDNA è stata condotta per 30 minuti a -20 °C

- la soluzione è stata centrifugata a 13000 g per 20 minuti a 4 °C
- dopo l'eliminazione del surnatante il precipitato è stato lavato con una soluzione di etanolo 75%
- la soluzione è stata centrifugata a 13000 g per 5 minuti a 4 °C
- dopo la completa rimozione del surnatante, il pellet è stato fatto asciugare all'aria fino alla completa evaporazione dell'etanolo
- il DNA è stato risospeso in 11 μ l di tampone di ibridazione 1.17X (formamide 50%, SSC 5X e SDS 0.1%)

3.5 Costruzione del *microarray*

Le sequenze sonda presenti sulla superficie del microarray derivano dall'amplificazione di tre librerie geniche. La prima libreria era stata ottenuta dall'Istituto di Genomica dell'Università dell'Arizona (*University of Arizona Genomics Institute, Tuscon, Arizona, US*). L'RNA utilizzato per la costruzione di questa libreria deriva da diversi tessuti di *N. benthamiana* tra i quali tessuto calloso, radici di piante cresciute in coltura liquida, foglie sottoposte a stress termici (3 ore e 6 ore a 38 °C; 3 ore e 6 ore a 5 °C), foglie infettate (12 ore con *Pseudomonas syringae* pv tomato; 12 e 18 ore con *Xanthomonas campestris* pv campestris; 18 ore con *Pseudomonas syringae* pv phaseolicola; 18 ore con *Xanthomonas campestris* pv vesicatoria). La medesima quantità di RNA isolato da questi tessuti è stata riunita in un'unica soluzione per la creazione di un'unica libreria genica di 3512 cloni unici. La seconda libreria genica utilizzata per la creazione delle molecole sonda proviene dal laboratorio del Dott. Steven A. Lommel (*Tobacco Genome Initiative, North Carolina State University, US*). In questo caso l'RNA messaggero è stato isolato da tessuti non trattati di *N. benthamiana* a formare una unica libreria di 769

cloni unici. La terza libreria contenente 5323 cloni unici proveniva dal laboratorio del Prof. Andy Maule (*Disease and Stress Biology, John Innes Centre, Norwich, UK*) ed era costituita con tessuto aereo di *N. benthamiana*. In totale quindi il *microarray* era composto da più di 9600 geni differenti.

L'amplificazione ed il sequenziamento delle librerie geniche è stata affidata al Genome Laboratory (*John Innes Centre, Norwich, UK*) mentre la stampa del vetrino è stata affidata all'Università di Nottingham (*Post-Genomic Technologies Facility, Nottingham, UK*).

3.5.1 Annotazione delle sequenze

Tutte le sequenze depositate sul vetrino sono state sequenziate e successivamente annotate mediante l'uso del software Blast2GO [Conesa et al., 2005]. In questo modo è stato possibile dare una funzione teorica ai geni depositati mediante ricerca di similarità con proteine note. Inoltre questo programma consente l'annotazione delle sequenze con il database Gene Ontology (GO) il quale raggruppa proteine con funzioni simili facilitando l'analisi e l'interpretazione dei risultati di esperimenti di *microarray*. Per facilitare il confronto con la pianta modello *A. thaliana* le sequenze depositate sono state anche annotate con il database proteico di *A. thaliana*. In questo modo è stato possibile utilizzare l'enorme quantità di dati disponibili di *A. thaliana*.

3.6 Esperimenti di ibridazione su *microarray* di cDNA

3.6.1 Disegno sperimentale

Per limitare le differenze di espressione genica dovute a variazioni nell'inoculo, l'RNA è stato isolato da due foglie della medesima pianta. Il tessuto derivante da tre repliche biologiche è stato raccolto per ogni *time point*. Un totale di 15 piante di *N. benthamiana* sono state inoculate con *Agrobacterium* contenente il plasmide pICH18000 (con inserito il genoma del virus TVCV)

e campionate a gruppi di tre a tempi successivi (4, 4.5, 5, 5.5, 6 giorni dopo l'inoculo). Per mascherare la risposta della pianta verso *Agrobacterium*, altrettante piante sono state inoculate con il battere contenente il plasmide pICBV24GFP, e campionate ai medesimi tempi. L'ibridazione è avvenuta confrontando le piante infettate con il virus ed il rispettivo controllo. Per ridurre l'errore prodotto dalla diversa fluorescenza delle due molecole usate per marcare il cDNA, ogni ibridazione è stata condotta in doppio invertendo le molecole di fluorocromi che venivano legate alle molecole sonda. In totale sono stati condotti 30 esperimenti di ibridazione.

3.6.2 Preibridazione

Prima che la fase di ibridazione avesse luogo, bisognava reidratare e pulire i vetrini per eliminare l'eventuale presenza di polvere sulla parte stampata. I vetrini inoltre contenevano numerosi siti attivi liberi che potevano legare in modo aspecifico il DNA. Questi siti sono stati saturati aggiungendo al tampone di preibridazione una grossa quantità di albumina del siero bovino, BSA. Il vetro è stato completamente immerso nel tampone di pre-ibridazione (SSC 0.5X, SDS 0.1% (v/v), BSA 1% (p/v)) ed incubato a 40 °C per almeno 45 minuti. Quindi il vetrino è stato lavato per 2 minuti in una soluzione SSC 0.5X e poi asciugato per centrifugazione (2 minuti a 4000 g).

3.6.3 Ibridazione

La procedura di ibridazione è stata condotta nel seguente modo:

- per evitare che durante l'ibridazione avvenissero reazioni aspecifiche tra le code di Adenina delle EST immobilizzate sul vetrino e le code di Timina presenti nel cDNA bersaglio, ai campioni di DNA marcato sono stati aggiunti 4 μL di oligo(dT)₁₆
- la miscela è stata riscaldata a 95 °C per 3 minuti e poi immediatamente riposta in ghiaccio. Questa procedura ha permesso la denaturazione

di eventuali strutture secondarie del cDNA che altrimenti avrebbero potuto interferire con il processo di ibridazione

- dopo preibridazione, il vetrino è stato alloggiato in apposite camerette di ibridazione di metallo, provviste di margini siliconici utili ad impedire l'entrata dell'acqua
- la miscela dei cDNA marcati è stata depositata sul vetrino e distribuita uniformemente su tutta la superficie del *microarray* grazie al lento adagiamento di un vetrino coprioggetto
- per evitare la disidratazione della superficie di ibridazione del vetrino, nei due pozzetti interni alla cameretta sono stati depositati circa 70 μL di una soluzione SSC 0.1X
- la camera è stata chiusa ermeticamente, poggiata su una superficie perfettamente orizzontale e immersa in un bagnetto termostato
- l'incubazione è stata protratta per 16 ore a 40 °C

3.6.4 Lavaggi

Terminata l'ibridazione, il vetrino è stato sottoposto ad una serie di lavaggi necessari all'eliminazione delle molecole bersaglio non legatesi alla sonda. I lavaggi sono stati condotti al buio nel seguente modo:

- il vetrino è stato rimosso dalla cameretta e posto in un tampone preriscaldato a bassa stringenza (SSC 1X , SDS 0.2%). Il coprioggetto è stato rimosso per delicata agitazione del vetrino con l'aiuto di una pinzetta
- il vetrino è stato quindi lavato per 15 minuti nella soluzione di lavaggio 1 (SSC 0.5X, SDS 0.01%)
- il vetrino è stato poi lavato per altri 15 minuti nella soluzione di lavaggio 2 (SSC0.06X , SDS 0.01%)

- l'ultimo lavaggio è avvenuto in soluzione di lavaggio 3 (0.06X SSC) per 15 minuti
- infine, il vetrino è stato asciugato per centrifugazione (2 minuti a 4000 g).

3.6.5 Scansione dei *microarray*

La lettura e digitalizzazione delle immagini è stata eseguita presso il *Genome Laboratory (John Innes Centre)* mediante l'uso dello scanner GenePix 4200A Professional. Lo scanner ha eseguito la lettura per ogni sonda stampata sul *microarray* delle lunghezze d'onda di emissione dei due fluorocromi Cy3 e Cy5, traducendole in immagini monocromatiche. La trasformazione dei punti luminosi in valori numerici è stata affidata al programma GenePix Pro, che come *output file* ha consegnato una tabella contenente varie informazioni sui punti (*spot*) della piattaforma: numero dello spot, posizione che occupa nella griglia del *microarray*, intensità media emessa dai due fluorocromi, vari altri indici statistici.

3.7 Analisi dei dati

Nelle due sezioni che seguono verranno descritti in dettaglio i metodi utilizzati per analizzare i dati degli esperimenti di ibridazione con il *microarray* e di PCR quantitativa.

3.7.1 Analisi dei dati di *microarray*

L'analisi dei dati degli esperimenti di ibridazione è stata eseguita con l'utilizzo del software gratuito di analisi statistica R. Se da una lato questo software è molto potente e versatile, tanto da poter essere utilizzato per l'analisi di qualsiasi tipo di *microarray* e di qualsiasi disegno sperimentale, da un altro lato è difficile da usare dal momento che fa uso di un linguaggio di programmazione proprio e non di una interfaccia grafica. Per questa ragione

descriverei nel dettaglio tutti i passaggi che portano ai risultati.

Per prima cosa ho dovuto caricare una particolare libreria chiamata `limma` [Smyth and Speed, 2003] che conteneva tutte le funzioni da utilizzare per la normalizzazione e l'individuazione dei geni differenzialmente espressi:

```
library(limma)
```

Successivamente ho letto il file `Targets.txt` che contiene le informazioni riguardanti ogni vetrino, come ad esempio i nomi dei file con i valori delle intensità luminose, i campioni ibridati, la replica biologica:

```
targets <- readTargets()
```

A questo punto, sfruttando le informazioni contenute nel file `Targets.txt`, il programma ha letto le intensità luminose rilevate dallo scanner e le informazioni relative all'annotazione della piattaforma (contenute nel file `Nbenth.gal`):

```
RG <- read.maimages(targets$FileName, source="genepix")
spottypes <- readSpotTypes()
RG$genes <- readGAL("Nbenth.gal")
RG$genes$Status <- controlStatus(spottypes, RG)
RG$printer <- getLayout(RG$genes)
```

La normalizzazione dei dati è avvenuta usando una procedura standard. Ad una iniziale sottrazione da ogni punto del proprio rumore di fondo, ho applicato una normalizzazione *print tip lowess* che riduce la variabilità dell'intensità luminosa dei due campioni ibridati su un medesimo vetrino [Yang et al., 2002]. Questo tipo di normalizzazione viene definita interna al *microarray* (*within array normalization*) dal momento che applica un fattore di normalizzazione differente per ogni *microarray*. Successivamente ho applicato ai dati una normalizzazione denominata *scale* che riduce la variabilità sia tra regioni interne di ogni *microarray* che tra i diversi *microarray*. Per questa ragione, questo secondo tipo di normalizzazione viene definito tra *microarray* (*between array normalization*) [Yang et al., 2002]:

```
RG.b <- backgroundCorrect(RG)
MA.p <- normalizeWithinArrays(RG.b, method="printTipLoess")
MA.pAq <- normalizeBetweenArrays(MA.p, method="scale")
```

A questo punto i dati di tutti i microarray erano stati normalizzati. I valori di intensità luminosa dei due campioni biologici ibridati su ogni *microarray* erano stati riuniti in un unico valore seguendo la formula:

$$M = \log_2 R - \log_2 G = \log_2 \left(\frac{R}{G} \right)$$

Dove R e G sono le intensità luminose rispettivamente della pianta infettata e di quella controllo per ogni punto contenuto nel *microarray*. Intensità luminose uguali in un punto, $M = 0$, indicano che non c'è variazione di espressione per quel gene tra i due campioni. Al contrario, se il valore di M è positivo, quel gene è maggiormente espresso in pianta infettata, altrimenti è negativo. A questo punto è stato costruito un modello lineare che tenga conto della presenza di 3 repliche biologiche per ogni *time point* e del confronto che deve essere fatto tra il campione infettato ed il rispettivo controllo:

```
design <- modelMatrix(targets, ref="GFP")
isGene <- RG$genes$Status == "cDNA"
fit <- lmFit(MA.pAq[isGene,], design)
contrasts <- makeContrasts(
  "4dpi"="(X1800041 + X1800042 + X1800043)/3",
  "4.5dpi"="(X18000451 + X18000452 + X18000453)/3",
  "5dpi"="(X1800051 + X1800052 + X1800053)/3",
  "5.5dpi"="(X18000551 + X18000552 + X18000553)/3",
  "6dpi"="(X1800061 + X1800062 + X1800063)/3",
  levels=design
)
fit2 <- contrasts.fit(fit, contrasts)
fit2 <- eBayes(fit2)
```

Per poter identificare i geni differenzialmente espressi si è considerato che il *microarray* sia formato da un numero elevato di punti e questi rappresenta-

vano una popolazione casuale dei geni all'interno del genoma della pianta. Inoltre, gli esperimenti venivano condotti su popolazioni cellulari simili. Ci si aspettava dunque che la maggior parte di geni analizzati non mostrassero variazione di espressione. Di conseguenza si è potuto ipotizzare che mediamente il valore di $M \simeq 0$ e che segua un andamento normale. In altre parole, considerando due popolazioni cellulari simili, ci si aspettava che i geni che non mostravano variazione di espressione dovessero essere di gran lunga in numero maggiore in confronto a quelli che mostravano variazione di espressione [Quackenbush, 2002]. Statisticamente, l'ipotesi che si formula per un gene che mantenga costante il proprio valore di espressione in due condizioni sperimentali è:

$$H_0 : \log_2 \left(\frac{R}{G} \right) = 0$$

dove i geni differenzialmente espressi lo sono sotto l'ipotesi alternativa:

$$H_1 : \log_2 \left(\frac{R}{G} \right) \neq 0$$

Bisognava dunque applicare il *test t* di Student ad ogni gene verificando la validità dell'ipotesi nulla.

Il problema sorto da questo approccio era dovuto all'altissimo numero di osservazioni (i 6000 geni presenti sul vetrino) che portavano ad avere un altissimo numero di gradi di libertà. In queste condizioni, la probabilità di incappare in un errore di tipo I (un gene ha casualmente valori di M lontani dallo zero) era molto alta. Infatti, applicando il *test t* con un valore di probabilità $p.value = 0.05$ ci si aspetta di compiere il 5% di errore nell'identificazione dei geni differenzialmente espressi. Significa che 300 geni (il 5% di 6000) potevano venir identificati per errore. Numerosi metodi statistici sono stati sviluppati per limitare questa grossa fonte di errore e quello attualmente più usato viene definito *FDR* (da *False Discovery Rate*) [Reiner et al., 2003; Klipper-Aurbach et al., 1995].

Per l'identificazione dei geni differenzialmente espressi applicando l'algoritmo *FDR*, è stata usata il seguente codice:

```
results <- decideTests(fit2, method = "separate",
```

```

adjust.method="fdr",p.value=0.01)
summary(results)
top1 <- topTable(fit2, coef=1, adjust.method="fdr", number=22)
top2 <- topTable(fit2, coef=2, adjust.method="fdr", number=78)
top3 <- topTable(fit2, coef=3, adjust.method="fdr", number=58)
top4 <- topTable(fit2, coef=4, adjust.method="fdr", number=118)
top5 <- topTable(fit2, coef=5, adjust.method="fdr", number=309)

```

Per stabilire se i processi di normalizzazione avevano portato a dei risultati accettabili, sono stati disegnati numerosi grafici. Di seguito viene riportato le istruzioni per molti di essi:

```

pdf(file='scatter_norm.pdf',width = 5, height = 5)
plotMA(RG)
plotDensities(RG.b)
plotDensities(MA.p)
plotDensities(MA.pAq)
plotPrintTipLoess(RG.b)
plotPrintTipLoess(MA.pAq)
boxplot(MA.pAq$M~col(MA.pAq$M), names=colnames(MA.pAq$M))
dev.off()
pdf(file='QQplot.pdf',width = 6, height = 5)
qqt(fit2$t,df=fit2$df.prior+fit2$df.residual,pch=16,cex=0.2)
abline(0,1)
dev.off()
#general plotMA with the name of differentially expressed genes
pdf(file='model_4.5dpi.pdf',width = 6, height = 5)
plotMA(fit2[,2])
ord <- order(fit2$p.value[,2],decreasing=F)
top30 <- ord[1:22]
points(fit2$Amean[top30],fit2$coef[top30,2],cex=0.6,col="red")
dev.off()
#venn diagram

```

```
pdf(file='venn_1_down.pdf',width = 6, height = 5)
vennDiagram(results[,c(1,2,3)],include="down", main="Downregulated genes")
dev.off()
pdf(file='venn_1_up.pdf',width = 6, height = 5)
vennDiagram(results[,c(1,2,3)],include="up", main="Downregulated genes")
dev.off()
pdf(file='venn_2_down.pdf',width = 6, height = 5)
vennDiagram(results[,c(4,5,6)],include="down", main="Downregulated genes")
dev.off()
pdf(file='venn_2_up.pdf',width = 6, height = 5)
vennDiagram(results[,c(4,5,6)],include="up", main="Downregulated genes")
dev.off()
#heat diagram
pdf(file='heat-5wpi.pdf',width = 4, height = 18)
heatDiagram(results, fit2$coef, primary="5dpi",
names=fit2$genes$ID, orientation="por", mar=c(1,1,4,6), cex=1.5)
dev.off()
#volcanoplot
pdf(file='volcanoplot.pdf',width = 6, height = 5)
volcanoplot(fit2, coef=1, highlight=10,names=fit$genes$Name)
dev.off()
```

Finite le analisi i dati sono stati salvati su file per essere accessibili anche con altri software come Excel:

```
write.csv(top1,"4dpi.csv")
write.csv(top2,"4.5dpi.csv")
write.csv(top3,"5dpi.csv")
write.csv(top4,"5.5dpi.csv")
write.csv(top5,"6dpi.csv")
write.csv(summary(results), "summary_results_p005.csv")
save.image(file='Nbenth_2007-01-18.RData')
```

3.7.2 Sviluppo del metodo *LOGIT* di analisi dei dati di qPCR

In generale, l'amplificazione degli acidi nucleici può essere descritta da una crescita esponenziale pura:

$$P = I \cdot E^n \quad (3.3)$$

dove P è la quantità del prodotto di reazione di PCR, I è la quantità iniziale di molecole bersaglio di DNA, E è l'efficienza di reazione (può variare da 1 a 2) e n è il numero di cicli di PCR. Solamente durante la fase iniziale della reazione, l'amplificazione ha un carattere esponenziale paragonabile a quello descritto nell'equazione 3.3. A stadi più tardivi la curva della reazione di amplificazione diverge significativamente dal modello esponenziale, diventando un processo maggiormente casuale. Per questa ragione la qPCR è caratterizzata dal ciclo di PCR nel quale l'amplificato bersaglio è rilevato per la prima volta. Questo valore è normalmente detto *cicle thresold* (C_t), ed è il momento nel quale l'intensità della fluorescenza emessa dall'amplificato supera la fluorescenza di *background*. Appena la reazione raggiunge il valore di fluorescenza prestabilito, si è in grado di ricavare l'informazione necessaria per valutare la quantità iniziale di molecole bersaglio. Il numero di cicli della reazione di amplificazione che rappresentano il C_t del gene bersaglio sono confrontati con il C_t di un gene di controllo. L'espressione di questo gene standard, chiamato anche *housekeeping gene* o *reference gene*, è ritenuta non influenzata dal trattamento sperimentale. A questo punto è possibile confrontare l'espressione del gene bersaglio prima e dopo il trattamento sperimentale mediante l'utilizzo della seguente equazione [Pfaffl, 2001]:

$$R = \frac{(E_{bersaglio})^{\Delta C_t_{bersaglio}(\text{controllo}-\text{trattato})}}{(E_{ref})^{\Delta C_t_{ref}(\text{controllo}-\text{trattato})}} \quad (3.4)$$

Recentemente in diversi lavori è riportato che spesso sequenze differenti amplificano con efficienze differenti [Liu and Saint, 2002; Meijerink et al., 2001]. Quindi, si rende necessario il calcolo separato dell'efficienza di amplificazione (E) tra il gene bersaglio e il gene controllo. Il metodo attualmente in uso per

determinare l'efficienza di una reazione di PCR è il metodo delle diluizioni seriali, ognuna analizzata in triplicato. Usando questo metodo, sono preparate delle diluizioni seriali del materiale di partenza e successivamente, il valore di C_t viene correlato con il valore del logaritmo della concentrazione nota del materiale di partenza. L'efficienza della reazione viene calcolata a partire dalla pendenza della retta di regressione che passa per questi punti. Tuttavia questa è una semplificazione del problema dal momento che l'efficienza può variare considerevolmente in relazione alla quantità iniziale di molecole stampo.

Quindi, ciò che viene richiesto è un metodo che determini l'efficienza sulla base la cinetica di ogni singola reazione di PCR [Liu and Saint, 2002]. Nel metodo proposto da Liu e Saint, viene presa in esame solamente la porzione di dati che seguono un andamento esponenziale. Infatti, dopo trasformazione logaritmica, la fluorescenza misurata segue un andamento lineare con il numero di cicli di PCR. Gli autori considerano la pendenza della retta di regressione che passa per questi punti come l'efficienza della reazione. L'idea di base di questo metodo è corretta, ma lo svantaggio principale consiste nella soggettività dello scegliere quali valori seguano un andamento esponenziale e quali no. Oltretutto, la necessità di una procedura di selezione soggettiva, rende impossibile l'automazione del processo. La parte più complessa in questo modello è la scelta del punto finale della fase esponenziale. In un altro lavoro, il massimo della derivata seconda della sigmoide che fitta i valori di fluorescenza viene considerato come il punto finale della fase esponenziale. Anche in questo caso se l'idea di base è corretta, sorgono alcuni dubbi sull'applicazione della stessa ai dati. Infatti i parametri della sigmoide non vengono scelti mediante un processo di regressione dei dati ma in modo arbitrario.

Per poter confrontare due reazioni di PCR quantitativa bisogna conoscere il C_t , o punto nel quale inizia la fase esponenziale, e l'efficienza della reazione. Per calcolare l'efficienza bisognava inoltre conoscere il ciclo al quale la fase esponenziale termina e lascia il posto ad un processo maggiormente casuale fino al raggiungimento di un *platou*. Un valore attualmente accettato per

questo punto è il massimo della derivata seconda (SDM da *Second Derivative Maximum*) della curva che interpola i dati di fluorescenza. Con il metodo di analisi suggerito di seguito questo processo viene semplificato, rendendo necessario solamente il calcolo dell'SDM e non del C_t .

La formula che lega la fluorescenza al numero di cicli in una reazione di qPCR è ben rappresentato da:

$$F_N = \frac{1}{1 + E^{-(\beta_0 + \beta_1 x)}}$$

con $F_N = \frac{F_i}{F_{MAX}}$, E l'efficienza della reazione, x il numero di cicli e β_0 e β_1 i parametri ignoti della curva.

Dal momento però che non si conosceva l'efficienza della reazione, questa poteva essere sostituita nella formula dall'esponenziale naturale:

$$F_N = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}} \quad (3.5)$$

Come si nota questa formula è identica a quella che viene normalmente utilizzata nei modelli lineari generalizzati per dati binari (modello LOG-IT) [Agresti, 1989]. Si è potuto così applicare ai dati di PCR quantitativa questo modello di regressione. In questo modo venivano facilmente calcolati i parametri della curva β_0 e β_1 .

Con i parametri ottenuti era dunque possibile calcolare l'SDM come il punto in cui la derivata terza dell'equazione 3.5 era uguale a zero:

$$f'''(X) = -\frac{e^{-X}[1 - 4e^{-X} + e^{-2X}]}{(1 + e^{-X})^4}$$

con $X = \beta_0 + \beta_1 x$

$$f'''(X) = 0$$

$$X_{SDM} = -\ln(4)$$

$$\beta_0 + \beta_1 x_{SDM} = -\ln(4)$$

$$x_{SDM} = \frac{-\beta_0 - \ln(4)}{\beta_1}$$

Il passo successivo che si deve compiere è il ricalcolo dei parametri della curva escludendo dall'analisi i dati di fluorescenza per i cicli maggiori dell'SDM.

Infatti in precedenza sono stati calcolati dei parametri con valori medi in modo tale che riescano ad interpolare tutti i dati dell'esperimento. Ma, come detto in precedenza, la parte importante della curva è quella che segue un andamento esponenziale puro, e cioè i valori di fluorescenza per i cicli minori dell'SDM. Solamente in questa regione della curva aveva senso calcolare l'efficienza.

Dopo aver ricalcolato β_0 e β_1 si può calcolare l'efficienza della reazione come:

$$E = e^{\beta_1} \quad (3.6)$$

La potenza di questa tecnica però si osserva maggiormente quando si è dovuto confrontare due curve differenti. Infatti, dal momento che il risultato di una singola reazione di PCR quantitativa riflette solamente la quantità della molecola bersaglio nella forma di unità di fluorescenza, essa deve essere considerata in funzione di un controllo. In altre parole la quantità di trascritto genico in un campione trattato doveva essere confrontata con la quantità del medesimo trascritto nel relativo campione controllo. Per fare ciò basta trovare la distanza tra le curve (qui indicate con il pedice 1 e 2) calcolate per il gene di interesse prima e dopo trattamento. La formula per ottenere la distanza è:

$$\Delta = \frac{(\beta_{1.1} - \beta_{2.1})x_1 + \beta_{1.0} - \beta_{2.0}}{\beta_{2.1}} \quad (3.7)$$

Infine, unendo l'equazione 3.7 appena calcolata con l'equazione 3.6 che calcola l'efficienza dalla reazione e il logaritmo della formula proposta da Pfaffl [Pfaffl, 2001] (vedi equazione 3.4) si ottiene l'equazione seguente:

$$\log_2 R = [(\beta_{0ctrl} - \beta_{0infetta})_{bersaglio} - (\beta_{0ctrl} - \beta_{0infetta})_{ref}] \log_2 e \quad (3.8)$$

nel quale β_{0ctrl} è il coefficiente β_0 calcolato sul campione controllo e $\beta_{0infetta}$ è il coefficiente β_0 calcolato sul campione infettato con il virus, entrambi calcolati sia per il gene bersaglio (*bersaglio*) e per il gene standard (*ref*). In questo modo è possibile confrontare i risultati delle analisi di PCR quantitativa con quelli ottenuti con il *microarray*.

Capitolo 4

Risultati e Discussione

4.1 Crescita delle piante ed inoculazione

Quando un virus entra in una cellula vegetale il suo genoma si replica e comincia a diffondersi alle cellule vicine propagando l'infezione a tutto il tessuto. La velocità di propagazione all'interno del tessuto è notevolmente inferiore alla velocità di replicazione virale. Ciò ha come risultato la formazione di piccole aree infettate circondate da una grande quantità di cellule sane, non ancora in contatto con il virus. È quindi inappropriato un campionare l'intero tessuto, perchè la variazione dell'espressione genica dovuta all'infezione viene mascherata dalla presenza di cellule sane.

Per aumentare la percentuale di cellule infettate nel campione si è usata una tecnica di infezione indiretta. Il genoma virale è stato inserito nel genoma della pianta mediante l'*Agrobacterium*. In questo lavoro, è stato utilizzato un ceppo di *Agrobacterium* contenente il plasmide pICH18000 sviluppato nel laboratorio del Prof Gleba (Icon Genetics, Germania), che trasporta nella regione del T-DNA l'informazione per generare tutto il genoma del *Turnip vein-clearing virus* e il gene reporter GFP (*Green Fluorescence Protein*) (Figura 4.1). Per consentire una replicazione efficiente, a monte del genoma virale è stata inserita la sequenza del promotore del gene *ACTINE2* di *A. thaliana*. Questo gene viene trascritto costitutivamente in tutti i tessuti della pinata

e la sua espressione non viene alterata da fattori esterni [Yang et al., 2003]. Per aumentare il tasso di trascrizione, all'interno del genoma virale sono state inserite delle sequenze introniche che mimano la struttura tipica di un gene eucariotico. In questo modo il trascritto sintetico viene riconosciuto dalle proteine di *splicing* e processato più facilmente, consentendo una percentuale di cellule infettate del 90% sul totale delle cellule del tessuto [Marillonnet et al., 2005].

Per mascherare il metodo di infezione, altrettante foglie di piante diverse sono state inoculate con *Agrobacterium* con il plasmide pICBV24GFP che conteneva una parte non funzionale della GFP (gruppo di piante controllo). Per verificare che nel tessuto inoculato il virus replicasse in modo uniforme e costante nel tempo sono state eseguite delle verifiche sia *in vivo* che *in vitro*.

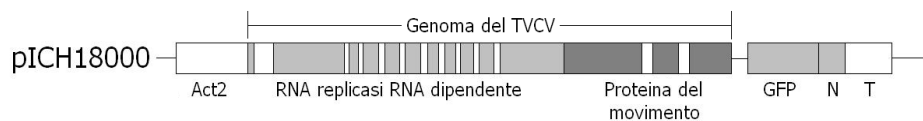


Figura 4.1: Rappresentazione schematica della regione del plasmide pICH18000 contenente il genoma virale. Act2, promotore del gene *ACTINE2* di *A. thaliana*. Le regioni colorate in bianco all'interno del genoma virale rappresentano le sequenze introniche inserite artificialmente. N, regione 3' terminale del genoma cr-TMV; T, terminatore Nos.

4.2 Verifica dell'infezione virale

4.2.1 Verifica *in vivo*

La verifica *in vivo* dell'infezione virale è stata eseguita monitorando la fluorescenza emessa dalla proteina GFP, presente nel plasmide pICH18000, quando questa viene eccitata da una sorgente luminosa di raggi UV. Dalla Figura 4.2 emerge che la fluorescenza è uniformemente distribuita su tutta la superficie fogliare indicando che circa il 90% delle cellule contengono la GFP e che

quindi sono state infettate dal virus. La GFP è un marcatore dell'infezione virale poichè viene espressa dal medesimo promotore usato per l'espressione del virus e la sua espressione è diretta conseguenza dell'espressione virale.

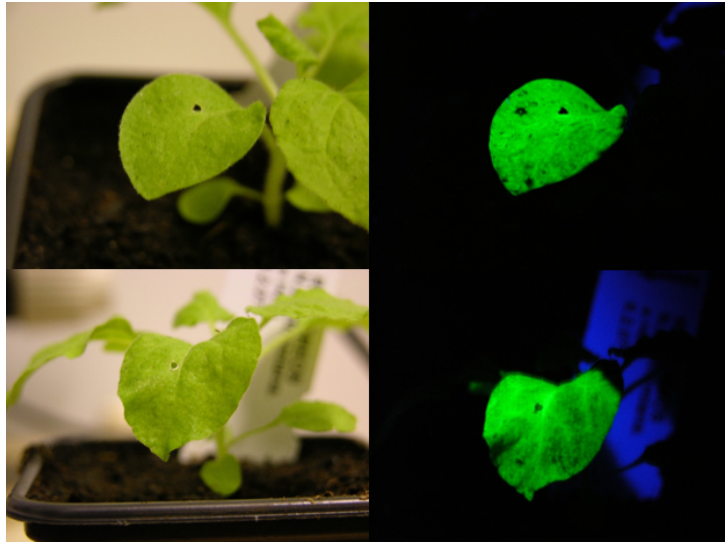


Figura 4.2: Saggio in vivo della presenza della GFP nel tessuto inoculato dopo 6 giorni prima e dopo irraggiamento della superficie fogliare con raggi UV. L'uniformità nella fluorescenza indica che la percentuale di cellule infettate è molto elevata (90%). Al centro delle foglie si nota il forellino praticato per l'inoculo.

4.2.2 Verifiche *in vitro*

Per le verifiche *in vitro* sono state usate la corsa in gel di acrilamide delle proteine (SDS-PAGE) estratte dalle foglie inoculate e il monitoraggio della concentrazione del genoma virale tramite qPCR.

SDS-PAGE

La corsa in gel di acrilamide al 10% delle proteine estratte dai campioni biologici mostrano l'incremento della concentrazione della proteina GFP all'interno del tessuto. La Figura 4.3 mostra il risultato di una corsa elettroforetica dopo colorazione con blu di Coomassie e dimostra che la proteina

GFP aumenta la propria concentrazione all'aumentare del tempo. Si nota inoltre che la GFP è assente nei campioni di controllo dal momento che il plasmide utilizzato conteneva una forma troncata e non funzionale della proteina.

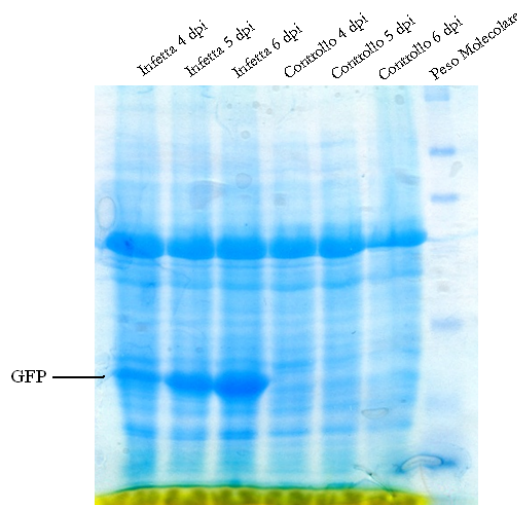


Figura 4.3: Saggio in vitro dell'infezione virale. È facilmente visibile l'incremento temporale della banda della proteina GFP

PCR quantitativa

Per confermare in modo diretto che il genoma virale aumenta la propria concentrazione all'interno delle cellule dell'ospite, il genoma del TVCV è stato amplificato tramite PCR quantitativa. Il calcolo della variazione della concentrazione del genoma virale è stato eseguito mediante il nuovo metodo di analisi Logit, il software di gestione del termociclatore e di analisi ABI 7000 SDS e il programma gratuito LinReg [Ramakers et al., 2003]. I risultati delle analisi riportati in Figura 4.4 mostrano che tutti e tre i metodi di analisi danno degli andamenti molto simili. Il metodo di analisi Logit restituisce dei valori di concentrazione più elevati dal momento che è capace, a differenza degli altri, di confrontare la concentrazione dei geni bersaglio al ciclo iniziale

di reazione. Infatti, con il procede della reazioneme di PCR le differenze di concentrazione tra i geni bersaglio diminuiscono. La possibilità di calcolo della concentrazione nelle prime fasi della reazione consente l'individuazione di valori più vicini alla realtà.

Nelle fasi iniziali dell'esperimento, la concentrazione del genoma virale rimane pressochè costante. Ciò indica che l'infezione non è ancora ben stabilita. Solo dopo 4.5 giorni dall'inoculo la concentrazione virale aumenta in modo esponenziale, per culminare 24 ore dopo con il picco a 5.5 giorni dopo l'inoculo. In questa fase, l'infezione è ben stabilita all'interno delle cellule. Il calo della concentrazione virale misurato a 6 giorni è probabilmente il risultato dell'attivazione dei meccanismi di difesa di *N. benthamiana*. I *primer* utilizzati per questa analisi sono riportati in Tabella 4.1.

Tabella 4.1: *Primer* utilizzati in qPCR

Nome	Sequenza 5' – 3'
<i>Primer for.</i> TVCV	TATGGCAGCAAAGGTGGTC
<i>Primer rev.</i> TVCV	GGAACGGAGCATCATTTTGT
<i>Primer for.</i> EF1- α	CCCTGGACACAGGGACTTTA
<i>Primer rev.</i> EF1- α	CGGGTCTGTCCATCCTTAGA

4.3 Verifica della qualità dell'RNA

L'RNA estratto dalle foglie inoculate è stato analizzato per valutare la sua degradazione e l'eventuale contaminazione da DNA genomico, carboidrati e proteine. In Figura 4.5 è riportato il risultato di una analisi condotta con il Bioanalyzer 2100 (Agilent). Il grafico a sinistra mostra i picchi di assorbimento che lo strumento rileva e che si possono associare a diversi RNA ribosomali (rRNA) presenti in grosse quantità in tutte le cellule. Si nota inoltre che l'RNA non è degradato dal momento che il segnale tra i picchi ritorna velocemente a valori prossimi al rumore di fondo e che non ci sono picchi associati a RNA di piccole dimensioni.

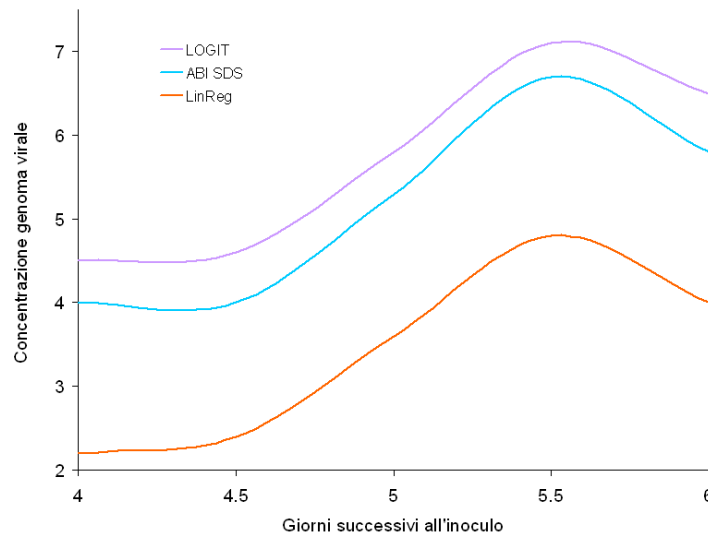


Figura 4.4: Saggio in vitro dell'infezione virale. Confronto dei tre diversi metodi di analisi utilizzati. Sull'asse delle ordinate viene riportato il logaritmo della concentrazione del genoma virale calcolato mediante il confronto con il gene EF1- α

La purezza dell'RNA è stata misurata tramite l'uso del NanoDrop prestando particolare attenzione al rapporto tra il valore di assorbanza misurato a 260 nm e 280 nm (P) e al rapporto dell'assorbanza misurata a 260 nm e 230 nm (C). Il rapporto P misura la contaminazione da DNA e proteine nella soluzione e solamente i valori di P compresi tra 2 e 2.2 sono da considerarsi accettabili per campioni di RNA da destinare ad analisi mediante qPCR o *microarray*. Il rapporto C invece misura la concentrazione di carboidrati presenti in soluzione. Sono stati scelti solamente i campioni con un valore di C superiore a 2 dal momento che una contaminazione da zuccheri pregiudica l'efficienza della reazione di marcatura competendo con il cDNA per il legame con le molecole fluorescenti durante un esperimento di *microarray*.

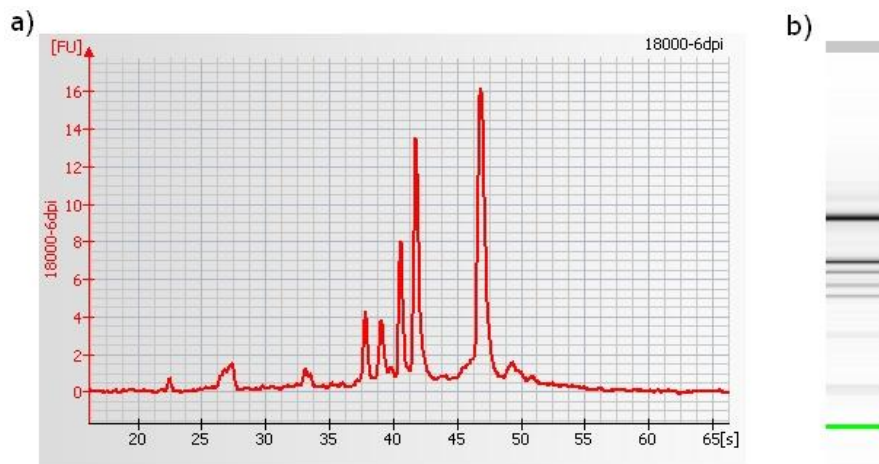


Figura 4.5: Analisi della qualità dell'RNA condotta con il Bioanalyzer 2100. a) Spettro di assorbimento del segnale misurato dallo strumento. I picchi più intensi corrispondono ad rRNA di varie dimensioni che si ritrovano in grosse quantità in tutte le cellule vegetali. Questo saggio dimostra come l'RNA non sia degradato dal momento che il segnale misurato tra i picchi ritorna velocemente a valori prossimi al rumore di fondo e che non sono presenti picchi da associare a RNA degradato di piccole dimensioni. b) Rappresentazione grafica simile alla corsa in gel di elettroforesi dello spettro di assorbimento misurato dallo strumento

4.4 Ibridazioni ed analisi dei dati

Il disegno sperimentale utilizzato ha avuto come scopo quello di enfatizzare la variazione di espressione dovuta ai due fattori principali, l'infezione virale e il tempo trascorso, e di ridurre al minimo la variazione di espressione dovuta al metodo sperimentale ed alle differenze intraspecifiche di *N. benthamiana*. Per ogni pianta sono state inoculate due foglie della medesima dimensione ed appartenenti al medesimo nodo in modo da limitare la variabilità nell'inoculo. In aggiunta, sono state eseguite tre repliche biologiche per ogni *time point* in modo tale da limitare la variabilità intraspecifica ed essere in grado di generalizzare i risultati. Per ridurre le variazioni di espressione dovute alla particolare tecnica di infezione utilizzata, le piante controllo sono state trattate esattamente come le piante infettate a differenza del plasmide utilizzato. Infatti, nelle piante infettate è stato usato il plasmide pICH18000 contenente il genoma virale, mentre per le piante controllo è stato usato il plasmide pICBV24GFP senza il genoma virale e con una proteina GFP non funzionale. Infine, per limitare le differenze di fluorescenza emessa dalle due molecole utilizzate per la marcatura dei campioni (Cy3 e Cy5), sono state eseguite due repliche tecniche per ogni ibridazione invertendo la marcatura. In questo modo ogni campione è stato marcato con entrambi i fluorocromi. In totale sono state eseguite 30 ibridazioni.

Le ibridazioni dei campioni con i *microarray* sono avvenute in condizioni controllate ad una temperatura di 40 °C per almeno 16 ore. Successivamente i vetrini sono stati lavati con tre soluzioni tampone differenti che permettono l'eliminazione del rumore di fondo e delle ibridazioni aspecifiche tramite un graduale aumento della stringenza. Quindi, i vetrini sono stati immediatamente letti con lo scanner GenePix 4200A e le immagini così ottenute (una per ogni fluorocromo) sono state analizzate con il software GenePix Pro.

La digitalizzazione delle intensità luminose effettuata dal software GenePix Pro non consente un confronto immediato dei dati per la ricerca dei geni differenzialmente espressi. Infatti, come si nota dal grafico a sinistra della Figura 4.6, i dati grezzi presentano un'enorme variabilità interna che pregiudica

ogni analisi. Essa è dovuta solo in minima parte alle differenze di espressione genica che si possono riscontrare tra le repliche biologiche. La maggior parte delle differenze nei segnali luminosi è dovuta da una parte alla variazione stocastica nella resa nella preparazione dei campioni e dall'altra alla variazione casuale dell'efficienza di ibridazione delle molecole sonda. Infatti, pur seguendo criteri di qualità rigorosi nella selezione del materiale di partenza, l'efficienza della retrotrascrizione e della marcatura varia tra i diversi campioni traducendosi in una variabilità nella concentrazione e luminosità del bersaglio. Anche l'ibridazione presenta alcuni aspetti di casualità. Alcune aree del vetrino possono presentare un rumore di fondo particolarmente elevato o un segnale luminoso mediamente più basso. Queste regioni a bassa qualità non sono correlate con un particolare trattamento subito dai vetrini, dal momento che non avvengono sempre nelle stesse zone e nel medesimo gruppo di vetrini. Per limitare le differenze dovute a questi due fattori, l'espreimento di ibridazione è stato eseguito suddividendolo in tre grossi gruppi corrispondenti alle tre repliche biologiche. Non è mai accaduto che un intero gruppo presentasse sempre il medesimo problema ma alcuni vetrini all'interno di un gruppo hanno una qualità leggermente inferiore rispetto alla media degli altri.

Quindi prima di procedere con la ricerca dei geni differenzialmente espressi, i dati grezzi necessitano di una correzione della distribuzione delle intensità luminose. Ai dati sono stati applicati tre processi di normalizzazione sequenziali che hanno lo scopo di ridurre le differenze all'interno di ogni vetrino (come aree più o meno luminose o con forte rumore di fondo) e tra i vetrini. Il primo processo di normalizzazione sottrae al segnale di ogni punto il rumore di fondo calcolato nelle immediate vicinanze del punto stesso; il secondo limita la differenza intrinseca di intensità luminosa tra i due fluorocromi; mentre il terzo rende uniforme la varianza tra i vetrini. I risultati dei processi di normalizzazione possono essere facilmente osservati confrontando i due grafici di Figura 4.6. Si nota infatti come la grande differenza tra i vetrini presente nei dati grezzi sia completamente corretta dalla normaliz-

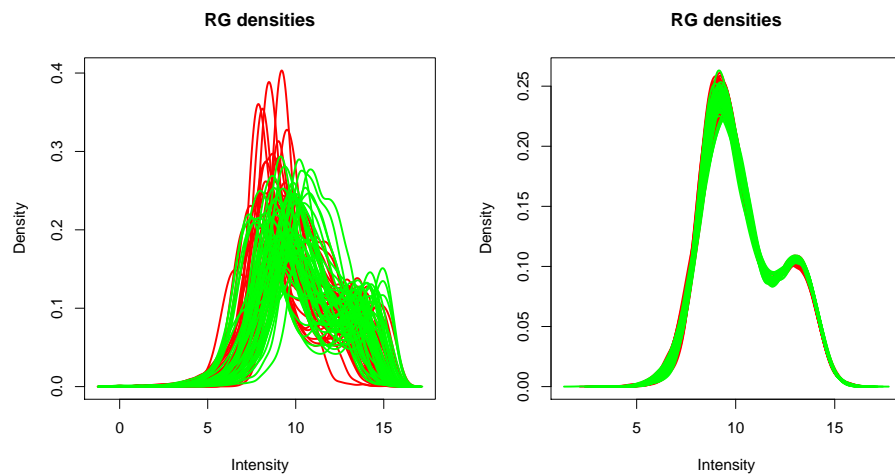


Figura 4.6: Densità di fluorescenza di entrambi i fluorocromi in tutti i vetrini analizzati prima e dopo normalizzazione. Nei grafici viene rappresentata la percentuale di punti in funzione della loro intensità luminosa. Il grafico a sinistra riporta le intensità luminose dei dati grezzi e dimostra la presenza di una enorme variabilità tra i vetrini e tra i fluorocromi utilizzati. Da notare come il fluorocromo rosso sia mediamente più intenso del verde. A destra viene riportato il grafico dopo normalizzazione nel quale si nota come tutti i vetrini abbiano un comportamento molto simile permettendo un confronto corretto tra i *microarray*

zazione.

A questo punto i dati sono pronti per la ricerca dei geni differenzialmente espressi. A tale scopo è stata usata una variante del comune test t di Student che riduce la probabilità di identificazione di falsi positivi (FDR). Il numero complessivo dei geni differenzialmente espressi individuato ad ogni *time point* viene riportato nella tabella 4.2, mentre la lista completa dei geni viene riportata in Appendice A.

Tabella 4.2: Numero dei geni differenzialmente espressi

	4 dpi	4.5 dpi	5 dpi	5.5 dpi	6 dpi
Geni sottoespressi	8	23	25	62	187
Geni sovraespressi	12	55	33	56	122

4.5 Interpretazione dei risultati

Una delle fasi più importanti e più complesse di un esperimento di *microarray* è l'interpretazione dei risultati. Gli esperimenti di ibridazione individuano un grande numero di geni differenzialmente espressi e capire come questi operino all'interno della cellula è un'impresa che può rivelarsi difficile e laboriosa. Bisogna inserire le variazioni di espressione nel contesto più generale dell'intera cellula vegetale, cercando di delineare con precisione il quadro dei cambiamenti fisiologici che la pianta sta mettendo in atto per difendersi dall'infezione virale. Lo studio dettagliato di ogni singolo gene è un processo molto lungo e non sempre è la soluzione migliore. Bisogna cercare di riassumere il comportamento della cellula cercando di classificare i cambiamenti di espressione in base alle loro funzioni piuttosto che studiando il singolo gene, almeno in una prima analisi. Per farlo, quindi, bisogna raggruppare i geni seguendo criteri rigorosi in modo da avere una buona confidenza delle interpretazioni che ne derivano. Uno dei metodi più utilizzati in esperimenti di questo tipo è l'assegnazione dei geni a classi funzionali che seguono la nomenclatura *Geneontology* (GO) [Harris et al., 2004]. GO è un progetto

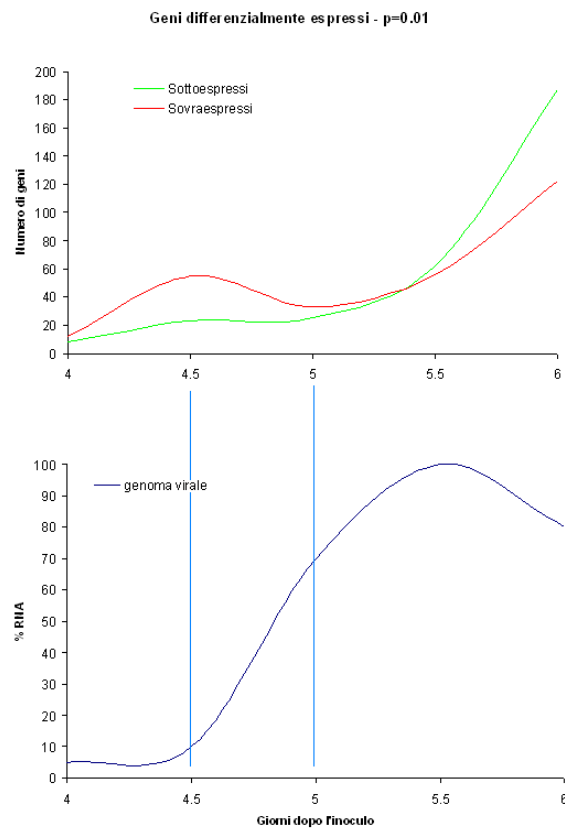


Figura 4.7: Correlazione tra il cambio dell'espressione genica e l'incremento della concentrazione del virus nelle cellule

internazionale che classifica i geni in funzione di tre importanti contesti biologici: la loro funzione molecolare, i processi fisiologici a cui appartengono e la loro localizzazione cellulare. In questo modo ad ogni gene viene assegnato un numero che lo classifica in base alla sua funzione nella cellula e alla sua localizzazione. Per implementare l'interpretazione dei risultati ho utilizzato anche il software MapMan [Thimm et al., 2004] che, utilizzando le definizioni della GO, inserisce i geni in contesti molto precisi, come la via metabolica di appartenenza, o più generali, come nello sviluppo cellulare.

Il primo passo di quest'analisi è stata un'osservazione del numero di geni durante le fasi dell'esperimento. Per prima cosa si nota un incremento del numero di geni che vengono coinvolti nella risposta della pianta in funzione

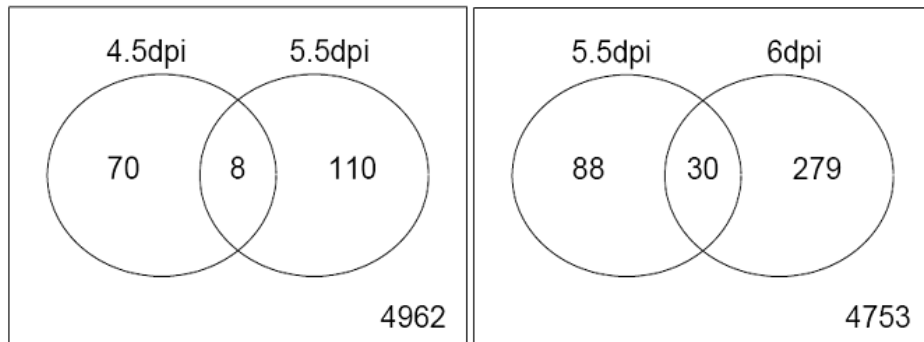


Figura 4.8: Diagrammi di Venn che dimostrano come il picco iniziale e le fasi terminali dell'infezione non condividano un grosso numero di geni. Nelle fasi finali dell'infezione il numero di geni condivisi aumenta

del tempo (Tabella 4.2). Questo comportamento è da ritenersi del tutto normale: la pianta, quando comincia a percepire l'infezione, impiega del tempo per attivare tutti i meccanismi di risposta e ciò avviene in maniera graduale nel tempo. L'incremento però non è lineare ma raggiunge un primo picco a 4.5 giorni successivi all'inoculo (Figura 4.7, grafico in alto). Tra 4.5 e 5 giorni dopo l'inoculo si assiste ad una fase di calo della risposta per riprendere 5.5 dpi e aumentare gradualmente a 6 dpi. Un'analisi più approfondita del picco iniziale e delle fasi terminali rivela che esse condividono un basso numero di geni, mentre le fasi terminali della risposta (5.5 dpi e 6 dpi) condividono un numero più elevato di geni. Infatti, come si nota dai due diagrammi di Venn della Figura 4.8, i due picchi hanno circa il 10% dei geni in comune, mentre il 30% dei geni identificati a 5.5 giorni viene identificato anche a 6 giorni dopo l'inoculo. Dai grafici della Figura 4.7 emerge anche che sussiste una relazione tra la concentrazione del genoma virale all'interno della cellula e le due fasi. Quando l'infezione virale è ancora nella prima fase, la concentrazione del suo genoma nella cellula non cresce molto velocemente. Solo dopo 4.5 giorni comincia una crescita esponenziale della replicazione del genoma virale che culmina a 5.5 giorni. Si nota allora che il passaggio tra la prima fase di risposta e la seconda coincide con l'inizio della crescita esponenziale del genoma virale nelle cellule dell'ospite.

Da alcuni dati preliminari non pubblicati del professor Andy Moule (*John Innes Centre, UK*) emerge che *Agrobacterium* sfrutta dei meccanismi di infezione simili ai meccanismi anti-attenuazione dell'RNA presenti nei virus. Infatti, come i virus, anche *Agrobacterium* sfrutta i meccanismi di espressione genica dell'ospite per infettare il tessuto. È allora plausibile ipotizzare che il ceppo di *Agrobacterium* contenente il plasmide pICH18000 sfrutti i meccanismi anti-attenuazione di TVCV per infettare e moltiplicarsi nell'ospite. Quindi, quando la concentrazione virale all'interno delle cellule dell'ospite è ancora bassa, *N. benthamiana* percepisce la crescita di *Agrobacterium* e attua una risposta di difesa che è maggiormente diretta nei suoi confronti, provocando il picco di risposta a 4.5 giorni. Solamente quando la concentrazione dell'RNA virale all'interno della cellula cresce in modo esponenziale la risposta della pianta cambia e si concentra contro l'infezione virale generando il picco delle fasi terminali dell'esperimento.

Un'analisi più dettagliata della risposta fisiologica che la pianta attua durante le due fasi sembra avvalorare questa ipotesi. Durante il primo picco di risposta, quando l'infezione virale non è ancora ben stabilita, si assiste ad una sovra-espressione dei geni importanti per l'ispessimento della parete cellulare e alla produzione di composti ad azione antimicrobica. Entrambe queste vie metaboliche sono tipiche di infezioni batteriche o fungine [La Camera et al., 2004; Zabala et al., 2006; Maher et al., 1994]. Queste stesse vie metaboliche sono bloccate nelle fasi più tardive quando l'infezione virale ha raggiunto il suo massimo. Il gene *Coumarate-CoA Ligase* (CK299019 e CK288563), importante per la sintesi della lignina, è sovraespresso a 4.5 giorni e sottoespresso a 5.5. Inoltre un gene della famiglia delle *transferasi* (G10-j001-plate-124), importante nella biosintesi dei fenilpropanoidi, è sovraespresso a 4.5 giorni. Questa via metabolica porta alla produzione di numerose molecole ad azione antimicrobica e all'ispessimento della parete cellulare. Il gene *endo-xyloglucan transferase* (F07-j001-plate-126), responsabile delle modifiche di parete tramite aggiunta e rimozione di unità di xyloglucano, viene sottoespresso a 5.5 giorni. Invece il gene *(1-4)-beta-mannan endohy-*

drolase (D09-j001-plate-213), importante per la degradazione della parete cellulare, è sovraespresso a 5.5 giorni.

Uno dei modi in cui le piante cercano di proteggersi da un'infezione virale è un rallentamento della sintesi delle proteine e dell'RNA e la degradazione delle proteine virali mediata dall'ubiquitina e dalle proteine *heat shock* [Gorovits et al., 2007; Aparicio et al., 2005; Senthil et al., 2005; Aranda et al., 1999; Aranda and Maule, 1998]. Infatti il forte tasso di replicazione virale e la produzione elevata delle sue proteine costringe la cellula infettata a attuare dei meccanismi che ne limitino i danni. Risposte simili sono riscontrate durante gli stress termici dal momento che in entrambi i casi sono presenti proteine con conformazioni strutturali non riconosciute come normali. Durante le fasi finali della risposta, si assiste infatti ad una repressione sia della sintesi dell'RNA che della sintesi proteica. L'espressione di 5 geni importanti per la sintesi o il processamento dell'RNA viene repressa nelle fasi finali della risposta (*RNase L inhibitor*, C08-j001-plate-71; *U-box domain-containing protein*, CK287506; *pathogenesis-related transcriptional factor and ERF*, CK297104; *EXO70 gene family*, CK292956; *scarecrow transcription factor family*, CK289500). Contemporaneamente viene repressa l'espressione di tre geni coinvolti nella sintesi proteica (*Putative peptide chain release factor subunit 1 protein*, CK287412; *chloroplast 30S ribosomal protein*, E08-j001-plate-215; *60S ribosomal protein L19*, A04-j001-plate-126). Inoltre, sia un gene codificante una proteina *heat shock* (CK285339) che un gene importante nella degradazione proteica mediata dall'ubiquitina (C10-j001-plate-35) sono sovraespressi nelle fasi tardive dell'esperimento. Il gene CK29820 codificante per la proteina CRULs è fondamentale per la degradazione proteica delle proteine mediata dall'ubiquitina ed è sottoespresso a 5.5 giorni. In un recente lavoro è stato notato che la funzione di questo gene viene modificata da numerosi virus in modo da amplificare i meccanismi virali che bloccano l'attenuazione dell'RNA [Barry and Fruh, 2006]. Una sua repressione potrebbe essere una diretta conseguenza di quest'attività. Un comportamento completamente opposto si osserva invece nelle prime fasi della risposta quan-

do vengono sovraespressi sia i geni per la sintesi e il processamento dell'RNA (*putative splicing factor*, CK287370; *IAA9 protein*, H12-a001-plate-66; *transcription factor*, CK297003; *similar to transcriptional factor B3 family protein*, G05-j001-plate-82; *nam-like protein 3*, CK292423; *DNA binding*, CK289726) sia i geni importanti per la sintesi proteica (*ribosomal Pr 117*, CK284181; *ribosomal protein S26*, CK288918).

Non è possibile fare una distinzione netta tra i due picchi dal momento che nelle fasi iniziali sono comunque presenti dei geni tipici delle infezioni virali come una proteina *heat shock* (CK293224) mentre nelle fasi finali è stato identificato un gene importante per la risposta contro le infezioni batteriche, la chitinasi (CK289519). Questo gene codifica per una proteina che degrada la chitina un componente fondamentale della parete delle cellule batteriche. Il confronto della risposta di *N. benthamiana* nelle due fasi è stato confrontato con il database AtGeneExpress presente nel software Genevestigator [Grennan, 2006]. Questo database contiene i dati di più di 3000 esperimenti di ibridazione con chip Affymetrix condotti su *A. thaliana* sotto l'azione di numerosi stress sia biotici che abiotici. I dati presenti in questo database derivano tutti dal medesimo organismo, sono stati misurati con la medesima tecnica e sono stati analizzati con il medesimo algoritmo di normalizzazione. Per queste ragioni vengono considerati come dei valori molto affidabili e precisi. Il confronto rivela che circa il 65% dei geni identificati nelle prime fasi si ritrovano anche nella risposta di *A. thaliana* verso gli stress biotici e più in particolare verso l'infezione di *Agrobacterium*. Nelle fasi finali invece la maggior parte dei geni appartengono alla risposta di tipo abiotico di stress indotto da elevate temperature.

Alla luce di tutte queste informazioni è possibile formulare alcune ipotesi su come *N. benthamiana* reagisca ad un'infezione virale condotta mediante questa particolare tecnica di infezione. Nelle fasi immediatamente successive all'inoculo sia nella pianta controllo che in quella infettata, *Agrobacterium* riconosce delle proteine di parete dell'ospite e tramite i geni della classe *vir* inserisce il genome del virus o un DNA non funzionale nel genoma della pi-

anta. In questa fase probabilmente la pianta risponde all'azione di *Agrobacterium* ma la risposta viene mascherata dalla presenza delle piante controllo trattate come le piante infettate. Successivamente il genoma del virus viene trascritto in RNA in modo tale da diventare completamente attivo e comincia a replicare all'interno della cellula dell'ospite. A questo punto il virus modifica il metabolismo dell'ospite mediante l'azione dei meccanismi virali che limitano l'azione del *silencing* dell'RNA. In queste condizioni di abbassamento delle difese della pianta, *Agrobacterium* riesce a moltiplicarsi. Questo processo avviene solamente nelle piante infettate con il virus e per questa ragione è possibile misurare nelle prime fasi della risposta una variazione dell'espressione dei geni tipici delle infezioni batteriche. Solamente quando l'infezione virale è ben stabilita, *N. benthamiana* comincia a rispondere in modo più diretto verso quest'ultima. In questa fase si assiste ad un cambio dell'espressione in *N. benthamiana* che porta alla produzione dei geni tipici delle infezioni virali fino a giungere al picco di risposta delle fasi tardive dell'esperimento.

Capitolo 5

Conclusioni

Lo studio approfondito dei meccanismi di difesa che la pianta attua contro le infezioni virali contribuisce sia alla scoperta di nuovi geni implicati nella difesa, sia di geni che vengono regolati dal patogeno e che sono importanti per il suo ciclo vitale. Entrambi questi gruppi di geni possono essere utilizzati come bersaglio per la produzione di organismi vegetali più resistenti o di pesticidi più specifici. Durante il periodo di dottorato ho voluto studiare il cambiamento fisiologico che avviene in *Nicotiana benthamiana* durante le prime fasi della risposta contro l'infezione del *Turnip Vein-Clearing Virus*. La finestra temporale è stata scelta in base alla concentrazione virale all'interno della cellula. Le prime fasi dell'esperimento rilevano una concentrazione virale molto bassa mentre le fasi finali corrispondono ad una concentrazione che ha raggiunto i valori massimi. In questo modo ho voluto investigare come la cellula vegetale percepisce la crescita della concentrazione virale e che meccanismi attua per difendersi.

Uno dei maggiori ostacoli da dover superare è stato l'impossibilità di infezione diretta della pianta, dal momento che solamente una piccola percentuale di cellule entra in contatto diretto con il virus. Per aggirare questo ostacolo ho deciso di utilizzare un approccio completamente diverso tramite l'utilizzo di un *carrier* che fosse in grado di inserire il genoma virale all'interno della cellula vegetale con un'efficienza elevata. A tale scopo ho utilizzato *Agrobac-*

terium, un battere patogenico che ha la capacità di inserire un frammento di DNA nel genoma dell'ospite. In questo modo il genoma virale viene integrato nel genoma vegetale e la sua trascrizione in RNA lo rende completamente attivo e libero di completare il ciclo di infezione. L'utilizzo di questo meccanismo di infezione da un lato ha aumentato notevolmente la percentuale di cellule infettate ma dall'altro ha aggiunto all'esperimento la variabile della risposta che la pianta attua contro *Agrobacterium*. Per ridurre questa variabile ho deciso di trattare le piante controllo esattamente come quelle infettate, utilizzando in questo caso un ceppo di *Agrobacterium* contenente un frammento di DNA non attivo. In questo modo sia le piante infettate che quelle controllo dovrebbero rispondere nello stesso modo alla presenza del battere, consentendo di individuare geni specifici alla risposta virale.

Ho verificato l'infezione virale con metodi sia *in vivo* che *in vitro*. Ho valutato la percentuale di cellule venute in contatto con il virus misurando la fluorescenza emessa dal gene *reporter* GFP inserito a valle del genoma virale. La trascrizione di questo gene è diretta conseguenza della trascrizione del genoma virale dal momento che entrambi vengono regolati dal medesimo promotore. Quindi se una cellula risulta fluorescente significa che contiene la GFP e di conseguenza contiene anche il genoma virale attivo. Con questo semplice saggio *in vivo* sono stato capace di evienziare che circa il 90% delle cellule del tessuto inoculato contengono il virus. Con le verifiche *in vitro* della concentrazione della GFP e del virus sono stato capace di monitorare l'aumento della concentrazione virale all'interno delle cellule. La crescita del virus non segue un andamento lineare, ma ad una prima fase rallentata ne segue una di crescita esponenziale che dura circa 24 ore. La fine della fase esponenziale corrisponde anche al momento in cui la concentrazione virale all'interno delle cellule ha raggiunto il valore massimo. Ho eseguito la misura della concentrazione del genoma virale con la tecnica della PCR quantitativa per la quale ho sviluppato un nuovo metodo di analisi. Il confronto dei risultati di questo nuovo metodo con altri due largamente utilizzati ha dimostrato l'adeguatezza del metodo di calcolo.

Ho studiato il cambiamento fisiologico di *N. benthamiana* con la tecnica delle ibridazioni con *microarray* di cDNA che è stato appositamente sviluppato. Questo strumento consente l'analisi simultanea di 6000 sequenze specifiche di *N. benthamiana*. Ho usato il disegno sperimentale *time course* tra 4 e 6 giorni successivi all'inoculo con intervalli di 12 ore, per un totale di 30 ibridazioni. I risultati delle analisi dimostrano la presenza di due fasi di risposta differenti. La prima raggiunge il suo massimo in concomitanza con l'inizio della fase di crescita esponenziale del genoma virale. La seconda invece reggiunge i valori masimi alla fine della fase di crescita. Uno studio dettagliato dei geni presenti in queste due fasi mi ha fatto ipotizzare una loro funzione diversa nella difesa della pianta. Il primo picco contiene dei geni che sono caratteristici di infezioni batteriche e un confronto con la risposta di *A. thaliana* ne accentua la somiglianza con una risposta diretta verso *Agrobacterium*. Ho ritenuto che questa risposta sia dovuta al fatto che il battere utilizza dei meccanismi anti-attenuazione dell'RNA simili a quelli utilizzati dal virus per infettare le piante. *Agrobacterium* sarebbe dunque in grado di sfruttare il virus per infettare a sua volta l'ospite. La seconda fase della risposta, invece, contiene geni caratteristici di un'infezione virale. Da un lato viene bloccata completamente la trascrizione dell'RNA e la traduzione proteica, dall'altro viene incrementata la degradazione proteica mediata dall'ubiquitina e dalle proteine della classe *heat shock*. Infatti il forte tasso di replicazione virale e la produzione elevata delle sue proteine costringe la cellula infettata ad attuare dei meccanismi che ne limitino i danni. Le due risposte non sono completamente separate dal momento che nelle prime fasi si trovano geni tipici di un'infezione virale, mentre nella fase tardiva si trovano geni tipici delle infezioni batteriche.

Questo lavoro ha aperto la strada allo studio di due aspetti biologici diversi. Da un lato si sono identificati numerosi geni la cui importanza nella risposta all'infezione va confermata ad esempio tramite la produzione di piante *knock-out* in modo da valutarne l'importanza durante la risposta all'infezioni virali. Un altro aspetto che merita uno studio approfondito è la conferma che

Agrobacterium sfrutti dei meccanismi anti-attenuazione simili a quelli del TVCV per infettare le cellule vegetali. Ad esempio si potrebbe inoculare una pinata con il battere contenete un plasmide con la proteina responsabile del meccanismo anti-attenuazione del TVCV. In questo modo se il battere riesce a crescere e ad infettare la pianta sarebbe dimostrata la validità della mia ipotesi.

Appendice A

Di seguito vengono riportate le liste complete dei geni differenzialmente espressi ai diversi tempi.

Tabella 5.1: Geni differenzialmente espressi a 4 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
A12-j001-plate-197	Putative protein kinase tousel	-2.59	8.17	-5.70	2.59E-03	3.87
B12-j001-plate-8	PTEN-like protein	-1.02	9.17	-5.25	3.63E-03	3.46
E11-j001-plate-173	No similarity found	-0.80	9.65	-5.49	2.53E-03	4.11
B02-j001-plate-115	Cyclin-like F-box	-0.67	10.14	-5.53	2.28E-03	4.30
CK290484	Unknown protein	-0.61	13.44	-6.41	2.68E-04	6.82
B09-j001-plate-90	No similarity found	-0.57	11.37	-5.13	3.64E-03	3.29
CK287604	Protein binding	-0.43	11.71	-5.15	3.64E-03	3.34
CK287492	FTSH-TOBACCell division protein ftsH homolog; chloro- plast precursor (DS9)	-0.31	11.31	-4.94	5.577E-03	2.77
G03-j001-plate-87	No similarity found	0.28	12.07	5.26	3.45E-03	3.57
E01-j001-plate-121	No similarity found	0.28	12.79	5.10	3.64E-03	3.21
B03-j001-plate-124	RNA binding / pseudouridine synthase/ synthase	0.31	12.88	6.57	2.68E-04	7.25
CK290539	No similarity found	0.59	9.83	5.15	3.64E-03	3.26
D04-j001-plate-84	Integrase, catalytic region	0.86	10.54	6.79	2.21E-04	7.85
CK294271	At1g03160	1.14	10.19	4.77	0.01	2.12
D11-j001-plate-148	RelA-SpoT like protein RSH4	1.84	8.35	4.95	7.30E-03	2.46
D10-j001-plate-66	No similarity found	1.87	8.82	5.96	1.58E-03	4.79
B01-j001-plate-206	No similarity found	1.94	9.82	5.66	1.58E-03	4.75
A11-j001-plate-92	No similarity found	2.01	8.18	6.98	2.68E-04	6.93

Continua...

Tabella 5.1: Geni differenzialmente espressi a 4 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
CK292657	Unknown protein	2.22	7.83	5.43	5.81E-03	2.71
A06-a001-plate-117	Hypoxia-responsive family protein	6.39	11.79	7.97	1.88E-05	10.70

^a Codice identificativo interno del laboratorio

^b Proteine con la più alta omologia. Ricavato da BLASTx

^c M: valore medio dei $\log_2(T/C)$, dove T è il segnale della pianta trattata e C del controllo. Indica la variazione di espressione del gene dovuta all'infezione virale.

^d A: valore medio del $\log_2(T^*C)$. Indica per ogni gene il valore medio di espressione.

^e Valore del test t di Student.

^f Probabilità associata al valore t

^g È il *log-odds* che il gene sia differenzialmente espresso. Un valore di B pari a zero indica una probabilità del 50% di differenza di espressione. Per valori positivi la probabilità aumenta (per B=1.5 la probabilità sale al 82%), per negativi diminuisce

Tabella 5.2: Geni differenzialmente espressi a 4.5 giorni

ID ^a	Funzione probabile ^b	M ^c	A ^d	t ^e	P ^f	B ^g
E03-j001-plate-8	unknown protein	-1.74	8.47	-4.53	8.69E-03	1.21
G02-j001-plate-165	No similarity found	-1.67	8.96	-5.99	3.55E-04	5.58
A12-a001-plate-104	No similarity found	-1.19	9.22	-5.41	1.47E-03	3.81
D03-j001-plate-143	No similarity found	-0.96	9.58	-4.52	6.62E-03	1.48
CK287372	unknown	-0.94	11.00	-6.36	1.32E-04	6.72
CK287604	protein binding	-0.72	11.72	-8.56	7.35E-07	12.64
E03-j001-plate-84	cathepsin B-like cysteine proteinase	-0.60	10.76	-5.17	1.97E-03	3.36
A08-j001-plate-89	No similarity found	-0.58	10.42	-4.31	9.50E-03	0.96
B11-j001-plate-37	No similarity found	-0.58	9.75	-4.27	9.99E-03	0.85
B09-j001-plate-90	No similarity found	-0.56	11.37	-5.02	2.13E-03	2.94
CK290484	unknown protein	-0.49	13.44	-5.09	2.07E-03	3.15
CK283646	Blood group Rhesus C/E and D polypeptide; Rh-like protein/ammonium transporter	-0.47	10.60	-4.60	5.25E-03	1.76
C09-a001-plate-179	electron transporter/ thiol-disulfide exchange intermediate	-0.46	13.71	-5.43	1.19E-03	4.10
F11-j001-plate-155	No similarity found	-0.46	12.80	-4.31	9.50E-03	0.98
G01-j001-plate-206	putative polyprotein	-0.44	10.62	-4.65	4.54E-03	1.91
CK295879	unknown protein	-0.43	12.10	-5.88	4.14E-04	5.37
C09-j001-plate-151	glycine hydroxymethyltransferase (EC 2.1.2.1)-like protein	-0.39	12.25	-4.56	5.58E-03	1.67
B01-a001-plate-51	No similarity found	-0.36	11.94	-4.51	6.23E-03	1.53
G12-j001-plate-1	unknown	-0.32	12.88	-4.90	2.56E-03	2.60

Continua...

Tabella 5.2: Geni differenzialmente espressi a 4.5 giorni

ID ^a	Funzione probabile ^b	M ^c	A ^d	t ^e	P ^f	B ^g
H05-j001-plate-200	No similarity found	-0.27	12.52	-4.39	8.18E-03	1.19
CK292627	putative phosphoribosylanthranilate transferase, 3' partial; 131493-134402	-0.27	14.39	-4.37	8.49E-03	1.14
CK283491	Cyclin-like F-box	-0.26	13.28	-4.29	9.89E-03	0.90
CK285383	unknown protein	-0.24	13.26	-5.17	1.97E-03	3.36
CK286842	putative proteasome regulatory non-ATPase subunit	0.23	13.07	5.12	1.99E-03	3.22
CK295599	putative cell survival CED-4-interacting protein MAC-1	0.25	13.19	4.80	2.99E-03	2.34
CK298371	putative transposase	0.26	12.83	4.81	2.99E-03	2.36
H12-a001-plate-66	IAA9 protein	0.27	13.03	4.36	8.53E-03	1.12
G01-j001-plate-161	No similarity found	0.30	13.61	5.04	2.13E-03	2.99
F07-j001-plate-94	ATTPS1 (TREHALOSE-6-PHOSPHATE SYNTHASE); transferase, transferring glycosyl groups	0.31	13.04	5.05	2.13E-03	3.01
CK289936	fw2.2	0.31	12.25	5.14	1.98E-03	3.29
CK283705	unknown protein	0.32	13.05	5.26	1.66E-03	3.63
CK298563	4-coumarate:CoA ligase 4	0.33	13.62	4.49	6.62E-03	1.45
CK287370	putative splicing factor, arginine/serine-rich 7	0.36	11.44	4.27	9.99E-03	0.86
D10-j001-plate-185	No similarity found	0.38	13.27	4.83	2.96E-03	2.40
B05-j001-plate-206	No similarity found	0.39	12.75	5.00	2.14E-03	2.88
CK291025	Glyoxalase/bleomycin resistance protein/dioxygenase	0.40	13.12	5.01	2.13E-03	2.93
G10-j001-plate-124	transferase	0.41	12.79	4.44	7.34E-03	1.32
CK288001	unknown protein	0.45	12.91	5.42	1.19E-03	4.06

Continua...

Tabella 5.2: Geni differenzialmente espressi a 4.5 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
C11-j001-plate-103	Integrase, catalytic region; Zinc finger, CCHC-type;	0.45	12.75	5.50	1.08E-03	4.29
	Peptidase aspartic, catalytic					
CK291411	No similarity found	0.46	11.06	4.52	6.17E-03	1.56
CK295607	beta-D-glucosidase	0.46	10.64	4.58	5.41E-03	1.71
CK293769	unknown protein	0.55	12.36	4.87	2.67E-03	2.52
CK284367	unknown protein	0.57	10.27	5.02	2.13E-03	2.94
CK294620	transporter	0.58	12.21	4.35	8.69E-03	1.09
E01-j001-plate-130	cytochrome P450 monooxygenase CYP72B	0.59	13.03	5.13	1.99E-03	3.24
CK296050	glutamate dehydrogenase	0.60	12.35	5.95	3.55E-04	5.57
C04-j001-plate-87	receptor-like protein kinase 3	0.61	13.52	4.96	2.31E-03	2.78
CK293503	unknown protein	0.68	11.25	4.81	2.99E-03	2.35
CK292854	unknown protein	0.71	11.17	5.66	7.34E-04	4.77
CK283792	Ubiquitin family, putative	0.71	13.79	16.13	5.37E-14	27.77
CK292423	nam-like protein 3	0.75	11.36	7.01	2.93E-05	8.53
CK288918	ribosomal protein S26	0.75	9.84	5.05	2.13E-03	2.97
CK295802	unknown protein	0.76	14.78	6.58	8.75E-05	7.20
CK286548	putative polyprotein	0.76	13.62	5.14	1.98E-03	3.29
CK284181	ribosomal Pr 117	0.77	10.20	7.46	1.08E-05	9.74
CK296072	ALY protein	0.78	13.32	5.39	1.24E-03	3.98
CK293224	putative DnaJ protein	0.80	12.35	5.19	1.97E-03	3.42
E05-j001-plate-103	No similarity found	0.81	9.37	5.06	2.13E-03	2.94
Continua...						

Tabella 5.2: Geni differenzialmente espressi a 4.5 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
C04-j001-plate-49	No similarity found	0.81	10.59	5.04	2.13E-03	3.00
E09-j001-plate-123	No similarity found	0.81	10.27	4.29	9.83E-03	0.92
C12-j001-plate-139	Membrane attack complex component / complement C9	0.99	9.44	4.40	7.99E-03	1.22
A08-j001-plate-49	No similarity found	1.04	10.27	6.34	1.48E-04	6.48
CK285339	DNAJ heat shock N-terminal domain-containing protein	1.08	9.45	4.50	6.94E-03	1.40
B11-j001-plate-140	Calcium-binding EF-hand	1.08	10.09	5.47	1.13E-03	4.21
CK286120	aquaporin 1	1.16	9.84	4.67	4.44E-03	1.95
H09-j001-plate-57	No similarity found	1.21	9.68	4.97	2.53E-03	2.69
D01-j001-plate-86	No similarity found	1.21	9.35	5.05	2.14E-03	2.89
H01-j001-plate-196	electron transporter / thiol-disulfide exchange intermediate	1.26	12.08	5.51	1.08E-03	4.33
D03-j001-plate-167	No similarity found	1.28	10.33	5.01	2.56E-03	2.65
CK297003	transcription factor	1.29	8.93	5.04	2.56E-03	2.69
H09-j001-plate-194	6-4 photolyase	1.35	8.54	4.89	2.67E-03	2.53
A01-j001-plate-17	No similarity found	1.39	9.95	6.79	7.34E-05	7.49
G05-j001-plate-82	similar to transcriptional factor B3 family protein	1.60	8.29	4.58	9.99E-03	1.04
CK289726	DNA binding	1.74	9.85	10.36	1.29E-08	16.15
B06-j001-plate-37	No similarity found	1.80	8.92	6.67	7.34E-05	7.46
H03-j001-plate-49	No similarity found	1.84	9.10	8.27	6.08E-06	10.11
G09-j001-plate-205	unknown protein	2.79	11.31	17.20	3.18E-14	27.29

Continua...

Tabella 5.2: Geni differenzialmente espressi a 4.5 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
A06-a001-plate-117	hypoxia-responsive family protein	5.53	11.80	7.18	2.46E-05	8.83

^a Codice identificativo interno del laboratorio

^b Proteine con la più alta omologia. Ricavato da BLASTx

^c M: valore medio dei $\log_2(T/C)$, dove T è il segnale della pianta trattata e C del controllo. Indica la variazione di espressione del gene dovuta all'infezione virale.

^d A: valore medio del $\log_2(T*C)$. Indica per ogni gene il valore medio di espressione.

^e Valore del test t di Student.

^f Probabilità associata al valore t

^g È il *log-odds* che il gene sia differenzialmente espresso. Un valore di B pari a zero indica una probabilità del 50% di differenza di espressione. Per valori positivi la probabilità aumenta (per B=1.5 la probabilità sale al 82%), per negativi diminuisce

Tabella 5.3: Geni differenzialmente espressi a 5 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
G06-j001-plate-162	No similarity found	-2.24	8.81	-4.74	5.39E-03	2.08
CK285769	putative kinesin-like protein	-1.30	9.54	-4.54	9.07E-03	1.33
CK294172	No similarity found	-1.27	10.54	-6.14	2.33E-04	5.52
C09-j001-plate-123	No similarity found	-1.13	8.94	-4.51	9.07E-03	1.30
CK298901	glutamate decarboxylase isozyme 1	-0.96	10.58	-6.61	5.80E-05	7.42
CK285623	parC	-0.96	9.83	-4.53	7.51E-03	1.52
G06-a001-plate-14	GHS1 (GLUCOSE HYPERSENSITIVE 1); structural constituent of ribosome	-0.94	9.89	-7.06	3.19E-05	8.29
B07-j001-plate-140	acetyl-CoA carboxylase beta subunit	-0.93	10.79	-6.66	5.49E-05	7.56
CK292211	unknown protein	-0.89	12.15	-7.62	5.78E-06	10.21
CK286217	unknown protein	-0.87	10.98	-8.07	1.96E-06	11.40
CK295879	unknown protein	-0.87	12.10	-10.52	9.13E-09	17.40
D08-j001-plate-121	At2g43020/MFL8.12	-0.86	10.03	-5.26	1.47E-03	3.45
E01-j001-plate-103	SUVH6	-0.84	12.68	-6.56	5.80E-05	7.30
CK284119	transporter	-0.77	11.88	-6.17	1.28E-04	6.18
CK287604	protein binding	-0.71	11.72	-7.55	6.17E-06	10.02
CK289227	pepsin A	-0.69	10.26	-6.27	1.04E-04	6.48
B09-j001-plate-90	No similarity found	-0.68	11.37	-5.46	7.01E-04	4.17
D12-j001-plate-6	No similarity found	-0.63	10.35	-4.58	6.63E-03	1.70
CK289910	No similarity found	-0.60	10.48	-4.59	6.60E-03	1.72

Continua...

Tabella 5.3: Geni differenzialmente espressi a 5 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
A01-j001-plate-58	No similarity found	-0.56	13.22	-10.89	9.13E-09	17.88
CK293802	amino acid binding / oxidoreductase, acting on the CH-OH group of donors, NAD or NADP as acceptor / phosphoglycerate dehydrogenase	-0.54	10.13	-5.15	1.81E-03	3.18
CK283682	dicyamin	-0.53	11.29	-5.59	4.99E-04	4.55
B01-j001-plate-70	No similarity found	-0.51	12.15	-6.78	4.62E-05	7.90
B11-j001-plate-126	No similarity found	-0.24	12.68	-4.43	9.07E-03	1.24
B12-j001-plate-163	unnamed protein product	-0.23	13.71	-4.55	7.01E-03	1.61
CK283839	No similarity found	0.31	13.83	5.91	2.33E-04	5.45
H11-a001-plate-90	BChain B, Crystal Structure Of Annexin 24(Ca32) From Capsicum Annuum	0.38	11.02	4.68	5.39E-03	1.98
G12-j001-plate-46	No similarity found	0.40	13.00	6.57	5.80E-05	7.32
G04-j001-plate-200	No similarity found	0.41	13.33	5.80	3.19E-04	5.05
CK292611	unknown protein	0.44	13.05	6.05	1.65E-04	5.85
A12-j001-plate-68	No similarity found	0.49	12.80	10.22	1.26E-08	16.73
CK290771	protein phosphatase 2C	0.50	10.75	5.07	1.93E-03	3.07
CK292597	secondary cell wall-related glycosyltransferase family 47	0.55	10.38	4.49	7.78E-03	1.45
CK287347	EMB1796	0.60	11.54	5.14	1.68E-03	3.28
C03-j001-plate-213	Retrotransposon gag protein, putative	0.67	10.78	5.12	1.77E-03	3.20
E08-j001-plate-6	No similarity found	0.75	9.37	4.55	7.51E-03	1.52
B06-j001-plate-91	P0592G05.9	0.76	11.01	5.88	2.68E-04	5.29
Continua...						

Tabella 5.3: Geni differenzialmente espressi a 5 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
G03-j001-plate-86	RL23A-TOBAC60S ribosomal protein L23a (L25)	0.77	11.68	4.57	6.77E-03	1.66
E03-j001-plate-150	No similarity found	0.78	10.78	6.15	1.28E-04	6.14
A06-a001-plate-118	No similarity found	0.85	11.12	6.20	1.19E-04	6.29
CK284067	hydroxyphenylpyruvate reductase (HPPR)	0.85	11.13	6.31	1.04E-04	6.48
B05-j001-plate-49	PHS1 (PROPYLAMIDE-HYPERSENSITIVE 1); phosphoprotein phosphatase/ protein tyrosine/serine/threonine phosphatase	0.94	11.11	6.34	9.28E-05	6.67
CK297882	glucose-inhibited division A family protein	0.97	9.12	4.55	9.53E-03	1.28
H03-j001-plate-49	No similarity found	0.98	9.10	4.78	5.96E-03	1.93
CK292295	hypothetical protein	1.02	8.96	4.96	2.56E-03	2.76
A01-j001-plate-17	No similarity found	1.03	9.95	4.50	8.80E-03	1.34
F12-j001-plate-138	No similarity found	1.14	9.44	5.73	3.43E-04	4.94
CK285839	PnFL-2	1.16	9.34	5.87	3.02E-04	5.15
CK297283	ATP:ADP antiporter/ binding	1.22	9.61	6.64	6.80E-05	7.11
C07-j001-plate-203	No similarity found	1.23	9.04	5.73	5.64E-04	4.46
CK294695	complement component 1, r subcomponent variant	1.26	10.14	4.71	5.39E-03	1.99
C07-j001-plate-214	No similarity found	1.26	10.41	9.67	5.30E-08	15.14
CK289457	DNA binding / transcription factor	1.28	8.87	4.66	5.92E-03	1.88
CK286035	G subunit of Vacuolar-type H ⁺ -ATPase	1.30	9.33	6.49	9.28E-05	6.71
H05-j001-plate-205	dihydrofolate reductase-thymidylate synthetase precursor	1.54	8.40	4.58	9.98E-03	1.22
CK296736	No similarity found	4.30	7.96	8.95	1.96E-06	11.22

Continua...

Tabella 5.3: Geni differenzialmente espressi a 5 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
A06-a001-plate-117	hypoxia-responsive family protein	5.82	11.80	6.76	5.22E-05	7.70

^a Codice identificativo interno del laboratorio

^b Proteine con la più alta omologia. Ricavato da BLASTx

^c M: valore medio dei $\log_2(T/C)$, dove T è il segnale della pianta trattata e C del controllo. Indica la variazione di espressione del gene dovuta all'infezione virale.

^d A: valore medio del $\log_2(T*C)$. Indica per ogni gene il valore medio di espressione.

^e Valore del test t di Student.

^f Probabilità associata al valore t

^g È il *log-odds* che il gene sia differenzialmente espresso. Un valore di B pari a zero indica una probabilità del 50% di differenza di espressione. Per valori positivi la probabilità aumenta (per B=1.5 la probabilità sale al 82%), per negativi diminuisce

Tabella 5.4: Geni differenzialmente espressi a 5.5 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
CK284995	oligouridylate binding protein	-1.90	9.10	-6.82	5.04E-05	7.60
F07-j001-plate-126	hydrolase, acting on glycosyl bonds / hydrolase, hydrolyzing O-glycosyl compounds	-1.80	7.98	-4.61	6.48E-03	1.18
E08-j001-plate-215	chloroplast 30S ribosomal protein	-1.70	8.75	-5.87	4.61E-04	4.73
B04-j001-plate-151	No similarity found	-1.63	9.12	-4.66	4.90E-03	1.56
D03-j001-plate-169	No similarity found	-1.48	9.20	-6.89	3.02E-05	8.21
G05-j001-plate-71	No similarity found	-1.36	8.64	-4.55	5.88E-03	1.29
CK295545	unknown protein	-1.32	9.26	-6.78	5.04E-05	7.61
B04-j001-plate-6	No similarity found	-1.29	9.47	-6.99	2.49E-05	8.49
CK297104	Pathogenesis-related transcriptional factor and ERF	-1.23	9.08	-5.08	1.89E-03	2.93
CK290735	T23J18.3	-1.13	9.02	-5.02	2.04E-03	2.76
CK294709	hypersensitive-induced reaction protein	-1.08	9.36	-4.21	9.66E-03	0.57
A08-j001-plate-89	No similarity found	-1.06	10.42	-7.84	2.71E-06	10.80
D01-j001-plate-124	peroxisomal acyl-CoA oxidase 1A	-1.05	8.34	-4.52	5.26E-03	1.39
C09-a001-plate-179	electron transporter / thiol-disulfide exchange intermediate	-1.04	13.71	-12.15	3.84E-10	20.94
B07-j001-plate-204	unknown protein	-0.99	9.41	-4.83	2.70E-03	2.27
H02-j001-plate-71	putative AAA-metalloprotease FtsH (fragment)	-0.96	9.30	-4.78	2.70E-03	2.24
CK290074	No similarity found	-0.92	10.44	-6.64	6.24E-05	7.22
H03-j001-plate-152	unknown protein	-0.91	12.65	-5.42	6.82E-04	4.06
CK290484	unknown protein	-0.90	13.44	-9.40	1.05E-07	14.78

Continua...

Tabella 5.4: Geni differenzialmente espressi a 5.5 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
CK292211	unknown protein	-0.90	12.15	-8.60	5.40E-07	12.77
H07-j001-plate-5	unknown protein	-0.90	11.85	-5.42	6.82E-04	4.06
B03-j001-plate-39	No similarity found	-0.87	9.49	-4.85	2.41E-03	2.44
CK287412	Putative peptide chain release factor subunit 1 protein	-0.87	13.13	-6.19	1.31E-04	6.25
CK298558	unknown protein	-0.83	9.34	-4.28	7.14E-03	0.85
CK298120	putative RING protein	-0.80	9.63	-4.69	3.49E-03	1.90
CK284203	ACC oxidase	-0.79	10.33	-4.40	5.88E-03	1.18
H08-a001-plate-36	No similarity found	-0.78	9.71	-4.17	9.59E-03	0.54
CK283697	No similarity found	-0.77	10.00	-4.64	3.52E-03	1.86
CK295879	unknown protein	-0.76	12.10	-10.34	1.43E-08	17.01
CK283808	transferase	-0.76	10.23	-5.39	7.11E-04	3.97
A04-j001-plate-126	60S ribosomal protein L19	-0.75	9.76	-4.36	6.27E-03	1.07
H08-j001-plate-171	hypothetical protein LOC-Os11g14070	-0.74	9.23	-4.67	3.36E-03	1.93
C12-j001-plate-143	unknown protein	-0.73	11.36	-4.96	2.04E-03	2.76
C08-j001-plate-68	EMB2745	-0.72	8.79	-4.72	3.28E-03	1.99
A01-j001-plate-85	Hypothetical protein	-0.71	9.87	-4.42	6.21E-03	1.18
CK284119	transporter	-0.71	11.88	-6.38	8.71E-05	6.79
CK297822	No similarity found	-0.70	12.36	-5.98	1.95E-04	5.67
B09-a001-plate-132	CYP1	-0.69	12.07	-4.62	3.68E-03	1.81
C08-j001-plate-71	RNase L inhibitor	-0.67	9.17	-4.41	7.88E-03	0.87
CK287506	U-box domain-containing protein	-0.66	14.18	-4.55	4.35E-03	1.60
Continua...						

Tabella 5.4: Geni differenzialmente espressi a 5.5 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
CK289500	scarecrow transcription factor family	-0.65	10.83	-4.90	2.19E-03	2.60
CK287471	chitinase precursor (AA -23 to 306)	-0.65	9.84	-5.24	1.02E-03	3.56
CK289648	chloroplastic outer envelope membrane protein (OEP75)	-0.62	10.88	-6.04	1.77E-04	5.83
CK299019	4CL1-TOBAC4-coumarate-CoA ligase 1 (4CL 1) (4-coumaroyl-CoA synthase 1)	-0.62	9.44	-4.14	9.66E-03	0.48
CK289940	EST AU070346(S12172) corresponds to a region of the predicted gene. Similar to AMP-binding protein. (X94625)	-0.60	12.54	-4.43	5.60E-03	1.28
H06-j001-plate-212	At3g19650	-0.58	10.15	-4.74	2.97E-03	2.13
D03-j001-plate-203	No similarity found	-0.57	12.48	-5.95	2.05E-04	5.57
CK297973	No similarity found	-0.56	11.22	-4.64	3.71E-03	1.81
CK289096	unknown	-0.55	10.44	-4.58	4.24E-03	1.66
CK284377	unknown protein	-0.55	10.59	-4.17	9.59E-03	0.53
H03-j001-plate-213	GHMP kinase-like protein	-0.54	11.17	-5.22	1.08E-03	3.49
CK287604	protein binding	-0.54	11.72	-6.42	8.25E-05	6.90
CK292971	serine acetyltransferase	-0.54	10.18	-5.38	7.11E-04	3.96
CK287558	Integrase, catalytic region; Peptidase aspartic, catalytic	-0.52	10.83	-4.39	5.96E-03	1.16
CK294404	E-class P450, group I	-0.52	11.77	-4.90	2.19E-03	2.59
H12-j001-plate-213	No similarity found	-0.46	11.31	-4.70	3.27E-03	2.01
E08-j001-plate-61	Pto-like serine/threonine kinase	-0.45	10.75	-4.32	6.48E-03	0.98
CK283682	dicyanin	-0.41	11.29	-4.83	2.50E-03	2.37
CK292956	EXO70 gene family	-0.40	11.02	-4.14	9.68E-03	0.47

Continua...

Tabella 5.4: Geni differenzialmente espressi a 5.5 giorni

ID ^a	Funzione probabile ^b	M ^c	A ^d	t ^e	P ^f	B ^g
CK292627	putative phosphoribosylanthranilate transferase, 3' partial; 131493-134402	-0.31	14.39	-5.16	1.23E-03	3.33
CK294393	AT5g53330/K19E1-13	-0.26	13.61	-4.33	6.48E-03	0.99
G12-j001-plate-46	No similarity found	0.24	13.00	4.42	5.71E-03	1.23
A06-a001-plate-142	At2g04280	0.30	12.15	4.30	6.81E-03	0.91
H11-j001-plate-5	No similarity found	0.31	11.48	4.68	3.28E-03	1.98
D09-j001-plate-213	hydrolase, hydrolyzing O-glycosyl compounds	0.31	12.80	6.08	1.75E-04	5.92
CK290609	DEAD/DEAH box helicase, N-terminal	0.47	11.31	4.80	2.64E-03	2.30
yeast3	yeast3	0.49	10.46	4.34	6.40E-03	1.03
CK290539	No similarity found	0.56	9.83	4.85	2.50E-03	2.38
CK289313	No similarity found	0.57	9.59	4.14	9.66E-03	0.48
D09-j001-plate-126	No similarity found	0.63	9.59	4.47	5.22E-03	1.37
CK287659	NAD-dependent isocitrate dehydrogenase	0.67	10.87	4.97	2.04E-03	2.72
C05-j001-plate-20	No similarity found	0.69	9.84	5.42	6.82E-04	4.06
E08-j001-plate-124	No similarity found	0.73	11.12	4.69	3.28E-03	1.98
D06-j001-plate-134	nicotinamide-nucleotide adenylyltransferase/ nucleotidyl- transferase	0.73	10.06	4.87	2.41E-03	2.45
CK286479	prephenate dehydratase	0.79	9.65	4.29	9.59E-03	0.63
CK295409	biotic cell death-associated protein	0.82	10.97	4.54	4.38E-03	1.58
CK296893	putative protein	0.84	10.65	4.21	8.51E-03	0.65
CK297807	No similarity found	0.86	9.46	4.63	3.71E-03	1.78
Continua...						

Tabella 5.4: Geni differenzialmente espressi a 5.5 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
G05-a001-plate-74	'putative endo-1,3;1,4-beta-D-glucanase'	0.90	10.95	7.17	1.64E-05	8.99
B04-j001-plate-122	putative beta-glucuronidase precursor	0.98	9.35	4.27	8.23E-03	0.77
G08-j001-plate-58	conserved hypothetical protein	0.99	9.41	4.27	7.88E-03	0.76
H08-j001-plate-134	unknown protein	1.02	9.43	4.37	6.28E-03	1.08
E10-j001-plate-86	triacylglycerol lipase	1.02	8.67	4.53	5.69E-03	1.35
A02-j001-plate-197	No similarity found	1.07	9.67	4.47	5.22E-03	1.38
B10-j001-plate-122	No similarity found	1.09	9.50	6.30	1.12E-04	6.46
F05-j001-plate-8	expressed protein	1.10	9.42	4.43	6.13E-03	1.15
C02-j001-plate-155	No similarity found	1.11	9.41	5.46	6.74E-04	4.18
G03-a001-plate-179	Harpin-induced 1	1.12	9.24	6.06	1.77E-04	5.86
CK285339	No similarity found	1.14	9.45	4.78	2.98E-03	2.13
D07-j001-plate-71	No similarity found	1.15	9.61	6.11	1.95E-04	5.68
F09-j001-plate-129	structural constituent of ribosome	1.19	9.26	4.94	2.20E-03	2.59
C10-j001-plate-35	UPL6; ubiquitin-protein ligase	1.22	9.06	4.80	2.70E-03	2.25
G09-j001-plate-214	No similarity found	1.22	7.98	4.52	6.22E-03	1.20
C06-a001-plate-53	unknown protein	1.26	8.74	5.35	1.24E-03	3.37
C03-j001-plate-201	unknown protein	1.28	9.60	6.27	1.12E-04	6.47
F01-j001-plate-82	No similarity found	1.31	8.72	4.22	9.75E-03	0.50
A11-j001-plate-91	unknown protein	1.34	9.54	6.00	2.67E-04	5.29
F09-j001-plate-150	putative gag/pol polyprotein, 3'-partial	1.35	8.84	5.45	8.29E-04	3.83
D09-j001-plate-142	No similarity found	1.38	9.17	5.70	6.82E-04	4.16

Continua...

Tabella 5.4: Geni differenzialmente espressi a 5.5 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
F05-j001-plate-171	similar to CDK5 regulatory subunit-associated protein	1.45	9.37	5.62	5.02E-04	4.56
G04-j001-plate-202	No similarity found	1.49	8.76	4.34	6.77E-03	0.95
A10-j001-plate-130	putative polyprotein	1.50	9.31	4.48	6.80E-03	1.04
B01-j001-plate-206	No similarity found	1.52	9.83	4.42	5.71E-03	1.24
A04-j001-plate-92	No similarity found	1.63	8.58	5.27	1.68E-03	3.08
E08-j001-plate-162	No similarity found	1.69	9.52	5.38	7.11E-04	3.95
E02-j001-plate-58	No similarity found	1.72	10.96	4.57	6.48E-03	1.18
CK289275	electron transporter	1.83	9.88	8.67	5.40E-07	12.96
CK295619	protein kinase-like protein	1.91	9.01	8.02	2.71E-06	10.86
CK292327	BONZAI1	1.97	8.92	7.16	5.27E-05	7.50
G03-j001-plate-33	LSTK-1-like kinase	1.99	8.21	5.83	5.54E-04	4.46
G09-j001-plate-162	No similarity found	2.02	8.84	4.91	2.23E-03	2.55
C05-j001-plate-141	No similarity found	2.04	8.58	6.23	2.35E-04	5.47
CK294482	unknown protein	2.07	8.15	5.43	2.19E-03	2.76
CK284185	FRO1-like protein	2.24	8.74	6.01	3.31E-04	5.08
B07-j001-plate-212	unknown protein	2.78	8.26	5.35	2.04E-03	2.88
H12-j001-plate-92	unknown protein	3.10	8.35	5.74	6.01E-04	4.39
A06-a001-plate-117	hypoxia-responsive family protein	6.51	11.80	8.45	8.60E-07	12.17

^a Codice identificativo interno del laboratorio

^b Proteine con la più alta omologia. Ricavato da BLASTx

^c M: valore medio dei $\log_2(T/C)$, dove T è il segnale della pianta trattata e C del controllo. Indica la variazione di espressione del gene dovuta all'infezione virale.

^d A: valore medio del $\log_2(T^*C)$. Indica per ogni gene il valore medio di espressione.

^e Valore del test t di Student.

^f Probabilità associata al valore t

^g È il *log-odds* che il gene sia differenzialmente espresso. Un valore di B pari a zero indica una probabilità del 50% di differenza di espressione. Per valori positivi la probabilità aumenta (per $B=1.5$ la probabilità sale al 82%), per negativi diminuisce

Tabella 5.5: Geni differenzialmente espressi a 6 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
F07-j001-plate-151	putative tubby protein	-2.98	11.80	-15.34	8.91E-13	27.14
CK295177	cytosolic 6-phosphogluconate dehydrogenase	-1.99	12.42	-11.91	1.44E-10	20.52
C11-a001-plate-20	DnaJ-like protein	-1.99	10.07	-10.82	8.39E-10	18.10
CK288026	nucleic acid binding protein	-1.95	11.25	-14.14	3.05E-12	24.97
G09-j001-plate-53	No similarity found	-1.88	8.08	-4.24	5.63E-03	0.07
D03-j001-plate-169	No similarity found	-1.78	9.20	-7.94	4.09E-07	11.06
B01-j001-plate-213	No similarity found	-1.59	8.32	-4.90	2.75E-03	1.28
C09-j001-plate-169	Ribosomal protein L29	-1.58	10.61	-10.64	1.20E-09	17.69
D10-j001-plate-39	Membrane attack	-1.54	10.57	-10.86	6.11E-10	18.59
	component/perforin/complement C9					
CK296258	pyruvate decarboxylase	-1.45	8.62	-4.27	4.06E-03	0.33
A12-j001-plate-65	No similarity found	-1.39	10.74	-9.72	8.20E-09	15.48
CK288290	respiratory burst oxidase homolog	-1.37	11.14	-12.62	3.58E-11	21.99
E08-j001-plate-61	Pto-like serine/threonine kinase	-1.36	10.75	-12.48	3.58E-11	22.28
CK292879	G3PA-TOBACGlyceraldehyde-3-phosphate dehydrogenase	-1.36	12.98	-12.92	1.71E-11	23.23
	A, chloroplast precursor (NADP-dependent glyceraldehy-					
	dephosphate dehydrogenase subunit A)					
CK293798	putative stress inducible protein	-1.33	13.48	-4.23	3.83E-03	0.37
CK285109	calcium-dependent protein kinase CPK4	-1.33	8.19	-5.14	8.41E-04	2.34
A07-a001-plate-38	putative anthranilate N-benzoyltransferase	-1.33	9.59	-6.31	3.37E-05	6.14
	Continua...					

Tabella 5.5: Geni differenzialmente espressi a 6 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
D11-j001-plate-214	No similarity found	-1.29	11.30	-10.42	2.49E-09	16.87
CK287604	protein binding	-1.28	11.72	-14.60	1.02E-12	26.63
CK284119	transporter	-1.27	11.88	-10.94	5.64E-10	18.79
B12-j001-plate-206	No similarity found	-1.27	10.39	-5.85	8.08E-05	5.07
CK298323	Zinc finger, RING-type	-1.25	10.61	-6.50	1.62E-05	6.93
CK293908	Zinc finger, DHHC-type	-1.22	13.86	-4.91	7.42E-04	2.30
CK291424	unknown	-1.22	11.06	-5.85	8.04E-05	5.08
CK283618	unknown	-1.21	10.21	-5.71	1.02E-04	4.64
D03-j001-plate-151	No similarity found	-1.20	10.71	-7.98	3.83E-07	11.15
C06-j001-plate-64	No similarity found	-1.20	12.46	-9.97	4.74E-09	16.12
D02-j001-plate-84	No similarity found	-1.19	12.33	-4.47	2.30E-03	1.03
B03-j001-plate-86	unknown protein	-1.17	11.24	-9.77	5.49E-09	15.91
D05-j001-plate-169	No similarity found	-1.15	10.29	-10.19	2.30E-09	16.97
B04-j001-plate-51	DNA-directed RNA polymerase II largest chain	-1.14	9.85	-3.99	7.31E-03	-0.27
H06-j001-plate-139	No similarity found	-1.14	8.84	-3.89	8.45E-03	-0.60
CK294774	PEX14 (PEROXISOME DEFECTIVE 2)	-1.13	10.99	-6.98	4.31E-06	8.35
CK295836	CONSTANS 1	-1.13	11.91	-9.59	1.09E-08	15.17
CK293784	S31196hypothetical protein - potato	-1.12	10.16	-8.76	6.58E-08	13.05
A05-a001-plate-52	ribosomal protein L29 protein-like	-1.12	10.21	-6.01	5.62E-05	5.55
CK287763	unknown protein	-1.10	11.73	-8.15	2.41E-07	11.64
G12-j001-plate-162	phospholipid transfer protein 1	-1.10	8.60	-4.07	6.30E-03	-0.19

Continua...

Tabella 5.5: Geni differenzialmente espressi a 6 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
CK290484	unknown protein	-1.09	13.44	-10.92	5.64E-10	18.74
CK287854	putative DNA/RNA binding protein	-1.08	10.25	-8.43	1.51E-07	12.17
D08-j001-plate-213	No similarity found	-1.07	11.10	-10.74	7.62E-10	18.30
CK289920	carbohydrate transporter/ sugar porter	-1.07	10.46	-7.85	5.17E-07	10.81
CK296918	similar to Arabidopsis thaliana DNA-directed RNA polymerase (EC 2.7.7.6) II largest chain (JDMU1)	-1.03	11.10	-6.90	6.23E-06	8.04
C09-a001-plate-179	electron transporter/ thiol-disulfide exchange intermediate	-1.03	13.71	-11.58	1.78E-10	20.27
B07-j001-plate-50	catalytic/ trehalose-phosphatase	-1.02	10.98	-7.59	9.36E-07	10.09
H12-j001-plate-35	No similarity found	-1.01	12.99	-4.44	2.42E-03	0.96
CK297091	unknown protein	-1.01	11.77	-7.74	6.51E-07	10.48
C03-j001-plate-200	No similarity found	-1.00	10.19	-5.87	7.33E-05	5.12
F03-a001-plate-190	ceramidase	-0.99	12.57	-5.57	1.43E-04	4.23
B09-j001-plate-90	No similarity found	-0.98	11.37	-8.44	1.17E-07	12.44
E05-j001-plate-57	No similarity found	-0.98	9.57	-4.29	3.90E-03	0.41
CK293251	PPR2, putative, expressed	-0.96	10.00	-4.57	1.76E-03	1.33
B08-j001-plate-103	unknown protein	-0.96	13.58	-4.01	6.30E-03	-0.25
F04-j001-plate-49	No similarity found	-0.95	10.09	-4.24	3.83E-03	0.38
G03-j001-plate-158	No similarity found	-0.94	11.82	-4.79	9.88E-04	1.95
B12-j001-plate-104	No similarity found	-0.93	9.73	-5.30	3.04E-04	3.37
CK287617	unnamed protein product	-0.90	11.41	-9.13	2.23E-08	14.26
A04-a001-plate-4	putative UDP-glucuronate decarboxylase 3	-0.90	9.38	-3.97	7.69E-03	-0.34
Continua...						

Tabella 5.5: Geni differenzialmente espressi a 6 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
G11-j001-plate-206	Hs1pro-1	-0.90	9.44	-6.14	3.69E-05	5.89
CK293240	Cyclin-like F-box	-0.89	9.32	-4.27	4.13E-03	0.34
CK291849	PDF1 (PROTODERMAL FACTOR 1)	-0.89	9.17	-4.05	5.76E-03	-0.14
C01-j001-plate-35	No similarity found	-0.89	12.78	-8.94	3.37E-08	13.75
CK296675	unknown protein	-0.89	11.50	-9.66	1.17E-08	15.04
CK293855	unknown protein	-0.89	13.34	-6.84	6.23E-06	7.93
CK288895	No similarity found	-0.88	11.25	-3.85	8.76E-03	-0.70
A08-j001-plate-89	No similarity found	-0.85	10.42	-6.07	4.41E-05	5.71
CK290794	hypothetical protein	-0.85	12.00	-4.37	2.84E-03	0.76
CK289493	putative SCARECROW protein	-0.84	10.79	-6.39	1.90E-05	6.62
G06-j001-plate-84	No similarity found	-0.83	11.60	-5.71	1.02E-04	4.64
G06-j001-plate-185	No similarity found	-0.81	11.61	-6.99	4.27E-06	8.38
CK292971	serine acetyltransferase	-0.80	10.18	-7.73	6.51E-07	10.47
CK287412	Putative peptide chain release factor subunit 1 protein	-0.79	13.13	-5.43	2.03E-04	3.82
CK287281	protein phosphatase 2C	-0.79	9.87	-4.55	1.96E-03	1.22
CK292211	unknown protein	-0.78	12.15	-7.16	2.90E-06	8.86
CK290539	No similarity found	-0.78	9.83	-5.38	2.48E-04	3.72
C11-j001-plate-67	conserved hypothetical protein	-0.78	9.02	-4.37	2.91E-03	0.73
G06-a001-plate-14	GHS1 (GLUCOSE HYPERSENSITIVE 1); structural constituent of ribosome	-0.76	9.89	-7.26	2.90E-06	8.83
C03-j001-plate-212	No similarity found	-0.76	10.48	-5.68	1.18E-04	4.47

Continua...

Tabella 5.5: Geni differenzialmente espressi a 6 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
CK298162	N-rich protein	-0.75	10.68	-5.18	3.83E-04	3.08
H07-j001-plate-5	unknown protein	-0.75	11.85	-4.31	3.21E-03	0.60
C06-j001-plate-215	ubiquitin-protein ligase/ zinc ion binding	-0.74	10.74	-4.92	7.18E-04	2.34
C12-j001-plate-143	unknown protein	-0.74	11.36	-4.87	8.02E-04	2.19
CK283481	No similarity found	-0.73	12.11	-11.02	5.29E-10	18.99
C08-j001-plate-175	putative p-coumaroyl 3'-hydroxylase CYP98A-C1	-0.72	10.57	-7.14	2.90E-06	8.80
CK295879	unknown protein	-0.71	12.10	-9.21	1.94E-08	14.46
CK294106	Helix-loop-helix DNA-binding	-0.70	10.33	-4.46	2.33E-03	1.01
F03-a001-plate-8	histone H2A	-0.70	10.37	-5.62	1.25E-04	4.38
CK288968	hsr201	-0.69	10.24	-5.03	5.48E-04	2.66
CK294121	unknown	-0.69	10.43	-5.66	1.16E-04	4.49
CK290367	Alpha/beta hydrolase fold	-0.67	9.72	-6.06	5.09E-05	5.55
CK292997	No similarity found	-0.67	11.92	-7.31	2.00E-06	9.30
H06-j001-plate-57	coiled-coil-helix-coiled-coil-helix domain containing protein	-0.67	11.40	-3.92	7.62E-03	-0.51
G03-j001-plate-86	RL23A-TOBAC60S ribosomal protein L23a (L25)	-0.67	11.68	-4.28	3.43E-03	0.51
CK283425	No similarity found	-0.67	9.55	-4.76	1.13E-03	1.82
CK294656	hydrolase, hydrolyzing O-glycosyl compounds	-0.66	9.85	-4.30	3.42E-03	0.53
CK284932	EFTUB-NICSYElongation factor TuB, chloroplast precursor (EF-TuB)	-0.66	9.92	-4.31	3.28E-03	0.57
CK289463	No similarity found	-0.66	11.44	-5.85	7.59E-05	5.06
CK291790	HMG-CoA synthase	-0.65	10.45	-3.84	9.13E-03	-0.75
Continua...						

Tabella 5.5: Geni differenzialmente espressi a 6 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
CK290196	catalytic/ hydrolase	-0.65	14.25	-6.63	1.06E-05	7.33
CK290771	protein phosphatase 2C	-0.64	10.75	-6.97	4.44E-06	8.30
G01-j001-plate-215	No similarity found	-0.63	11.24	-5.65	1.18E-04	4.46
B09-a001-plate-132	CYP1	-0.63	12.07	-4.06	5.65E-03	-0.12
C10-j001-plate-169	No similarity found	-0.63	13.06	-6.92	6.66E-06	7.96
F11-a001-plate-70	cystathionine gamma synthase	-0.63	10.23	-4.15	4.62E-03	0.13
D12-j001-plate-202	unknown protein	-0.62	12.78	-5.37	2.38E-04	3.64
B03-j001-plate-49	No similarity found	-0.62	13.88	-5.56	1.55E-04	4.23
B10-a001-plate-65	Glutaredoxin	-0.62	10.99	-7.14	2.90E-06	8.80
CK284181	ribosomal Pr 117	-0.62	10.20	-5.73	9.93E-05	4.70
F01-j001-plate-103	EF-1-alpha-related GTP-binding protein	-0.61	12.30	-7.85	6.19E-07	10.66
CK291350	GP1-LYCESPolygalacturonase-1 non-catalytic beta sub-unit precursor (AroGPI) (Polygalacturonase converter) (PG converter)	-0.61	10.54	-5.40	2.20E-04	3.73
E05-j001-plate-6	DGL1; dolichyl-diphosphooligosaccharide-protein glycosyltransferase	-0.61	10.34	-6.42	1.76E-05	6.71
CK293647	perchloric acid soluble translation inhibitor protein homolog	-0.60	11.26	-4.24	3.89E-03	0.51
CK292851	sulfate transporter	-0.60	13.51	-11.31	2.96E-10	19.66
D12-j001-plate-166	unknown protein	-0.60	10.16	-4.15	4.62E-03	0.14
D09-j001-plate-129	transcription factor	-0.60	14.13	-5.66	1.16E-04	4.48
Continua...						

Tabella 5.5: Geni differenzialmente espressi a 6 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
H12-j001-plate-213	No similarity found	-0.59	11.31	-5.84	7.68E-05	5.02
CK291576	No similarity found	-0.59	13.59	-9.01	2.96E-08	13.94
A01-j001-plate-104	protein kinase AtSIK	-0.59	13.25	-4.07	5.54E-03	-0.08
G04-j001-plate-133	RB11B-TOBACRas-related protein Rab11B	-0.59	11.64	-5.12	4.46E-04	2.91
A12-j001-plate-82	putative S-receptor kinase (EC 2.7.1.-) homolog 2 precursor	-0.58	11.75	-4.18	4.34E-03	0.21
CK293653	TUB8	-0.57	10.13	-6.51	1.43E-05	6.98
B02-j001-plate-168	No similarity found	-0.57	10.78	-5.29	3.37E-04	3.26
C06-a001-plate-128	No similarity found	-0.57	10.10	-4.72	1.33E-03	1.65
A10-j001-plate-128	No similarity found	-0.57	13.33	-3.95	7.21E-03	-0.42
CK292956	unnamed protein product	-0.56	11.02	-5.53	1.54E-04	4.12
CK295732	homeobox-leucine zipper protein	-0.56	11.59	-5.97	6.14E-05	5.31
D03-j001-plate-203	No similarity found	-0.56	12.48	-5.62	1.25E-04	4.38
G07-a001-plate-70	putative plastid protein	-0.56	10.65	-5.31	2.77E-04	3.46
CK283305	putative beta-amylase PCT-BMYI	-0.55	10.62	-5.10	4.68E-04	2.85
CK298871	putative Citrate synthase, glyoxysomal precursor (GCS)	-0.55	14.07	-5.03	5.48E-04	2.67
CK294404	E-class P450, group I	-0.55	11.77	-4.97	6.45E-04	2.47
G09-j001-plate-8	No similarity found	-0.54	10.60	-5.75	9.45E-05	4.76
CK297822	No similarity found	-0.54	12.36	-4.39	2.77E-03	0.81
C05-j001-plate-204	ATCSLA02; transferase, transferring glycosyl groups	-0.53	13.56	-5.57	1.54E-04	4.26
CK290444	No similarity found	-0.53	10.71	-3.91	7.73E-03	-0.54
F08-j001-plate-2	unknown protein	-0.52	11.80	-5.84	7.68E-05	5.02

Continua...

Tabella 5.5: Geni differenzialmente espressi a 6 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
E10-j001-plate-57	No similarity found	-0.52	12.19	-5.22	3.44E-04	3.19
CK287558	Integrase, catalytic region; Peptidase aspartic, catalytic	-0.51	10.83	-4.15	4.62E-03	0.13
G04-j001-plate-31	No similarity found	-0.50	10.62	-4.10	5.19E-03	-0.01
E01-j001-plate-103	SUVH6	-0.50	12.68	-4.21	3.97E-03	0.31
G02-j001-plate-156	protein carrier	-0.50	10.51	-4.91	7.42E-04	2.30
CK291908	protein binding / ubiquitin-protein ligase/ zinc ion binding	-0.49	11.37	-4.01	6.30E-03	-0.26
E02-j001-plate-64	pectin methylesterase	-0.49	10.69	-3.98	6.71E-03	-0.33
CK295145	No similarity found	-0.49	10.27	-3.93	7.31E-03	-0.46
A06-j001-plate-128	No similarity found	-0.49	13.02	-7.86	6.08E-07	10.70
CK287621	unknown protein	-0.48	12.10	-6.34	2.12E-05	6.49
CK289961	No similarity found	-0.47	10.51	-4.39	2.84E-03	0.76
C01-j001-plate-130	No similarity found	-0.47	12.62	-7.13	2.90E-06	8.78
CK294116	unknown protein	-0.47	11.20	-4.40	2.75E-03	0.83
F03-j001-plate-87	RPB5d	-0.46	11.04	-3.85	8.70E-03	-0.69
H12-j001-plate-206	SANT/MYB domain protein	-0.46	12.04	-6.36	2.03E-05	6.55
B10-j001-plate-206	catalytic	-0.46	13.15	-5.63	1.33E-04	4.44
CK294087	No similarity found	-0.44	10.32	-3.87	8.30E-03	-0.63
G12-j001-plate-1	unknown	-0.44	12.88	-6.47	1.55E-05	6.88
C10-j001-plate-141	No similarity found	-0.44	11.88	-4.26	3.62E-03	0.44
B02-j001-plate-206	No similarity found	-0.42	13.28	-4.76	1.13E-03	1.96
A08-j001-plate-84	Ribosomal L22e protein	-0.42	15.33	-6.43	1.74E-05	6.74

Continua...

Tabella 5.5: Geni differenzialmente espressi a 6 giorni

ID ^a	Funzione probabile ^b	M ^c	A ^d	t ^e	P ^f	B ^g
CK285252	unnamed protein product	-0.42	13.73	-5.53	1.65E-04	4.16
E12-j001-plate-74	No similarity found	-0.41	11.41	-4.07	5.53E-03	-0.08
B11-a001-plate-120	lipid transfer protein	-0.40	12.28	-6.25	2.75E-05	6.22
CK289637	NAD-dependent epimerase/dehydratase; TonB box, N-terminal	-0.39	10.26	-4.39	2.77E-03	0.80
PCV-P15	PCV-P15	-0.39	12.50	-3.92	7.76E-03	-0.54
H10-j001-plate-196	No similarity found	-0.37	12.17	-3.86	8.54E-03	-0.67
CK289958	No similarity found	-0.37	11.54	-3.97	6.87E-03	-0.36
CK284063	NtPRp27	-0.36	10.94	-4.90	7.62E-04	2.26
CK288961	No similarity found	-0.35	12.10	-4.89	7.65E-04	2.25
E09-j001-plate-81	No similarity found	-0.35	14.36	-4.36	3.03E-03	0.82
G01-a001-plate-130	putative nucleosome assembly protein	-0.34	13.14	-4.09	5.28E-03	-0.03
CK292627	putative phosphoribosylanthranilate transferase, 3' partial; 131493-134402	-0.34	14.39	-5.28	2.95E-04	3.39
CK291149	Aspartyl-tRNA synthetase, class IIb; tRNA synthetase, class II (D, K and N)	-0.33	13.91	-5.11	4.51E-04	2.89
F07-j001-plate-81	No similarity found	-0.32	13.38	-4.54	2.00E-03	1.34
CK287096	Strong similarity to gi	-0.31	13.74	-5.55	1.51E-04	4.16
CK287332	No similarity found	-0.30	13.63	-4.19	4.21E-03	0.24
C05-j001-plate-128	Endonuclease/exonuclease/phosphatase	-0.30	12.81	-4.93	7.13E-04	2.36
CK294390	ARG1 (ALTERED RESPONSE TO GRAVITY 1)	-0.29	13.11	-3.92	7.62E-03	-0.51

Continua...

Tabella 5.5: Geni differenzialmente espressi a 6 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
H06-j001-plate-185	hypothetical protein MtrDRAFT-AC136139g13v1	-0.28	11.81	-3.84	8.82E-03	-0.71
D08-j001-plate-150	No similarity found	-0.26	14.27	-4.51	2.17E-03	1.26
B04-j001-plate-115	acyltransferase-like protein	-0.26	12.92	-4.57	1.78E-03	1.32
CK284009	unknown	-0.25	13.42	-4.48	2.25E-03	1.05
A03-a001-plate-132	CYPH-LYCESPeptidyl-prolyl cis-trans isomerase (PPIase)	-0.21	12.41	-3.97	6.96E-03	-0.38
	(Rotamase) (Cyclophilin) (Cyclosporin A-binding protein)					
C11-j001-plate-139	No similarity found	-0.20	13.71	-4.13	4.88E-03	0.07
CK289764	AT5g52920/MXC20-15	-0.20	12.90	-3.98	6.71E-03	-0.33
G11-j001-plate-173	No similarity found	0.24	13.28	4.30	3.28E-03	0.57
B05-j001-plate-148	protein binding / ubiquitin-protein ligase/ zinc ion binding	0.24	13.21	4.12	4.90E-03	0.06
D09-j001-plate-47	No similarity found	0.28	12.75	4.99	5.98E-04	2.55
CK290350	Protein kinase; Concanavalin A-like lectin/glucanase	0.28	13.05	4.81	9.36E-04	2.01
H01-j001-plate-63	No similarity found	0.29	12.65	5.36	2.44E-04	3.61
D07-j001-plate-168	putative nitrate-induced NOI protein	0.29	12.54	4.03	6.11E-03	-0.20
CK292619	protein binding / ubiquitin-protein ligase/ zinc ion binding	0.31	13.78	3.98	6.75E-03	-0.34
B03-j001-plate-130	No similarity found	0.33	12.21	4.31	3.21E-03	0.60
E03-j001-plate-134	No similarity found	0.37	13.14	4.03	6.11E-03	-0.21
C11-j001-plate-103	Integrase, catalytic region; Zinc finger, CCHC-type; Peptidase aspartic, catalytic	0.38	12.75	4.39	2.77E-03	0.80
CK287356	AT-TCP20; transcription factor	0.39	10.79	4.07	5.54E-03	-0.09
CK283387	DNA-binding protein phosphatase 2C	0.39	10.91	4.35	2.92E-03	0.70
Continua...						

Tabella 5.5: Geni differenzialmente espressi a 6 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
CK291025	Glyoxalase/bleomycin resistance protein/dioxygenase	0.40	13.12	4.81	9.36E-04	2.01
CK290178	putative ZF-HD homeobox protein	0.50	11.25	4.82	9.05E-04	2.05
CK286487	No similarity found	0.51	12.58	4.03	6.11E-03	-0.21
G05-a001-plate-74	'putative endo-1,3;1,4-beta-D-glucanase'	0.52	10.95	3.95	7.22E-03	-0.43
A01-j001-plate-58	No similarity found	0.52	13.22	8.93	4.40E-08	13.50
CK298468	No similarity found	0.54	12.77	3.86	8.54E-03	-0.66
F02-j001-plate-122	putative retroelement pol polyprotein-like	0.55	10.10	3.86	9.66E-03	-0.79
C06-j001-plate-175	unnamed protein product	0.57	10.41	4.23	3.83E-03	0.36
CK292423	nam-like protein 3	0.57	11.36	5.12	4.41E-04	2.92
CK289519	chitinase C class I	0.57	10.52	3.84	8.93E-03	-0.73
G01-j001-plate-205	No similarity found	0.58	9.69	3.87	8.71E-03	-0.68
CK290308	multi-protein bridging factor 1	0.58	9.88	4.86	8.75E-04	2.10
C11-j001-plate-198	No similarity found	0.61	9.76	4.49	2.21E-03	1.08
C08-j001-plate-132	unknown protein	0.61	10.72	4.46	2.33E-03	1.00
A03-j001-plate-52	putative ubiquitin-conjugating enzyme	0.61	10.01	4.90	7.62E-04	2.26
G03-j001-plate-151	No similarity found	0.61	9.64	3.95	7.15E-03	-0.41
C07-a001-plate-118	Histone-fold/TFIID-TAF/NF-Y	0.65	10.35	4.85	8.52E-04	2.12
D07-j001-plate-192	No similarity found	0.65	9.51	3.82	9.28E-03	-0.77
CK291270	hypothetical protein	0.66	9.37	3.91	7.63E-03	-0.52
CK290465	transferase, transferring glycosyl groups / transferase, transferring hexosyl groups	0.67	13.20	8.95	3.32E-08	13.80

Continua...

Tabella 5.5: Geni differenzialmente espressi a 6 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
E01-j001-plate-148	conserved hypothetical protein	0.68	10.51	5.07	5.01E-04	2.77
B03-j001-plate-214	splicing factor-like protein	0.68	11.12	3.94	7.27E-03	-0.44
H08-j001-plate-47	No similarity found	0.69	9.90	4.65	1.42E-03	1.56
D06-j001-plate-152	No similarity found	0.70	10.00	4.08	5.58E-03	-0.09
F12-j001-plate-154	No similarity found	0.71	10.48	5.27	3.07E-04	3.34
H10-j001-plate-111	No similarity found	0.71	9.26	4.29	3.62E-03	0.46
G03-j001-plate-5	unknown protein	0.71	12.76	4.23	3.83E-03	0.37
CK291489	cytochrome-c oxidase	0.72	11.35	5.63	1.46E-04	4.23
C12-a001-plate-72	small subunit ribulose 1,5-bisphosphate carboxylase	0.73	10.08	5.24	3.32E-04	3.25
D03-j001-plate-155	heat shock protein binding / unfolded protein binding	0.73	10.58	4.61	1.71E-03	1.43
G09-j001-plate-68	catalytic/ hydrolase	0.74	13.03	9.85	4.74E-09	16.12
C02-j001-plate-86	No similarity found	0.75	9.11	4.00	7.49E-03	-0.44
A02-j001-plate-198	cis-prenyltransferase	0.76	10.06	4.73	1.16E-03	1.77
G11-j001-plate-36	No similarity found	0.78	11.14	5.77	9.00E-05	4.82
B01-j001-plate-123	protein binding / ubiquitin-protein ligase/ zinc ion binding	0.80	10.29	5.09	4.75E-04	2.82
E11-j001-plate-2	No similarity found	0.81	9.94	3.88	8.13E-03	-0.60
D07-j001-plate-132	No similarity found	0.83	10.50	7.74	6.51E-07	10.50
D02-j001-plate-151	No similarity found	0.83	9.94	4.48	2.23E-03	1.07
CK284337	No similarity found	0.83	9.39	4.96	7.83E-04	2.26
C04-j001-plate-31	CRR2 (CHLORORESPIRATORY REDUCTION 2)	0.83	9.35	4.27	3.81E-03	0.41
D12-j001-plate-6	No similarity found	0.84	10.35	6.51	1.43E-05	6.97

Continua...

Tabella 5.5: Geni differenzialmente espressi a 6 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
C10-j001-plate-81	short-chain dehydrogenase Tic32	0.85	9.32	4.43	2.83E-03	0.82
E02-j001-plate-8	No similarity found	0.85	10.41	5.21	3.44E-04	3.19
CK287260	ERF3-TOBACEthylene-responsive transcription factor 3 (Ethylene-responsive element-binding factor 3 homolog) (EREBP-5) (NtERF5)	0.85	11.22	6.63	1.06E-05	7.33
E09-j001-plate-123	No similarity found	0.86	10.27	4.36	2.89E-03	0.73
B04-j001-plate-132	DNA binding / protein binding / zinc ion binding	0.87	9.90	3.85	8.67E-03	-0.68
CK293808	phi-1	0.89	9.66	4.17	4.62E-03	0.30
CK286714	phytochrome kinase substrate 1-like protein, putative	0.89	9.75	4.41	2.82E-03	0.80
D05-j001-plate-150	unknown protein	0.90	10.21	3.81	1.04E-02	-0.69
B03-j001-plate-65	No similarity found	0.92	10.97	7.16	2.90E-06	8.85
CK296070	beta 1,3-glycosyltransferase-like protein I	0.92	8.94	4.10	5.72E-03	-0.11
CK290489	syntaxin-related protein Nt-syr1	0.93	10.12	6.03	4.97E-05	5.58
C01-j001-plate-202	No similarity found	0.94	11.62	4.22	3.90E-03	0.33
CK297104	Pathogenesis-related transcriptional factor and ERF	0.95	9.08	5.00	7.15E-04	2.39
E09-j001-plate-88	No similarity found	0.96	9.98	3.99	7.21E-03	-0.39
CK297140	RK4-TOBAC50S ribosomal protein L4, chloroplast precursor (R-protein L4)	1.00	10.22	3.99	6.87E-03	-0.35
CK285339	No similarity found	1.01	9.45	3.90	8.24E-03	-0.57
E08-j001-plate-6	No similarity found	1.02	9.37	6.64	1.34E-05	7.11
F12-j001-plate-6	No similarity found	1.03	9.13	5.74	9.63E-05	4.74

Continua...

Tabella 5.5: Geni differenzialmente espressi a 6 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
G10-j001-plate-64	No similarity found	1.03	11.51	10.34	1.68E-09	17.33
A12-j001-plate-67	putative peroxisomal membrane protein(22-kDa)(PMP22)	1.03	9.87	5.56	2.25E-04	3.77
E03-j001-plate-154	CAX	1.04	10.86	6.62	1.40E-05	7.05
D02-j001-plate-58	CG6277-PA	1.04	9.62	4.49	2.77E-03	1.00
CK284256	phosphate/phosphoenolpyruvate translocator precursor	1.05	9.48	3.97	7.95E-03	-0.47
D10-j001-plate-121	PREDICTED: similar to MEGF10 protein	1.06	9.49	5.82	1.02E-04	4.69
D01-j001-plate-63	type-A response regulator	1.06	10.86	6.07	5.54E-05	5.47
D06-j001-plate-113	Peptidase M20	1.08	9.75	4.10	5.14E-03	0.01
CK297005	Hypothetical plant protein	1.08	9.72	4.03	6.36E-03	-0.25
G11-j001-plate-92	OSJNBa0074L08.23	1.09	11.49	5.95	5.85E-05	5.36
CK297233	At1g54180	1.11	8.73	3.89	9.66E-03	-0.76
G07-j001-plate-126	No similarity found	1.12	8.92	3.99	7.31E-03	-0.42
D02-j001-plate-128	No similarity found	1.14	9.49	4.23	4.88E-03	0.13
B07-j001-plate-17	hypothetical protein	1.14	9.15	3.96	7.73E-03	-0.50
B01-j001-plate-67	No similarity found	1.16	9.46	4.05	6.21E-03	-0.21
D07-j001-plate-71	No similarity found	1.18	9.61	5.98	7.48E-05	5.13
D05-j001-plate-132	No similarity found	1.19	9.19	4.67	1.53E-03	1.51
C06-j001-plate-89	Putative protein kinase	1.20	11.29	7.43	1.78E-06	9.52
E03-j001-plate-2	No similarity found	1.21	9.42	5.40	2.52E-04	3.71
E11-j001-plate-1	No similarity found	1.22	9.66	5.04	5.48E-04	2.67
C01-j001-plate-125	HMG-I and HMG-Y, DNA-binding	1.24	9.61	3.90	8.48E-03	-0.63

Continua...

Tabella 5.5: Geni differenzialmente espressi a 6 giorni

ID^a	Funzione probabile^b	M^c	A^d	t^e	P^f	B^g
F01-j001-plate-199	DNA-binding protein 1	1.24	9.15	5.23	4.14E-04	3.16
B04-j001-plate-122	putative beta-glucuronidase precursor	1.24	9.35	5.78	1.14E-04	4.56
CK297573	electron transporter/ thiol-disulfide exchange intermediate	1.25	9.48	5.00	5.86E-04	2.57
H07-j001-plate-112	No similarity found	1.25	10.30	4.43	2.77E-03	0.98
C08-j001-plate-162	No similarity found	1.25	10.99	5.27	3.44E-04	3.27
D06-j001-plate-121	transcription factor	1.26	9.77	6.82	7.38E-06	7.74
CK294059	Endonuclease/exonuclease/phosphatase	1.29	11.86	6.70	1.01E-05	7.41
CK286391	No similarity found	1.29	9.99	5.09	5.76E-04	2.63
C12-j001-plate-123	RL34-TOBAC60S ribosomal protein L34	1.30	8.18	4.33	4.16E-03	0.41
E02-j001-plate-111	binding	1.35	9.81	5.88	7.68E-05	5.06
E09-j001-plate-126	No similarity found	1.35	9.74	5.04	5.69E-04	2.63
H03-j001-plate-151	dirigent protein	1.35	8.24	3.87	9.02E-03	-0.52
E05-j001-plate-151	No similarity found	1.35	8.74	4.41	3.56E-03	0.55
CK285532	unknown protein	1.38	9.10	7.42	6.23E-06	8.08
E06-j001-plate-5	No similarity found	1.40	8.79	4.27	4.98E-03	0.16
D06-j001-plate-134	nicotinamide-nucleotide adenylyltransferase/ nucleotidyl-transferase	1.45	10.06	9.28	2.08E-08	14.38
B09-j001-plate-88	unknown protein	1.46	9.77	9.41	1.17E-08	15.00
E02-j001-plate-123	ATP binding / kinase/ protein kinase/ serine/threonine kinase/ protein-tyrosine kinase	1.51	9.95	5.83	7.86E-05	4.98
H07-j001-plate-123	No similarity found	1.52	8.51	4.06	7.22E-03	-0.33

Continua...

Tabella 5.5: Geni differenzialmente espressi a 6 giorni

ID ^a	Funzione probabile ^b	M ^c	A ^d	t ^e	P ^f	B ^g
E03-j001-plate-124	unnamed protein product	1.53	9.82	7.39	2.71E-06	9.01
G12-j001-plate-158	non-ltr retrotransposon reverse transcriptase-like protein	1.57	8.18	4.04	7.27E-03	-0.36
E01-j001-plate-112	putative transposon MuDR mudrA-like protein	1.62	7.92	4.68	2.91E-03	1.00
G08-j001-plate-155	No similarity found	1.72	9.27	4.39	2.83E-03	0.77
C10-j001-plate-6	At5g59250	1.73	8.20	3.85	1.16E-02	-0.75
B01-j001-plate-206	No similarity found	1.98	9.83	5.53	1.54E-04	4.11
G04-j001-plate-39	SUI1 protein	1.99	8.11	3.83	9.99E-03	-0.78
B03-j001-plate-111	No similarity found	2.07	8.17	4.68	1.80E-03	1.50
H04-j001-plate-154	kinesin heavy chain, putative; 55116-47986	2.15	8.35	7.58	2.09E-06	9.33
C08-j001-plate-123	No similarity found	2.25	8.02	4.46	3.62E-03	0.56
G07-j001-plate-198	putative receptor kinase	3.03	7.94	6.31	5.76E-05	5.64
A06-a001-plate-117	hypoxia-responsive family protein	6.62	11.80	8.26	2.33E-07	11.71

^a Codice identificativo interno del laboratorio

^b Proteine con la più alta omologia. Ricavato da BLASTx

^c M: valore medio dei $\log_2(T/C)$, dove T è il segnale della pianta trattata e C del controllo. Indica la variazione di espressione del gene dovuta all'infezione virale.

^d A: valore medio del $\log_2(T^*C)$. Indica per ogni gene il valore medio di espressione.

^e Valore del test t di Student.

^f Probabilità associata al valore t

^a È il *log-odds* che il gene sia differenzialmente espresso. Un valore di B pari a zero indica una probabilità del 50% di differenza di espressione. Per valori positivi la probabilità aumenta (per B=1.5 la probabilità sale al 82%), per negativi diminuisce

Bibliografia

- A. Abu-Arish, D. Frenkiel-Krispin, T. Fricke, T. Tzfira, V. Citovsky, S. G. Wolf, and M. Elbaum. Three-dimensional reconstruction of agrobacterium vire2 protein with single-stranded dna. *J. Biol. Chem.*, 279(24):25359–25363, June 2004. doi: 10.1074/jbc.M401804200. URL <http://dx.doi.org/10.1074/jbc.M401804200>.
- A. Agresti. Tutorial on modeling ordered categorical response data. *Psychological Bulletin*, 105(2):290–301, March 1989.
- P. Ahlquist. Rna-dependent rna polymerases, viruses, and rna silencing. *Science*, 296(5571):1270–1273, May 2002. doi: 10.1126/science.1069132. URL <http://view.ncbi.nlm.nih.gov/pubmed/12016304>.
- A. Al-Taher, A. Bashein, T. Nolan, M. Hollingsworth, and G. Brady. Global cdna amplification combined with real-time rt-pcr: accurate quantification of multiple human potassium channel genes at the single cell level. *Yeast*, 17(3):201–210, 2000. doi: 10.1002/1097-0061(20000930)17:3<201::AID-YEA30>3.0.CO;2-R. URL [http://dx.doi.org/10.1002/1097-0061\(20000930\)17:3<201::AID-YEA30>3.0.CO;2-R](http://dx.doi.org/10.1002/1097-0061(20000930)17:3<201::AID-YEA30>3.0.CO;2-R).
- R. Anandalakshmi, G. J. Pruss, X. Ge, R. Marathe, A. C. Mallory, T. H. Smith, and V. B. Vance. A viral suppressor of gene silencing in plants. *Proc Natl Acad Sci U S A*, 95(22):13079–13084, October 1998. ISSN 0027-8424. doi: 10.1073/pnas.95.22.13079. URL <http://dx.doi.org/10.1073/pnas.95.22.13079>.

- F. Aparicio, C. L. Thomas, C. Lederer, Y. Niu, D. Wang, and A. J. Maule. Virus induction of heat shock protein 70 reflects a general response to protein accumulation in the plant cytosol. *Plant Physiol.*, 138(1):529–536, May 2005. doi: 10.1104/pp.104.058958. URL <http://dx.doi.org/10.1104%2Fpp.104.058958>.
- M. Aranda and A. Maule. Virus-induced host gene shutoff in animals and plants. *Virology*, 243(2):261–267, April 1998. doi: 10.1006/viro.1998.9032. URL <http://dx.doi.org/10.1006%2Fviro.1998.9032>.
- M. A. Aranda, M. Escaler, C. L. Thomas, and A. J. Maule. A heat shock transcription factor in pea is differentially controlled by heat and virus replication. *The Plant Journal*, 20(2):153–161, 1999. doi: 10.1046/j.1365-313x.1999.00586.x. URL <http://www.blackwell-synergy.com/doi/abs/10.1046/j.1365-313x.1999.00586.x>.
- F. M. Ausubel. Are innate immune signaling pathways in plants and animals conserved? *Nature Immunology*, 6(10):973–979, September 2005. ISSN 1529-2908. doi: 10.1038/ni1253. URL <http://dx.doi.org/10.1038/ni1253>.
- D. Baldwin, V. Crane, and D. Rice. A comparison of gel-based, nylon filter and microarray techniques to detect differential rna expression in plants. *Curr Opin Plant Biol*, 2(2):96–103, April 1999. ISSN 1369-5266. doi: 10.1016/S1369-5266(99)80020-X. URL [http://dx.doi.org/10.1016/S1369-5266\(99\)80020-X](http://dx.doi.org/10.1016/S1369-5266(99)80020-X).
- N. Ballas and V. Citovsky. Nuclear localization signal binding protein from arabidopsis mediates nuclear import of agrobacterium vird2 protein. *Proceedings of the National Academy of Sciences*, 94(20):10723–10728, September 1997. doi: 10.1073/pnas.94.20.10723. URL <http://dx.doi.org/10.1073/pnas.94.20.10723>.
- C. J. Barrett and E. S. Kawasaki. Microarrays: the use of oligonucleotides and cDNA for the analysis of gene expression. *Drug Discovery Today*, 8

- (3):134–141, January 2003. doi: 10.1016/S1359-6446(02)02578-3. URL [http://dx.doi.org/10.1016/S1359-6446\(02\)02578-3](http://dx.doi.org/10.1016/S1359-6446(02)02578-3).
- M. Barry and K. Fruh. Viral modulators of cullin ring ubiquitin ligases: Culling the host defense. *Sci. STKE*, 2006(335):pe21+, May 2006. doi: 10.1126/stke.3352006pe21. URL <http://dx.doi.org/10.1126/stke.3352006pe21>.
- D. Baulcombe. Rna silencing in plants. *Nature*, 431(7006):356–363, 2004. doi: 10.1038/nature02874. URL <http://dx.doi.org/10.1038/nature02874>.
- G. Brigneti, O. Voinnet, W.-X. Li, L.-H. Ji, S.-W. Ding, and D. C. Baulcombe. Viral pathogenicity determinants are suppressors of transgene silencing in nicotiana benthamiana. *The EMBO Journal*, 17:6739–6746, 1998.
- K. W. Buck. Replication of tobacco mosaic virus rna. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 354(1383):613–627, March 1999. doi: 10.1098/rstb.1999.0413. URL <http://dx.doi.org/10.1098/rstb.1999.0413>.
- E. J. Chapman, A. I. Prokhnovsky, K. Gopinath, V. V. Dolja, and J. C. Carrington. Viral rna silencing suppressors inhibit the microrna pathway at an intermediate step. *Genes Dev.*, 18(10):1179–1186, May 2004. doi: 10.1101/gad.1201204. URL <http://dx.doi.org/10.1101/gad.1201204>.
- B. Chen, J. H. Jiang, and X. P. Zhou. A tom1 homologue is required for multiplication of tobacco mosaic virus in nicotiana benthamiana. *J Zhejiang Univ Sci B*, 8(4):256–259, April 2007. ISSN 1673-1581. doi: 10.1631/jzus.2007.B0256. URL <http://dx.doi.org/10.1631/jzus.2007.B0256>.
- W. Chen, N. J. Provart, J. Glazebrook, F. Katagiri, H. S. Chang, T. Eulgem, F. Mauch, S. Luan, G. Zou, S. A. Whitham, P. R. Budworth, Y. Tao, Z. Xie, X. Chen, S. Lam, J. A. Kreps, J. F. Harper, A. Si-Ammour, B. Mauch-Mani, M. Heinlein, K. Kobayashi, T. Hohn, J. L.

- Dangl, X. Wang, and T. Zhu. Expression profile matrix of arabidopsis transcription factor genes suggests their putative functions in response to environmental stresses. *Plant Cell*, 14(3):559–574, March 2002. ISSN 1040-4651. doi: 10.1105/tpc.010410. URL <http://dx.doi.org/10.1105/tpc.010410>.
- S. T. Chisholm, G. Coaker, B. Day, and B. J. Staskawicz. Host-microbe interactions: shaping the evolution of the plant immune response. *Cell*, 124(4):803–814, February 2006. ISSN 0092-8674. doi: 10.1016/j.cell.2006.02.008. URL <http://dx.doi.org/10.1016/j.cell.2006.02.008>.
- P. J. Christie. Type iv secretion: the agrobacterium virb/d4 and related conjugation systems. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1694(1-3):219–234, November 2004. doi: 10.1016/j.bbamcr.2004.02.013. URL <http://dx.doi.org/10.1016/j.bbamcr.2004.02.013>.
- V. Citovsky, M. L. Wong, and P. Zambryski. Cooperative interaction of agrobacterium vire2 protein with single-stranded dna: Implications for the t-dna transfer process. *Proceedings of the National Academy of Sciences*, 86(4):1193–1197, February 1989. doi: 10.1073/pnas.86.4.1193. URL <http://dx.doi.org/10.1073/pnas.86.4.1193>.
- A. Conesa, S. Götz, J. M. García-Gómez, J. Terol, M. Talón, and M. Robles. Blast2go: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, 21(18):3674–3676, September 2005. ISSN 1367-4803. doi: 10.1093/bioinformatics/bti610. URL <http://dx.doi.org/10.1093/bioinformatics/bti610>.
- P. Costantino, P. J. J. Hooykaas, H. den Dulk-Ras, and R. A. Schilperoort. Tumor formation and rhizogenicity of agro bacterium rhizogenes carrying ti plasmids. *Gene*, 11(1-2):79–87, October 1980. doi: 10.1016/0378-1119(80)90088-8. URL [http://dx.doi.org/10.1016/0378-1119\(80\)90088-8](http://dx.doi.org/10.1016/0378-1119(80)90088-8).
- B. De Nardi, R. Dreos, L. Del Terra, C. Martellosi, E. Asquini, P. Tornincasa, D. Gasperini, B. Pacchioni, R. Rathinavelu, A. Pallavicini, and

- G. Graziosi. Differential responses of *coffea arabica* l. leaves and roots to chemically induced systemic acquired resistance. *Genome*, 49(12):1594–1605, December 2006. URL <http://rparticle.web-p.cisti.nrc.ca/rparticle/AbstractTemplateServlet?journal=gen&volume=49&year=\&issue=\&msno=g06-125&calyLang=eng>.
- M. Delledonne, Y. Xia, R. A. Dixon, and C. Lamb. Nitric oxide functions as a signal in plant disease resistance. *Nature*, 394(6693):585–588, August 1998. ISSN 0028-0836. doi: 10.1038/29087. URL <http://dx.doi.org/10.1038/29087>.
- Y. L. Dorokhov, P. A. Ivanov, V. K. Novikov, A. A. Agranovsky, Y. S. Morozov, V. A. Efimov, R. Casper, and J. G. Atabekov. Complete nucleotide sequence and genome organization of a tobamovirus infecting cruciferae plants. *FEBS Letters*, 350(1):5–8, April 1994.
- J. Durner, D. Wendehenne, and D. F. Klessig. Defense gene induction in tobacco by nitric oxide, cyclic gmp, and cyclic adp-ribose. *Proc Natl Acad Sci U S A*, 95(17):10328–10333, August 1998. ISSN 0027-8424. doi: 10.1073/pnas.95.17.10328. URL <http://dx.doi.org/10.1073/pnas.95.17.10328>.
- V. A. Efimov, O. G. Chakhmakhcheva, J. Archdeacon, J. M. Fernandez, O. N. Fedorkin, Y. Dorokhov, and J. G. Atabekov. Detection of the 5'-cap structure of messenger rnas with the use of the cap-jumping approach. *Nucl. Acids Res.*, 29(22):4751–4759, November 2001. doi: 10.1093/nar/29.22.4751. URL <http://dx.doi.org/10.1093/nar/29.22.4751>.
- O. N. Fedorkin, O. N. Denisenko, A. S. Sitikov, D. A. Zelenina, L. Lukasheva, I. S. Morozov, and I. G. Atabekov. A protein product of the tobamovirus open translation frame forms a stable complex with translation elongation factor eef-1alpha. *Doklady Akademii nauk*, 343(5):703–704, August 1995.
- S. A. Filichkin and S. B. Gelvin. Formation of a putative relaxation intermediate during t-dna processing directed by the agrobac-

- terium tumefaciens virD1, D2 endonuclease. *Molecular Microbiology*, 8(5):915–926, 1993. doi: 10.1111/j.1365-2958.1993.tb01637.x. URL <http://www.blackwell-synergy.com/doi/abs/10.1111/j.1365-2958.1993.tb01637.x>.
- S. B. Gelvin. Agrobacterium-mediated plant transformation: the biology behind the gene-jockeying tool. *Microbiol Mol Biol Rev*, 67(1), March 2003. ISSN 1092-2172. doi: 10.1128/MMBR.67.1.16-37.2003. URL <http://dx.doi.org/10.1128/MMBR.67.1.16-37.2003>.
- Z. Gong, C.-H. Dong, H. Lee, J. Zhu, L. Xiong, D. Gong, B. Stevenson, and J.-K. Zhu. A dead box rna helicase is essential for mrna export and important for development and stress responses in arabidopsis. *Plant Cell*, 17(1):256–267, January 2005. doi: 10.1105/tpc.104.027557. URL <http://dx.doi.org/10.1105/tpc.104.027557>.
- R. Gorovits, F. Akad, H. Beery, F. Vidavsky, A. Mahadav, and H. Czosnek. Expression of stress-response proteins upon whitefly-mediated inoculation of tomato yellow leaf curl virus in susceptible and resistant tomato plants. *Molecular Plant-Microbe Interactions*, 20(11):1376–1383, November 2007. URL <http://apsjournals.apsnet.org/doi/abs/10.1094/MPMI-20-11-1376>.
- A. K. Grennan. Genevestigator. facilitating web-based gene-expression analysis. *Plant Physiol.*, 141(4):1164–1166, August 2006. doi: 10.1104/pp.104.900198. URL <http://dx.doi.org/10.1104%2Fpp.104.900198>.
- F.-Q. Guo, M. Okamoto, and N. M. Crawford. Identification of a plant nitric oxide synthase gene involved in hormonal signaling. *Science*, 302(5642):100–103, October 2003. doi: 10.1126/science.1086770. URL <http://dx.doi.org/10.1126/science.1086770>.
- Y. Hagiwara, K. Komoda, T. Yamanaka, A. Tamai, T. Meshi, R. Funada, T. Tsuchiya, S. Naito, and M. Ishikawa. Subcellular localization of host

- and viral proteins associated with tobamovirus rna replication. *The EMBO Journal*, 22:344–353, 2003. doi: doi:10.1093/emboj/cdg033. URL <http://dx.doi.org/doi:10.1093/emboj/cdg033>.
- G. J. Hannon. Rna interference. *Nature*, 418(6894):244–251, July 2002. ISSN 0028-0836. doi: 10.1038/418244a. URL <http://dx.doi.org/10.1038/418244a>.
- L. Harrell, U. Melcher, and J. F. Atkins. Predominance of six different hexanucleotide recoding signals 3' of read-through stop codons. *Nucl. Acids Res.*, 30(9):2011–2017, May 2002. doi: 10.1093/nar/30.9.2011. URL <http://dx.doi.org/10.1093/nar/30.9.2011>.
- M. A. Harris, J. Clark, A. Ireland, J. Lomax, M. Ashburner, R. Foulger, K. Eilbeck, S. Lewis, B. Marshall, C. Mungall, J. Richter, G. M. Rubin, J. A. Blake, C. Bult, M. Dolan, H. Drabkin, J. T. Eppig, D. P. Hill, L. Ni, M. Ringwald, R. Balakrishnan, J. M. Cherry, K. R. Christie, M. C. Costanzo, S. S. Dwight, S. Engel, D. G. Fisk, J. E. Hirschman, E. L. Hong, R. S. Nash, A. Sethuraman, C. L. Theesfeld, D. Botstein, K. Dolinski, B. Feierbach, T. Berardini, S. Mundodi, S. Y. Rhee, R. Apweiler, D. Barrell, E. Camon, E. Dimmer, V. Lee, R. Chisholm, P. Gaudet, W. Kibbe, R. Kishore, E. M. Schwarz, P. Sternberg, M. Gwinn, L. Hannick, J. Wortman, M. Berri-man, V. Wood, N. de la Cruz, P. Tonellato, P. Jaiswal, T. Seigfried, and R. a. White. The gene ontology (go) database and informatics resource. *Nucleic Acids Res*, 32(Database issue), January 2004. ISSN 1362-4962. URL <http://view.ncbi.nlm.nih.gov/pubmed/14681407>.
- M. Heil and R. M. Bostock. Induced systemic resistance (isr) against pathogens in the context of induced plant defences. *Ann Bot*, 89(5):503–512, May 2002. doi: 10.1093/aob/mcf076. URL <http://dx.doi.org/10.1093/aob/mcf076>.
- H.-H. Hwang and S. B. Gelvin. Plant proteins that interact with virb2, the agrobacterium tumefaciens pilin protein, mediate plant transforma-

- tion. *Plant Cell*, 16(11):3148–3167, November 2004. doi: 10.1105/tpc.104.026476. URL <http://dx.doi.org/10.1105/tpc.104.026476>.
- K. I. Ivanov, P. A. Ivanov, E. K. Timofeeva, Y. L. Dorokhov, and J. G. Atabekov. The immobilized movement proteins of two tobamoviruses form stable ribonucleoprotein complexes with full-length viral genomic rna. *FEBS Letters*, 346(2-3):217–220, June 1994. doi: 10.1016/0014-5793(94)00477-3. URL [http://dx.doi.org/10.1016/0014-5793\(94\)00477-3](http://dx.doi.org/10.1016/0014-5793(94)00477-3).
- J. D. G. Jones and J. L. Dangl. The plant immune system. *Nature*, 444(7117):323–329, November 2006. ISSN 0028-0836. doi: 10.1038/nature05286. URL <http://dx.doi.org/10.1038/nature05286>.
- A. V. Kajava. Structural diversity of leucine-rich repeat proteins. *Journal of Molecular Biology*, 277(3):519–527, April 1998. doi: 10.1006/jmbi.1998.1643. URL <http://dx.doi.org/10.1006/jmbi.1998.1643>.
- B. C. Kang, Yeam, and M. M. Jahn. Genetics of plant virus resistance. *Annual Review of Phytopathology*, 43(1):581–621, 2005. URL <http://arjournals.annualreviews.org/doi/full/10.1146/annurev.phyto.43.011205.141140?cookieSet=1>.
- O. V. Karpova, N. P. Rodionova, K. I. Ivanov, S. V. Kozlovsky, Y. Dorokhov, and J. G. Atabekov. Phosphorylation of tobacco mosaic virus movement protein abolishes its translation repressing ability. *Virology*, 261(1):20–24, August 1999. doi: 10.1006/viro.1999.9842. URL <http://dx.doi.org/10.1006/viro.1999.9842>.
- K. D. Kasschau, Z. Xie, E. Allen, C. Llave, E. J. Chapman, K. A. Krizan, and J. C. Carrington. P1/hc-pro, a viral suppressor of rna silencing, interferes with arabidopsis development and mirna function. *Developmental Cell*, 4:205–217, 2003.
- Y. Klipper-Aurbach, M. Wasserman, N. Braunsriegel-Weintrob, D. Borstein, S. Peleg, S. Assa, M. Karp, Y. Benjamini, Y. Hochberg, and Z. Laron.

- Mathematical formulae for the prediction of the residual beta cell function during the first two years of disease in children and adolescents with insulin-dependent diabetes mellitus. *Medical Hypotheses*, 45(5):486–490, November 1995. doi: 10.1016/0306-9877(95)90228-7. URL [http://dx.doi.org/10.1016/0306-9877\(95\)90228-7](http://dx.doi.org/10.1016/0306-9877(95)90228-7).
- K. Klumpp, R. Ruigrok, and F. Baudin. Roles of the influenza virus polymerase and nucleoprotein in forming a functional rnp structure. *The EMBO Journal*, 16:1248–1257, 1997.
- K. Kubota, S. Tsuda, A. Tamai, and T. Meshi. Tomato mosaic virus replication protein suppresses virus-targeted posttranscriptional gene silencing. *J Virol*, 77(20):11016–11026, October 2003. ISSN 0022-538X. URL <http://view.ncbi.nlm.nih.gov/pubmed/14512550>.
- S. La Camera, G. Gouzerh, S. Dhondt, L. Hoffmann, B. Fritig, M. Legrand, and T. Heitz. Metabolic reprogramming in plant innate immunity: the contributions of phenylpropanoid and oxylipin pathways. *Immunological Reviews*, 198(1):267–284, 2004. doi: 10.1111/j.0105-2896.2004.0129.x. URL <http://www.blackwell-synergy.com/doi/abs/10.1111/j.0105-2896.2004.0129.x>.
- R. T. Lartey, T. C. Voss, and U. Melcher. Completion of a cDNA sequence from a tobamovirus pathogenic to crucifers. *Gene*, 166(2):331–332, December 1995. doi: 10.1016/0378-1119(95)00674-5. URL [http://dx.doi.org/10.1016/0378-1119\(95\)00674-5](http://dx.doi.org/10.1016/0378-1119(95)00674-5).
- W. Liu and D. A. Saint. Validation of a quantitative method for real time pcr kinetics. *Biochem Biophys Res Commun*, 294(2):347–353, June 2002. ISSN 0006-291X. doi: 10.1016/S0006-291X(02)00478-3. URL [http://dx.doi.org/10.1016/S0006-291X\(02\)00478-3](http://dx.doi.org/10.1016/S0006-291X(02)00478-3).
- W. Ma and G. A. Berkowitz. The grateful dead: calcium and cell death in plant innate immunity. *Cellular Microbiology*, 9(11):2571–2585, November 2007. ISSN 1462-5814. doi: 10.1111/j.1462-5822.2007.

- 01031.x. URL <http://www.blackwell-synergy.com/doi/abs/10.1111/j.1462-5822.2007.01031.x>.
- E. A. Maher, N. J. Bate, W. Ni, Y. Elkind, R. A. Dixon, and C. J. Lamb. Increased disease susceptibility of transgenic tobacco plants with suppressed levels of preformed phenylpropanoid products. *Proceedings of the National Academy of Sciences*, 91(16):7802–7806, August 1994. doi: 10.1073/pnas.91.16.7802. URL <http://dx.doi.org/10.1073%2Fpnas.91.16.7802>.
- R. Marathe, Z. Guan, R. Anandalakshmi, H. Zhao, and S. P. Dinesh-Kumar. Study of arabidopsis thaliana resistome in response to cucumber mosaic virus infection using whole genome microarray. *Plant Mol Biol*, 55(4):501–520, July 2004. ISSN 0167-4412. doi: 10.1007/s11103-004-0439-0. URL <http://dx.doi.org/10.1007/s11103-004-0439-0>.
- S. Marillonnet, C. Thoeringer, R. Kandzia, V. Klimyuk, and Y. Gleba. Systemic agrobacterium tumefaciens-mediated transfection of viral replicons for efficient transient expression in plants. *Nature Biotechnology*, 23(6):718–723, May 2005. ISSN 1087-0156. doi: 10.1038/nbt1094. URL <http://dx.doi.org/10.1038/nbt1094>.
- G. B. Martin, A. J. Bogdanove, and G. Sessa. Understanding the functions of plant disease resistance proteins. *Annual Review of Plant Biology*, 54(1):23–61, 2003. URL <http://arjournals.annualreviews.org/doi/abs/10.1146%2Fannurev.arplant.54.031902.135035>.
- C. A. McCullen and A. N. Binns. Agrobacterium tumefaciens and plant cell interactions and activities required for interkingdom macromolecular transfer. *Annual Review of Cell and Developmental Biology*, 22(1):101–127, 2006. URL <http://arjournals.annualreviews.org/doi/abs/10.1146/annurev.cellbio.22.011105.102022>.
- L. Mchale, X. Tan, P. Koehl, and R. Michelmore. Plant nbs-llr proteins: adaptable guards. *Genome Biology*, 7(4), 2006. doi: 10.1186/gb-2006-7-4-212. URL <http://dx.doi.org/10.1186/gb-2006-7-4-212>.

- J. Meijerink, C. Mandigers, L. van de Locht, E. Tonnissen, F. Goodsaid, and J. Raemaekers. A novel method to compensate for different amplification efficiencies between patient dna samples in quantitative real-time pcr. *J Mol Diagn*, 3(2):55–61, May 2001. URL <http://jmd.amjpathol.org/cgi/content/abstract/3/2/55>.
- G. Meister and T. Tuschl. Mechanisms of gene silencing by double-stranded rna. *Nature*, 431(7006):343–349, 2004. doi: 10.1038/nature02873. URL <http://dx.doi.org/10.1038/nature02873>.
- U. Melcher. The ?30k? superfamily of viral movement proteins. *J Gen Virol*, 81(1):257–266, January 2000. URL <http://vir.sgmjournals.org/cgi/content/abstract/81/1/257>.
- U. Melcher. Turnip vein-clearing virus, from pathogen to host expression profile. pages 133–140, May 2003. URL <http://www.ingentaconnect.com/content/bsc/mpp/2003/00000004/00000003/art00001>.
- B. C. Meyers, A. Kozik, A. Griego, H. Kuang, and R. W. Michelmore. Genome-wide analysis of nbs-llr-encoding genes in arabidopsis. *Plant Cell*, 15(4):809–834, April 2003. doi: 10.1105/tpc.009308. URL <http://dx.doi.org/10.1105/tpc.009308>.
- B. Monosi, R. Wisser, L. Pennill, and S. Hulbert. Full-genome analysis of resistance gene homologues in rice. *TAG Theoretical and Applied Genetics*, 109(7):1434–1447, November 2004. doi: 10.1007/s00122-004-1758-x. URL <http://dx.doi.org/10.1007/s00122-004-1758-x>.
- S. Morozov, O. N. Fedorkin, G. Juttner, J. Schiemann, D. C. Baulcombe, and J. G. Atabekov. Complementation of a potato virus x mutant mediated by bombardment of plant tissues with cloned viral movement protein genes. *J Gen Virol*, 78(8):2077–2083, August 1997. URL <http://vir.sgmjournals.org/cgi/content/abstract/78/8/2077>.

- Y. S. Morozov, O. N. Denisenko, D. A. Zelenina, O. N. Fedorkin, A. G. Solovyev, E. Maiss, R. Casper, and J. G. Atabekov. A novel open reading frame in tobacco mosaic virus genome coding for a putative small, positively charged protein. *Biochimie*, 75(8):659–665, 1993.
- A. Naderi, A. Ahmed, N. B. Morais, S. Aparicio, J. Brenton, and C. Caldas. Expression microarray reproducibility is improved by optimising purification steps in rna amplification and labelling. *BMC Genomics*, 5(1), 2004. doi: 10.1186/1471-2164-5-9. URL <http://dx.doi.org/10.1186/1471-2164-5-9>.
- L. Navazio, B. Baldan, R. Moscatiello, A. Zuppini, S. L. Woo, P. Mariani, and M. Lorito. Calcium-mediated perception and defense responses activated in plant cells by metabolite mixtures secreted by the biocontrol fungus *trichoderma atroviride*. *BMC Plant Biology*, 7:41+, July 2007. ISSN 1471-2229. doi: 10.1186/1471-2229-7-41. URL <http://dx.doi.org/10.1186/1471-2229-7-41>.
- T. Nurnberger and B. Kemmerling. Receptor protein kinases - pattern recognition receptors in plant immunity. *Trends in Plant Science*, 11(11):519–522, November 2006. doi: 10.1016/j.tplants.2006.09.005. URL <http://dx.doi.org/10.1016/j.tplants.2006.09.005>.
- L. Otten, H. Greve, J. Leemans, R. Hain, P. Hooykaas, and J. Schell. Restoration of virulence of vir region mutants of *agrobacterium tumefaciens* strain b6s3 by coinfection with normal and mutant *agrobacterium* strains. *Molecular and General Genetics MGG*, 195(1):159–163, June 1984. doi: 10.1007/BF00332739. URL <http://dx.doi.org/10.1007/BF00332739>.
- W. Park, J. Li, R. Song, J. Messing, and X. Chen. Carpel factory, a dicer homolog, and hen1, a novel protein, act in microRNA metabolism in *arabidopsis thaliana*. *Current Biology*, 12:1484–1495, 2002.

- M. W. Pfaffl. A new mathematical model for relative quantification in real-time rt-pcr. *Nucl. Acids Res.*, 29(9):e45+, May 2001. doi: 10.1093/nar/29.9.e45. URL <http://dx.doi.org/10.1093/nar/29.9.e45>.
- J. Quackenbush. Microarray data normalization and transformation. *Nature Genetics*, 32:496–501, 2002.
- C. Ramakers, J. M. Ruijter, R. H. Deprez, and A. F. Moorman. Assumption-free analysis of quantitative real-time polymerase chain reaction (pcr) data. *Neuroscience Letters*, 339(1):62–66, March 2003. doi: 10.1016/S0304-3940(02)01423-4. URL <http://dx.doi.org/10.1016%2FS0304-3940%2802%2901423-4>.
- F. Ratcliff, B. D. Harrison, and D. Baulcombe. A similarity between viral defense and gene silencing in plants. *Science*, 276(5318):1558–1560, June 1997. doi: 10.1126/science.276.5318.1558. URL <http://dx.doi.org/10.1126/science.276.5318.1558>.
- A. Reiner, D. Yekutieli, and Y. Benjamini. Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics*, 19(3):368–375, February 2003. ISSN 1367-4803. doi: 10.1093/bioinformatics/btf877. URL <http://dx.doi.org/10.1093/bioinformatics/btf877>.
- T. Romeis, P. Piedras, and J. D. Jones. Resistance gene-dependent activation of a calcium-dependent protein kinase in the plant defense response. *Plant Cell*, 12(5):803–816, May 2000. doi: 10.1105/tpc.12.5.803. URL <http://dx.doi.org/10.1105/tpc.12.5.803>.
- D. Rotenberg, T. S. Thompson, T. L. German, and D. K. Willis. Methods for effective real-time rt-pcr analysis of virus-induced gene silencing. *Journal of Virological Methods*, 138(1-2):49–59, December 2006. doi: 10.1016/j.jviromet.2006.07.017. URL <http://dx.doi.org/10.1016/j.jviromet.2006.07.017>.

- H. Salman, A. Abu-Arish, S. Oliel, A. Loyter, J. Klafter, R. Granek, and M. Elbaum. Nuclear localization signal peptides induce molecular delivery along microtubules. *Biophys. J.*, 89(3):2134–2145, September 2005. doi: 10.1529/biophysj.105.060160. URL <http://dx.doi.org/10.1529/biophysj.105.060160>.
- M. Scheideler, N. L. Schlaich, K. Fellenberg, T. Beissbarth, N. C. Hauser, M. Vingron, A. J. Slusarenko, and J. D. Hoheisel. Monitoring the switch from housekeeping to pathogen defense metabolism in arabidopsis thaliana using cdna arrays. *J. Biol. Chem.*, 277(12):10555–10561, March 2002. doi: 10.1074/jbc.M104863200. URL <http://dx.doi.org/10.1074/jbc.M104863200>.
- M. Schena, D. Shalon, R. W. Davis, and P. O. Brown. Quantitative monitoring of gene expression patterns with a complementary dna microarray. *Science*, 270(5235):467–470, October 1995. ISSN 0036-8075. doi: 10.1126/science.270.5235.467. URL <http://dx.doi.org/10.1126/science.270.5235.467>.
- P. M. Schenk, K. Kazan, I. Wilson, J. P. Anderson, T. Richmond, S. C. Somerville, and J. M. Manners. Coordinated plant defense responses in arabidopsis revealed by microarray analysis. *Proceedings of the National Academy of Sciences*, 97(21):11655–11660, October 2000. doi: 10.1073/pnas.97.21.11655. URL <http://dx.doi.org/10.1073/pnas.97.21.11655>.
- J. Schuchhardt, D. Beule, A. Malik, E. Wolski, H. Eickhoff, H. Lehrach, and H. Herzel. Normalization strategies for cdna microarrays. *Nucleic Acids Res*, 28(10), May 2000. ISSN 1362-4962. doi: 10.1093/nar/28.10.e47. URL <http://dx.doi.org/10.1093/nar/28.10.e47>.
- G. Senthil, H. Liu, V. G. Puram, A. Clark, A. Stromberg, and M. M. Goodin. Specific and common changes in nicotiana benthamiana gene expression in response to infection by enveloped viruses. *J Gen Virol*, 86(9):2615–2625,

- September 2005. doi: 10.1099/vir.0.81043-0. URL <http://dx.doi.org/10.1099%2Fvir.0.81043-0>.
- G. K. Smyth and T. P. Speed. Normalization of cdna microarray data. *Methods*, 31:265–273, 2003.
- J. L. Soosaar, T. M. Burch-Smith, and S. P. Dinesh-Kumar. Mechanisms of plant resistance to viruses. *Nat Rev Micro*, 3(10):789–798, October 2005. doi: 10.1038/nrmicro1239. URL <http://dx.doi.org/10.1038/nrmicro1239>.
- P. A. C. 't Hoen, F. de Kort, G. J. B. van Ommen, and J. T. den Dunnen. Fluorescent labelling of crna for microarray applications. *Nucl. Acids Res.*, 31(5):e20+, March 2003. doi: 10.1093/nar/gng020. URL <http://dx.doi.org/10.1093/nar/gng020>.
- W. Tameling and F. Takken. Resistance proteins: scouts of the plant innate immune system. *European Journal of Plant Pathology*, September 2007. doi: 10.1007/s10658-007-9187-8. URL <http://dx.doi.org/10.1007/s10658-007-9187-8>.
- O. Thimm, O. Blasing, Y. Gibon, A. Nagel, S. Meyer, P. Kruger, J. Selbig, L. A. Muller, S. Y. Rhee, and M. Stitt. mapman: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *The Plant Journal*, 37(6):914–939, 2004. doi: 10.1111/j.1365-313X.2004.02016.x. URL <http://www.blackwell-synergy.com/doi/abs/10.1111/j.1365-313X.2004.02016.x>.
- H. Thordal-Christensen. Fresh insights into processes of nonhost resistance. *Current Opinion in Plant Biology*, 6(4):351–357, August 2003. doi: 10.1016/S1369-5266(03)00063-3. URL [http://dx.doi.org/10.1016/S1369-5266\(03\)00063-3](http://dx.doi.org/10.1016/S1369-5266(03)00063-3).
- S. W. Tighe, P. de Lajudie, K. Dipietro, K. Lindstrom, G. Nick, and B. D. W. Jarvis. Analysis of cellular fatty acids and phenotypic re-

- relationships of agrobacterium, bradyrhizobium, mesorhizobium, rhizobium and sinorhizobium species using the sherlock microbial identification system. *Int J Syst Evol Microbiol*, 50(2):787–801, March 2000. URL <http://ijs.sgmjournals.org/cgi/content/abstract/50/2/787>.
- T. W. Traut. The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites. *European Journal of Biochemistry*, 222(1):9–19, 1994. doi: 10.1111/j.1432-1033.1994.tb18835.x. URL <http://www.blackwell-synergy.com/doi/abs/10.1111/j.1432-1033.1994.tb18835.x>.
- Y. Tsujimoto, T. Numaga, K. Ohshima, M.-A. Yano, R. Ohsawa, D. B. Goto, S. Naito, and M. Ishikawa. Arabidopsis tobamovirus multiplication (tom) 2 locus encodes a transmembrane protein that interacts with tom1. *The EMBO journal*, 22:335–343, 2003. doi: doi:10.1093/emboj/cdg034. URL <http://dx.doi.org/doi:10.1093/emboj/cdg034>.
- T. Tzfira and V. Citovsky. Agrobacterium-mediated genetic transformation of plants: biology and biotechnology. *Current Opinion in Biotechnology*, 17(2):147–154, April 2006. doi: 10.1016/j.copbio.2006.01.009. URL <http://dx.doi.org/10.1016/j.copbio.2006.01.009>.
- E. A. van der Biezen and J. D. Jones. The nb-arc domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. *Current Biology*, 8(7):R226–R228, March 1998. doi: 10.1016/S0960-9822(98)70145-9. URL [http://dx.doi.org/10.1016/S0960-9822\(98\)70145-9](http://dx.doi.org/10.1016/S0960-9822(98)70145-9).
- E. A. Van Der Biezen and J. D. Jones. Plant disease-resistance proteins and the gene-for-gene concept. *Trends in Biochemical Sciences*, 23(12):454–456, December 1998. doi: 10.1016/S0968-0004(98)01311-5. URL [http://dx.doi.org/10.1016/S0968-0004\(98\)01311-5](http://dx.doi.org/10.1016/S0968-0004(98)01311-5).

- H. Vaucheret, C. Beclin, and M. Fagard. Post-transcriptional gene silencing in plants. *J Cell Sci*, 114(17):3083–3091, September 2001. URL <http://view.ncbi.nlm.nih.gov/pubmed/11590235>.
- A. C. Vergunst, B. Schrammeijer, A. den Dulk-Ras, C. M. de Vlaam, T. J. Regensburg-Tuink, and P. J. Hooykaas. Virb/d4-dependent protein translocation from agrobacterium into plant cells. *Science*, 290(5493):979–982, November 2000. doi: 10.1126/science.290.5493.979. URL <http://dx.doi.org/10.1126/science.290.5493.979>.
- P. M. Waterhouse, M. W. Graham, and M.-B. Wang. Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense rna. *Proceedings of the National Academy of Sciences*, 95(23):13959–13964, November 1998. doi: 10.1073/pnas.95.23.13959. URL <http://view.ncbi.nlm.nih.gov/pubmed/9811908>.
- F. F. White and E. W. Nester. Hairy root: plasmid encodes virulence traits in agrobacterium rhizogenes. *J. Bacteriol.*, 141(3):1134–1141, March 1980. URL <http://jb.asm.org/cgi/content/abstract/141/3/1134>.
- S. A. Wingard. Hosts and symptoms of ring spot, a virus disease of plants. *Journal of Agricultural Research*, 37:127–153, 1928.
- C. C. Xiang and Y. Chen. cdna microarray technology and its applications. *Biotechnology Advances*, 18(1):35–46, March 2000. doi: 10.1016/S0734-9750(99)00035-X. URL [http://dx.doi.org/10.1016/S0734-9750\(99\)00035-X](http://dx.doi.org/10.1016/S0734-9750(99)00035-X).
- Z. Xie, L. K. Johansen, A. M. Gustafson, K. D. Kasschau, A. D. Lellis, D. Zilberman, S. E. Jacobsen, and J. C. Carrington. Genetic and functional diversification of small rna pathways in plants. *PLoS Biol*, 2(5), May 2004. ISSN 1545-7885. doi: 10.1371/journal.pbio.0020104. URL <http://dx.doi.org/10.1371/journal.pbio.0020104>.

- T. Yamanaka, T. Ohta, M. Takahashi, T. Meshi, R. Schmidt, C. Dean, S. Naito, and M. Ishikawa. Tom1, an arabidopsis gene required for efficient multiplication of a tobamovirus, encodes a putative transmembrane protein. *Proc Natl Acad Sci U S A*, 97(18):10107–10112, August 2000. ISSN 0027-8424. doi: 10.1073/pnas.170295097. URL <http://dx.doi.org/10.1073/pnas.170295097>.
- T. Yamanaka, T. Imai, R. Satoh, A. Kawashima, M. Takahashi, K. Tomita, K. Kubota, T. Meshi, S. Naito, and M. Ishikawa. Complete inhibition of tobamovirus multiplication by simultaneous mutations in two homologous host genes. *J Virol*, 76(5):2491–2497, March 2002. ISSN 0022-538X. doi: 10.1128/jvi.76.5.2491-2497.2002. URL <http://dx.doi.org/10.1128/jvi.76.5.2491-2497.2002>.
- H. Yang, Nairn, and P. Ozias-Akins. Transformation of peanut using a modified bacterial mercuric ion reductase gene driven by an actin promoter from arabidopsis thaliana. *Journal of Plant Physiology*, 160(8):945–952, 2003. doi: 10.1078/0176-1617-01087. URL <http://dx.doi.org/10.1078/0176-1617-01087>.
- Y. H. Yang, S. Dudoit, P. Luu, D. M. Lin, V. Peng, J. Ngai, and T. P. Speed. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res*, 30(4), February 2002. ISSN 1362-4962. doi: 10.1093/nar/30.4.e15. URL <http://dx.doi.org/10.1093/nar/30.4.e15>.
- K. Ye, L. Malinina, and D. J. Patel. Recognition of small interfering rna by a viral suppressor of rna silencing. *Nature*, 426(6968):874–878, December 2003. doi: 10.1038/nature02213. URL <http://dx.doi.org/10.1038/nature02213>.
- G. Zabala, J. Zou, J. Tuteja, D. O. Gonzalez, S. J. Clough, and L. O. Vodkin. Transcriptome changes in the phenylpropanoid pathway of glycine max in response to pseudomonas syringae infection. *BMC Plant Biology*, 6:26+,

November 2006. ISSN 1471-2229. doi: 10.1186/1471-2229-6-26. URL <http://dx.doi.org/10.1186%2F1471-2229-6-26>.

D. Zeidler, U. Zahringer, I. Gerber, I. Dubery, T. Hartung, W. Bors, P. Hut-
zler, and J. Durner. From the cover: Innate immunity in arabidopsis
thaliana: Lipopolysaccharides activate nitric oxide synthase (nos) and in-
duce defense genes. *Proceedings of the National Academy of Sciences*, 101
(44):15811–15816, November 2004. doi: 10.1073/pnas.0404536101. URL
<http://dx.doi.org/10.1073/pnas.0404536101>.