

ALESSIO APRILE^{*,**}, ANNA MARIA MASTRANGELO^{***},
ANNA MARIA DE LEONARDIS^{***}, GABOR GALIBA^{****},
ENRICA RONCAGLIA^{*****}, FRANCESCO FERRARI^{*****},
LUIGI DE BELLIS^{**}, LUANA TURCHI^{*****}, GIOVANNI GIULIANO^{*****},
LUIGI CATTIVELLI^{***}

Studio dei profili trascrizionali di frumento tenero e duro in risposta a stress idrico durante lo stadio di riempimento delle cariossidi

INTRODUZIONE

Le piante di interesse agrario sono organismi che trascorrono la quasi totalità del loro ciclo vitale fissati a un substrato e come tali sono fortemente dipendenti dal loro habitat. Tuttavia le piante hanno colonizzato ambienti estremamente ostili dimostrando capacità di adattamento alle condizioni avverse.

La crescita e lo sviluppo delle piante dipendono dalle interazioni del genotipo con vari fattori esterni, come intensità luminosa, temperatura, disponibilità idrica, dei nutrienti, ecc. Ogni volta che nell'ambiente si verifica una variazione, in difetto o in eccesso, si parla di "stress" perché l'organismo vivente soggiace a modifiche potenzialmente dannose.

Tra i fattori abiotici, che hanno caratterizzato e continuano a caratterizzare l'evoluzione delle piante, la disponibilità d'acqua è senza ombra di dubbio il più importante. Lo stress idrico nel suo più ampio significato ingloba sia la carenza d'acqua che lo stress da salinità e parte dello stress da congelamento. Gli stress da siccità unitamente allo stress da basse temperature rappresentano il problema maggiore per l'agricoltura poiché di difficile controllo con le tecniche colturali e impediscono alla pianta di esprimere il suo potenziale produttivo.

* Dipartimento di Scienze e Tecnologie Biologiche ed Ambientali, Università del Salento

** CRA-Centro di Genomica e Post-Genomica, Fiorenzuola d'Arda, Piacenza

*** CRA-Centro di Ricerca per la Cerealicoltura, Foggia

**** Agricultural Research Institute of the Hungarian Academy of Sciences, Martonvasar, Hungary

***** Dipartimento di Scienze Biomediche, Università di Modena e Reggio Emilia

***** Dipartimento di Biologia, Università di Padova

***** ENEA-Centro di Ricerca Casaccia, Roma

È stato stimato che la quantità di acqua sulla terra è di 14 milioni di km³, e solo lo 0,003% di questi, circa 45.000 km³ sono definiti “fresh water resources”, acqua che teoricamente può essere destinata a uso potabile, per l'agricoltura e l'industria. Di questi solo 9.000-14.000 km³ sono disponibili per l'uso umano.

Nell'area mediterranea, che comprende 25 Paesi, 8 (con circa 115 milioni di abitanti) sono sotto la soglia minima di disponibilità idrica e i Paesi più esposti sono Spagna, Portogallo, Grecia e Italia. Inoltre si calcola che la desertificazione minacci direttamente 250 milioni di persone e indirettamente circa un miliardo (dati dal sito: www.fao.org). In questa prospettiva appare sempre più importante la razionalizzazione dell'uso dell'acqua sia attraverso le pratiche colturali che con l'ottenimento di genotipi tolleranti la carenza idrica e dalle spiccate attitudini nella WUE (water use efficiency) (Rizza et al., 2004).

Lo stress idrico determina un'ampia gamma di cambiamenti fisiologici e biochimici nelle piante. Dal punto di vista fisiologico le risposte delle piante hanno lo scopo principale di ridurre una ulteriore perdita di acqua per traspirazione, chiudendo gli stomi o reprimendo la crescita cellulare e della fotosintesi (Cattivelli et al., 2008). L'inibizione dell'espansione cellulare si verifica con un rallentamento dell'espansione fogliare e risulta uno dei primi meccanismi di risposta al deficit idrico. Un'area fogliare ridotta traspira meno, preservando e utilizzando acqua per un periodo più lungo.

La risposta allo stress inizia quando una pianta percepisce uno stress a livello cellulare. Il riconoscimento dello stress attiva le vie di trasduzione del segnale che trasmettono l'informazione all'interno delle singole cellule e in tutta la pianta. I cambiamenti dell'espressione genica che alla fine avvengono a livello cellulare sono integrati in una risposta da parte della pianta che può modificare la crescita e lo sviluppo e influenzare anche le capacità riproduttive. Non sono ancora completamente chiare le complesse vie di trasduzione del segnale che si pensa partecipino ai cambiamenti dell'espressione genica della pianta in risposta allo stress idrico. Tuttavia ci sono prove evidenti che indicano come la regolazione del segnale di risposta delle piante allo stress coinvolge gli ormoni, specialmente l'ABA, secondi messaggeri come il Ca²⁺ (Sanders et al., 1999; Knight, 2000) e le specie reattive dell'ossigeno (ROS, reactive oxigene species) (Mittler et al., 2002).

I fattori trascrizionali (TF) rappresentano l'ultimo anello della catena di trasduzione del segnale, prima dell'attivazione vera e propria di geni specifici coinvolti nella risposta allo stress. Diverse famiglie di fattori di trascrizione sono in grado di attivare i geni tipici della risposta allo stress idrico. I fattori

di trascrizione di tipo DREB2 e WRKY38, per esempio, sono stati descritti essere coinvolti nella risposta agli stress di tipo osmotico e sono in grado di regolare la trascrizione di geni inducibili da stress idrico (Sakuma et al., 2006; Marè et al., 2004). Altri TF della famiglia dei DREB inducibili da stress idrico sono CBF4/DREB1D (Haake et al., 2002), DDF1/DREB1F e DDF2/DREB1E (Sakuma et al., 2002). Inoltre, noti per essere a valle del segnale di trasduzione dell'ABA, sono i geni appartenenti alle famiglie bZip (Choi et al., 2000), MYC e MYB (Shinozaki and Yamaguchi-Shinozaki 2007).

Spesso i geni target di questi fattori di trascrizione sono enzimi coinvolti nella biosintesi degli osmoliti e proteine con funzione specifica nella tolleranza allo stress. L'accumulo di osmoliti determina un aumento del potenziale osmotico in modo da contrastare la perdita d'acqua dovuta a stress idrico (un processo definito aggiustamento osmotico). Effettori di questa risposta sono gli osmoliti, un piccolo gruppo di composti organici chimicamente diversi, altamente solubili, che non interferiscono con il metabolismo cellulare e che consentono alle piante di abbassare il potenziale idrico e mantenere il turgore durante condizioni di essiccamento o di salinità (prolina, sorbitolo, mannitolo, glicina-betaina, ecc.).

Una delle risposte primarie alla carenza d'acqua è la produzione di acido abscissico (ABA) che causa sia la chiusura degli stomi al fine di regolare la traspirazione fogliare, sia risposte di tipo molecolare inducendo l'espressione di geni indotti da stress (come i geni per la sintesi di osmoliti). È stato ampiamente descritto come molti geni indotti da stress idrico siano anche indotti da applicazioni esogene di ABA (Seki et al., 2002b), mentre altri siano del tutto indifferenti alla presenza di ABA esogeno, dimostrando che esistono due vie di regolazione dei geni indotti da stress idrico (Shinozaki and Yamaguchi-Shinozaki, 2007).

I geni che vengono repressi o indotti in seguito allo stress idrico sono numerosi (Seki et al., 2002a). Molti di questi prodotti genici hanno un ruolo diretto nella risposta e tolleranza allo stress idrico, altri sono semplicemente coregolati. Piante modificate geneticamente allo scopo di introdurre alcuni di questi geni stress-inducibili hanno portato all'ottenimento di piante più tolleranti allo stress idrico (Zhang et al., 2004; Umezawa, 2006a; Umezawa, 2006b). Nel corso degli ultimi anni numerosi geni indotti da stress idrico sono stati identificati in *Arabidopsis*, riso, vite, mais e orzo (Cramer et al., 2007; Atienza et al., 2004) utilizzando l'approccio microarray.

Per poter agire in modo consapevole al fine di aumentare la tolleranza delle piante allo stress idrico è fondamentale identificare il maggior numero di geni indotti da stress, capire le funzioni dei singoli geni, ma soprattutto capire

le relazioni esistenti fra questi, chi sono gli effettori, e quali sono i promotori della risposta che porta alla tolleranza.

In questo lavoro è stato preso come oggetto di studio il frumento, una specie di grande interesse commerciale nell'area mediterranea, la cui produttività può essere seriamente compromessa dalla carenza idrica. Per identificare cambiamenti dei livelli di espressione genica coinvolti nella risposta allo stress idrico sono stati analizzati i profili di trascrizione di un frumento duro (Creso), uno tenero (Chinese Spring) e una sua linea di delezione, priva di una porzione del braccio lungo del cromosoma 5A recante geni di risposta agli stress abiotici (Cattivelli et al., 2002). A ognuno di questi tre genotipi sono stati imposti tre livelli progressivi di stress idrico.

MATERIALI E METODI

Materiali genetici e condizioni di crescita

Il genotipo di frumento duro (cultivar Creso) e i due genotipi di frumento tenero (cultivar Chinese Spring (CS) e una sua linea di delezione per una porzione del braccio lungo del cromosoma 5A, Chinese Spring-5AL) sono stati fatti crescere in celle climatiche a temperatura, umidità, luce e fotoperiodo controllati. Le sementi sono state poste in vaso in un terreno misto di terra, sabbia e torba (6:3:1) e le condizioni di crescita erano le seguenti: a 10°C di giorno, 7°C la notte, umidità relativa del 60%, 12 ore di luce, densità della luce 500 μmol di fotoni $\text{m}^{-2} \text{s}^{-1}$ fino allo spuntare della terza foglia; 22°C di giorno, 18°C la notte, umidità relativa del 55%, 16 ore di luce, densità della

GENOTIPO	TRATTAMENTO	Ψ_w (MPa)	SWC	DURATA STRESS
Creso	CTRL	14±0,58	28%	
	MS	26±1,73	18%	4 days
	SS	39±1,15	12.5%	4 days
CS	CTRL	12±0,58	28%	
	MS	22±0,58	18%	6 days
	SS	33±0,58	12.5%	6 days
CS-5AL	CTRL	11±0,58	28%	
	MS	21±1,15	18%	6 days
	SS	34±1,73	12.5%	6 days

Tab. 1 Nella tabella sono riassunte le condizioni di stress imposte ai singoli campioni, il valore del potenziale idrico di riferimento per definire lo stress moderato o intenso e, ultima colonna, i giorni necessari per il raggiungimento di quel potenziale osmotico nelle relative condizioni di SWC

luce 500 μmol di fotoni $\text{m}^{-2} \text{s}^{-1}$ dallo stadio di terza foglia fino a tre giorni dopo l'antesi. In queste condizioni sono stati necessari circa due mesi per raggiungere lo stadio di antesi.

Disegno sperimentale

Il contenuto idrico del terreno (Soil Water Content SWC) è stato attentamente mantenuto vicino alle condizioni di campo (SWC=28%) finché le piante non hanno raggiunto la fase di antesi. Le piante sono state monitorate mediante misurazione del potenziale idrico con una camera a pressione (PMS Instrument Co., Corvallis, OR, USA).

Il contenuto idrico del terreno è stato imposto a partire dallo stadio di antesi. Le condizioni del controllo (CTRL) sono state mantenute costanti, mentre per quanto riguarda il terreno delle piante stressate si è apportata una riduzione del SWC fino al 18% (per lo stress moderato, MS) e fino al 12,5% per lo stress intenso (SS). Le condizioni di stress sono state prolungate per il tempo necessario affinché la pianta raggiungesse nella foglia bandiera un potenziale idrico pari a 2.1/2.6 Mpa e 3.3/3.9 Mpa rispettivamente nello stress moderato e nello stress intenso. Per ognuna di queste condizioni sono state eseguite tre repliche biologiche, per un totale di 27 campioni.

Obiettivo di questo lavoro era il monitoraggio dell'espressione genica in condizioni di stress idrico durante la fase di riempimento delle cariossidi. A tal scopo sono stati prelevati campioni di glume e glumelle, tessuti che rimangono verdi e fotosintetizzanti anche durante le ultime fasi di crescita delle piante di frumento.

Estrazione dell'RNA e ibridazione dei GeneChip® Wheat Genome Array

L'RNA totale è stato estratto dai tessuti di rivestimento delle spighe, glume e glumelle. Il protocollo utilizzato è lo stesso pubblicato sul sito web dedicato ai progetti dell'Arabidopsis Functional genomics Consortium (<https://www.arabidopsis.org/portals/masc/AFGC/RevisedAFGC/site2RnaL.htm#isolation>), derivato dalla modifica del protocollo sviluppato da Chomczynski e Sacchi (1987). L'RNA totale ottenuto è stato successivamente purificato utilizzando le colonnine RNeasy seguendo raccomandazioni e protocollo descritto nel Qiagen RNeasy Mini Handbook, Aprile 2006, pp. 56-57 (Qiagen Inc., Valencia, CA). La qualità e la quantità dell'RNA estratto sono

state misurate eseguendo diluzioni seriali dei campioni ottenuti e analizzandole con lo strumento Agilent Bioanalyzer 2100 con il nano kit Agilent RNA 6000 (Agilent Technologies, Palo Alto, CA).

L'RNA di alta qualità è stato quindi processato seguendo le istruzioni raccomandate da Affymetrix (Affymetrix GeneChip Expression Analysis Technical Manual, Affymetrix, Inc., Santa Clara, CA). In breve: si sintetizza il primo filamento di cDNA e successivamente anche il complementare sfruttando i polyA-RNA presenti nell'RNA totale (5 µg di RNA totale vengono utilizzati come materiale di partenza) utilizzando i reagenti contenuti nel kit Affymetrix One-Cycle labeling kit and Control reagents. In questa fase vengono aggiunti i controlli di RNA messaggero batterico che serviranno per appurare l'efficienza di sintesi del cRNA. Il ds-DNA così ottenuto viene purificato dai residui enzimatici delle reazioni precedenti e utilizzato come stampo per generare il cRNA marcato con biotina. I reagenti impiegati in queste reazioni sono contenuti nel Affymetrix GeneChip IVT Labeling kit.

L'incubazione dei microarray con cRNA biotinilato e la lettura allo scanner sono state effettuate presso i laboratori del dipartimento di Biochimica Molecolare dell'Università di Modena e Reggio Emilia procedendo nel seguente modo: 15 µg di cRNA biotinilato sono stati frammentati in segmenti lunghi 35-200 bp circa e ibridati sugli array secondo un processo che prevede l'incubazione degli array in un rotore a 45°C per 16 ore (Affymetrix GeneChip Hybridization Oven 640). Dopo le 16 ore i GeneChip® Wheat Genome Array vengono lavati dai residui di cRNA non ibridato e dai cRNA che non hanno un'ibridazione forte (e quindi aspecifici). I cRNA biotinilati fissati al microarray vengono coniugati con la streptavidina che porta un colorante (ficoeritrina) in grado di essere letto da uno scanner. Questi ultimi due passaggi sono stati eseguiti attraverso l'utilizzo di una stazione fluidica (Affymetrix Fluidics Station 450). La lettura dei microarray è stata condotta con il GeneChip Scanner 3000.

Analisi e processing dei dati microarray

Per assicurare la qualità delle ibridazioni effettuate sono state analizzate sia le immagini dei microarray che i file riassuntivi output del software GCOS, che gestisce il GeneChip Scanner 3000. Sono stati valutati anche i parametri definiti da Affymetrix come fondamentali per testare la qualità delle ibridazioni: le performance degli OligoB2, la media dei background, i controlli

sulla sintesi del cRNA, i controlli sull'ibridazione e i controlli sui geni housekeeping.

I dati grezzi sono stati normalizzati primariamente con algoritmo RMA (Irizarry et al., 2003) utilizzando il software R (<http://www.R-project.org>) e la libreria "affy" (Irizarry et al., 2006; Gautier et al., 2004). La stessa libreria è stata utilizzata per eseguire l'algoritmo di normalizzazione, calcolo del segnale e delle "detection call" MAS 5.0 (Affymetrix, 2001). Le detection call ottenute (Present, Absent o Marginal) sono state utilizzate per applicare un filtro preliminare ai probe set da analizzare, poiché i probe set dichiarati Absent in questo algoritmo sono definiti come non analizzabili e quindi costituiscono rumore di fondo nell'analisi dei dati. Il principio applicato per rimuovere tali probe set è stato quello di dichiarare Present solo quei probe set che presentano almeno tre call Present in tutte e tre le repliche di un campione.

Utilizzando i dati in scala logaritmica (in base 2) derivati dalla normalizzazione RMA sono stati calcolati i coefficienti di correlazione lineare R^2 per ciascuna coppia possibile fra le tre repliche biologiche dei nove campioni.

I dati normalizzati RMA sono stati importati nell'ambiente di lavoro GeneSpring GX 7.3 (Agilent Technologies, Santa Clara, CA) e tutte le analisi successive di processamento dei dati sono state condotte con questo software.

Sono stati eseguiti tre confronti per ognuno dei tre genotipi presi in considerazione in questo esperimento: CTRL ∇ MS, CTRL ∇ SS and MS ∇ SS.

Per identificare geni differenzialmente espressi, sono stati applicati due indici di significatività: uno statistico e uno biologico (Chen et al., 2007). Statisticamente i dati sono stati analizzati per mezzo di un Welch t-test con correzione di Benjamini e Hochberg del False Discovery Rate per i test multipli (Benjamini and Hochberg 1995). Sono stati considerati statisticamente rilevanti probe set che hanno mostrato valori di p-value inferiori a 0.05. Dal punto di vista della significatività biologica sono stati considerati differenzialmente espressi solo i probe set con livelli di espressione due volte maggiori o inferiori rispetto al confronto.

Nel disegno sperimentale di questo lavoro vi sono due variabili principali: il genotipo e il grado di stress idrico. Per identificare le componenti effettive che hanno determinato la varianza fra i campioni in esame è stata condotta una Principal Component Analysis (PCA, Yeung e Ruzzo 2001).

Successivamente è stata eseguita una analisi QT-clustering (Heyer et al., 1999) allo scopo di raggruppare geni con andamenti d'espressione simili. Il QT (Quality Threshold) clustering è un algoritmo che raggruppa i geni in base a due parametri: il numero minimo di geni necessari per determinare un

cluster e la correlazione minima che deve sussistere tra elementi dello stesso cluster. In questo lavoro è stata impostata una dimensiona minima di 30 geni e una correlazione di almeno 0.75 (correlazione di Pearson).

Una volta ottenuti i cluster, per determinare se certe categorie funzionali fossero maggiormente rappresentate in un cluster rispetto alla frequenza osservata nell'intero genoma, è stata sfruttata l'informazione disponibile nel "MIPS Arabidopsis thaliana Database" (MAtdB) (mips.gsf.de/projects/funcat) e lo strumento "Search for Functional Distribution of Gene Lists" (Ruepp et al., 2004). MIPS FunCat è un database gerarchico che mette in relazione i locus identifier di Arabidopsis con la relativa categoria funzionale. MIPS FunCat contiene 28 categorie funzionali principali suddivise in 1289 sottocategorie. Come soglia di significatività, nel considerare una categoria sovra-rappresentata rispetto la sua normale distribuzione nel genoma, si è considerato un p-value inferiore a 0.005.

Sulle sequenze rappresentate dai probe set del GeneChip* Wheat Genome Array è stata condotta un'analisi BLAST con il software HarvEST: Affymetrix Wheat1 Chip 1.50 (<http://www.harvest.unc.edu>) al fine di annotare i probe set sulla base dell'omologia di sequenza. Per determinare una soglia di significatività è stata presa in considerazione l'E-score dell'analisi BLAST. Similarità di sequenze con E-score maggiori di 10^{-10} non sono stati considerati idonei.

qRT-PCR per la validazione dei dati microarray

Per validare i dati microarray sono stati selezionati 21 geni dell'array (tab. 3) in base alla funzione genica (appartenenza a pathway biosintetici coinvolti nello stress idrico) e al livello di espressione (appartenenza a cluster di interesse). È stata condotta una quantificazione dei livelli di trascritto tramite real-time PCR.

La trascrizione inversa per sintetizzare il cDNA da utilizzare nelle reazioni di quantificazione in qRT-PCR è stata eseguita utilizzando il protocollo standard e i reagenti del kit Super ScriptTM II Reverse Transcriptase (Invitrogen). 4 µg di RNA totale, i dNTP e i primer oligo (dT)₁₈ sono stati incubati a 65°C per 5 minuti. A questa mix di reazione sono stati aggiunti il Super ScriptTM II RT Reaction Buffer e il DTT e il tutto incubato a 42°C per 2 minuti. Dopo l'aggiunta della Super ScriptTM II Reverse Transcriptase si incuba nuovamente a 42°C per 50 minuti e successivamente si riscalda a 70°C per 15 minuti allo scopo di inattivare la trascrittasi inversa.

Il cDNA ottenuto è stato quantificato utilizzando il Fluorometro Qubit™ (Invitrogen), diluito, portato a concentrazioni uguali tra i campioni e utilizzato per le reazioni di qRT-PCR.

La qRT-PCR è stata eseguita sfruttando la tecnologia SYBR Green e utilizzando il termociclatore per real-time PCR ABI PRISM 7300 Sequence Detection System (Applied Biosystems). La reazione di qRT-PCR è stata condotta su 1 ng di cDNA, 10 µl di SYBR Green PCR Master Mix (Applied Biosystems) e primer alla concentrazione di 400nM, il tutto in un volume di 25 µl. Le condizioni di PCR utilizzate sono state: 95°C per 10 minuti, 40 cicli di 95°C per 15 secondi e 60°C per un minuto.

Alla fine della reazione di amplificazione è stata eseguita una analisi per identificare le curve di dissociazione degli amplificati. La normalizzazione è stata condotta con il *reference* gene selezionato secondo i criteri descritti in seguito.

La quantificazione dei risultati ottenuti tramite qRT-PCR è stata determinata con il metodo comparativo $\Delta\Delta C_t$. Questo metodo confronta i valori di C_t del gene di interesse (target) nel campione da esaminare con quelli di un calibratore. I due valori di C_t sono normalizzati secondo i valori di un gene endogeno (reference gene) e paragonati a quelli del calibratore. Tale metodo è noto come metodo $2^{-\Delta\Delta C_t}$, dove:

$$\begin{aligned}\Delta\Delta C_t &= \Delta C_{t_{\text{trattamento}}} - \Delta C_{t_{\text{calibratore}}}; \\ \Delta C_{t_{\text{campione}}} &= (C_{t_{\text{target}}} - C_{t_{\text{reference gene}}})_{\text{trattamento}}; \\ \Delta C_{t_{\text{calibratore}}} &= (C_{t_{\text{target}}} - C_{t_{\text{reference gene}}})_{\text{calibratore}}\end{aligned}$$

Quindi si è trovato il valore del Fold Change per ciascun gene analizzato: $Fold\ Change = 2^{-\Delta\Delta C_t}$.

I dati dei C_t sono espressi come medie di tre repliche sperimentali. Successivamente il *fold change* del dato qRT-PCR è stato comparato con il *fold change* del dato Array, che si ottiene dividendo il valore del dato array di un campione (trattato) per il valore del suo calibratore (controllo).

Selezione di reference gene per analisi qRT-PCR

I geni che hanno un livello di espressione costitutivo (cioè mostrano sempre lo stesso livello di espressione al variare delle condizione e dello stadio di sviluppo) sono ideali per essere utilizzati come geni normalizzatori in una analisi di real-time PCR. Tuttavia, molti dei geni più comunemente utiliz-

zati a tale scopo nelle piante e negli animali, i così detti *housekeeping gene*, non sono sempre così costanti nel livello d'espressione (Thellin et al., 1999; Suzuki et al., 2000; Lee et al., 2001; Czechowski et al., 2005; Faccioli et al., 2007). Per capire la consistenza di espressione di un gene è necessario avere dei dati di espressione di moltissimi geni (teoricamente tutti i geni del genoma) provenienti da condizioni di crescita e tessuti differenti. L'analisi microarray si avvicina molto a queste condizioni ideali e sicuramente permette di identificare i geni che meno variano nei campioni analizzati nell'esperimento.

Per identificare dei *reference gene* con i dati array disponibili sono stati presi in considerazione solo i geni che mostravano *Present call* in tutti i campioni (27) sulla base dell'algoritmo MAS 5.0. 17134 probe set su 61290 hanno superato questo filtro (27.9%). Il valore di espressione di questi geni è stato normalizzato al valore di espressione della mediana per ogni singolo gene nelle 27 ibridazioni. La lista di geni è stata ulteriormente filtrata eliminando i probe set che anche in una sola ibridazione mostravano valore di espressione doppio o la metà rispetto alla sua mediana nei 27 campioni. 4768 hanno superato anche questo filtro. Successivamente, sulla base del coefficiente di variazione $CV = SD/MV$ (Deviazione standard / valore di espressione medio) si è ottenuta una lista dei geni che mostrano un minore livello di variazione. I tre probe set migliori, anche in base al livello d'espressione assoluto, che codificano rispettivamente per un fattore di iniziazione della trascrizione TIF (Ta.9068.1.S1_at $CV=0.052$), una spastina (Ta4093.1.S1_at, $CV=0.055$) e una poliubiquitina (Ta.24299.1.S1_at, $CV=0.062$) sono stati testati in qRT-PCR per stabilire quale fosse il gene effettivamente più idoneo alla funzione di reference gene. In seguito a queste analisi la poliubiquitina si è dimostrata essere la più stabile ed è stata utilizzata in questo lavoro come reference gene.

Mappaggio di geni sulla base del livello di espressione

In questo esperimento sono stati messi a confronto i trascrittomi di tre genotipi che hanno assetti cromosomici particolari e strettamente connessi tra loro: il CS è un frumento esaploide (AA BB DD) e possiede 7 cromosomi per ogni genoma aploide. CS-5AL è una sua linea di delezione ed è priva di una porzione del braccio lungo del cromosoma 5A.

Essendo 42 i cromosomi del frumento tenero, potremmo dire che CS-5AL presenta circa il 95% della struttura cromosomica di un normale fru-

mento esaploide. Creso, invece, frumento duro e quindi tetraploide (AA BB) ha un genoma che è un sottoinsieme di quello del frumento tenero e copre circa il 66% del genoma del frumento tenero. È razionale ipotizzare che i geni che non si esprimono nel frumento duro, ma che si esprimono nel tenero siano posizionati sul genoma D, oppure che si trovino sui genomi A o B e che siano regolati da elementi del genoma D. Per identificare geni con queste caratteristiche si è proceduto nel seguente modo.

È stato applicato un primo filtro stringente sulla base delle detection call: per superare tale filtro un probe set deve mostrare 18 Present call su tutte le 18 repliche dei campioni di frumento tenero, e allo stesso tempo mostrare 9 Absent call su tutte le 9 repliche di Creso (frumento duro). La lista di probe set ottenuti è stata nuovamente filtrata in base al livello di espressione che doveva risultare inferiore al livello di background in tutte le repliche del frumento duro, e tre volte maggiore del background in tutti i frumenti teneri.

Con lo stesso identico criterio applicato alle 9 repliche del genotipo CS-5AL e al suo wild-type CS, sono stati identificati geni putativamente posizionati sul braccio lungo del cromosoma 5A.

Una volta identificate le liste di geni con queste caratteristiche si è proceduto alla verifica della posizione genomica attraverso PCR su DNA genomico. Tali analisi sono state effettuate in collaborazione con il Centro ricerche dell'ENEA di Casaccia. In questa fase è introdotta una ulteriore varietà di frumento duro (Trinakria), per escludere che l'effetto osservato fosse varietà-specifico e non specie-specifico.

In base alle sequenze di interesse e alle loro omologie con altre sequenze di frumento, sono state disegnate coppie di primer specifiche e altre comuni alle sequenze omologhe. Quindi, mentre i primer comuni a più sequenze dovrebbero amplificare sia in frumento duro che tenero (controllo positivo), i primer specifici dovrebbero amplificare solo la sequenza individuata come espressa solo nei frumenti teneri e di conseguenza non mostrare nessuna banda in Creso e Trinakria.

Si sono testate coppie di primer per 70 geni putativamente mappati sul genoma D e 23 geni putativamente appartenenti al braccio lungo del cromosoma 5A.

Per verificare se l'ipotesi nulla è compatibile con i dati ottenuti, cioè che il gene selezionato appartenga al genoma D o al braccio lungo del cromosoma 5A per cause probabilistiche (circa il 33% dei geni di frumento tenero è sul genoma D e circa il 5% è sulla parte deleta del braccio lungo del 5A), è stato eseguito un test chi quadrato.

RISULTATI E DISCUSSIONE

Analisi dei parametri di qualità dell'ibridazione

La qualità delle immagini ottenute dalle 27 ibridazioni è stata messa in evidenza usando i controlli oligoB2. Tutti hanno mostrato le strutture a scacchiera attese agli angoli degli array e nessuna immagine aberrante è stata evidenziata.

Sono stati effettuati anche i controlli sui geni PolyA di origine batterica (*lys, phe, thr, dap*) e sui controlli di ibridazione (*BioB, BioC, BioD and Cre*). Tutti hanno mostrato valori compresi nei parametri forniti da Affymetrix.

Il valore medio di background osservato è stato di 39.96, ampiamente nei parametri definiti dal produttore dei microarray.

Le percentuali delle "Present call" osservata oscillava tra il 42.35% e il 51.78% dei 61mila probe set dell'array. Si è osservato, inoltre che la percentuale delle "Present call" nei genotipi di frumento duro era costantemente inferiore rispetto a quella dei frumenti teneri di almeno il 7%. Questo può essere spiegato sulla base dell'algoritmo che porta alle detection call. Infatti se un gene non è espresso nel campione di interesse il segnale sul PM probe cell sarà simile a quello del background che sarà simile a quello del MM probe cell. Di conseguenza l'algoritmo MAS 5.0 definirà tale probe set come "Absent" probe set. Visto che i frumenti duri hanno un genoma in meno rispetto ai frumenti teneri è razionale pensare che anche il numero di geni che si esprimono in una cellula sia inferiore. Il 7% di differenza costante nelle "detection call" corrisponde a circa 4200 trascritti di differenza presenti nei frumenti teneri e assenti nei frumenti duri.

L'analisi preliminare dei dati e il confronto dei parametri di qualità hanno evidenziato l'ottima qualità delle ibridazioni effettuate.

Qualità delle repliche biologiche

Per determinare la qualità delle repliche biologiche sono stati calcolati i coefficienti di correlazione lineare R^2 sui valori di espressione in scala logaritmica normalizzati con algoritmo RMA. Tutti i coefficienti oscillano tra il valore 0.93 e 0.99 indicando un alto livello di correlazione fra repliche.

L'analisi delle componenti principali della varianza ha messo in evidenza quattro componenti principali che nella totalità spiegano il 75.43% della variazione osservata nelle 27 ibridazioni (fig. 1). In particolare la componente

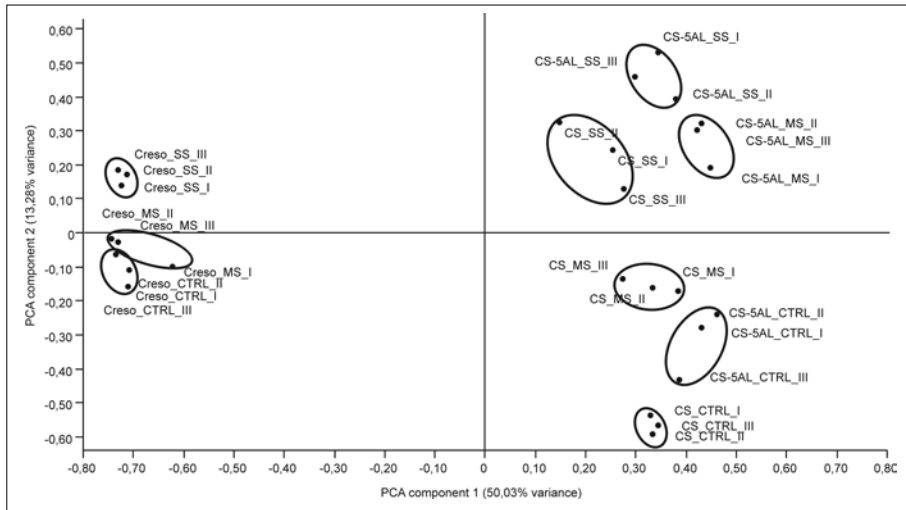


Fig. 1 Scatter plot delle repliche biologiche dei campioni in funzione delle due componenti principali della varianza ottenute in seguito a PCA. Gli ovali rappresentano i raggruppamenti in base alla replica biologica

primaria spiega ben il 50% della varianza. Ponendo su di un grafico cartesiano tale componente sull'asse x, si osserva che i campioni si separano in base alla specie a cui appartengono. I frumenti teneri si dispongono tutti alla destra del grafico, sulla sinistra i frumenti duri. Si deduce che la componente primaria corrisponde alla specie botanica di appartenenza, anche perché il CS e la sua linea di delezione non si separano sulla base di questa componente. Questo dato, ampiamente atteso, conferma che i dati microarray sono di buona qualità.

Un'altra evidenza sulla bontà dei dati ottenuti deriva dall'osservazione che le tre repliche per ciascun campione vengono raggruppate insieme nel grafico funzione delle due componenti principali, come è possibile osservare in figura 1 dove gli ovali racchiudono repliche dello stesso campione.

La componente secondaria, rappresentata dall'asse y (fig. 1) corrisponde alla variabile stress idrico. Infatti sull'asse y si separano i campioni sulla base dello stress imposto. In basso si ritrovano i campioni CTRL, in posizione intermedia i MS e in alto i SS relativi allo stesso genotipo. Da notare come la distanza tra CTRL e MS in Creso sia molto ridotta, suggerendo che dal punto di vista del trascrittoma non ci sono grosse differenze tra questi due campioni. Situazione opposta in CS-5AL dove i campioni dello stress intenso e moderato sono molto più vicini rispetto a quelli del controllo.

Confronto dei profili trascrizionali

Per identificare geni differenzialmente espressi fra i differenti genotipi o fra campioni sottoposti a crescenti livelli di stress idrico, i dati di espressione sono stati analizzati per mezzo di un Welch t-test con correzione di Benjamini e Hochberg per i test multipli (Benjamini and Hochberg, 1995). I geni identificati come differenzialmente espressi sono stati successivamente filtrati in base a un 2-fold change come parametro di significatività biologica (Chen et al., 2007), come descritto nel paragrafo sui materiali e metodi.

Come riportato in tabella 2, sono stati eseguiti nove confronti e sono stati identificati 3056 (non ridondanti) probe set differenzialmente espressi in almeno un confronto fra i diversi livelli di stress. Tre ulteriori confronti sono stati effettuati tra i tre genotipi nelle condizioni di controllo. Ben 8552 probe set sono risultati differenzialmente espressi di cui solo 607 nel confronto tra CS e la sua linea di delezione. Questi numeri testimoniano le differenze in termini di trascrittoma tra le due specie e tra CS e Cs-5AL, differenze attribuibili in larga parte alla diversa costituzione genomica (tab. 2). Il numero di geni con livello di espressione maggiore in CS rispetto a Creso è sensibilmente maggiore (circa 1300 probe set) rispetto a quelli che mostrano livello di espressione inferiore.

CONFRONTI	PROBESETS 2X UP-REGULATED	PROBESETS 2X DOWN-REGULATED	TOTALE DEI PROBE SET DIF- FERENZIALMENTE ESPRESSI
Creso			
MS Vs CTRL	0	0	0
SS Vs MS	191	302	493
SS Vs CTRL	661	809	1470
CS			
MS Vs CTRL	54	52	106
SS Vs MS	29	0	29
SS Vs CTRL	579	263	842
CS-5AL			
MS Vs CTRL	287	106	393
SS Vs MS	47	0	47
SS Vs CTRL	660	527	1187
CTRL			
CS Vs Creso	3974	2662	6636
CS-5AL Vs Creso	3889	3062	6951
CS-5AL Vs CS	226	381	607

Tab. 2 *Riassunto dei confronti eseguiti. Sono riportati il numero di geni che mostrano livelli di espressione 2 volte maggiori o minori rispetto al confronto*

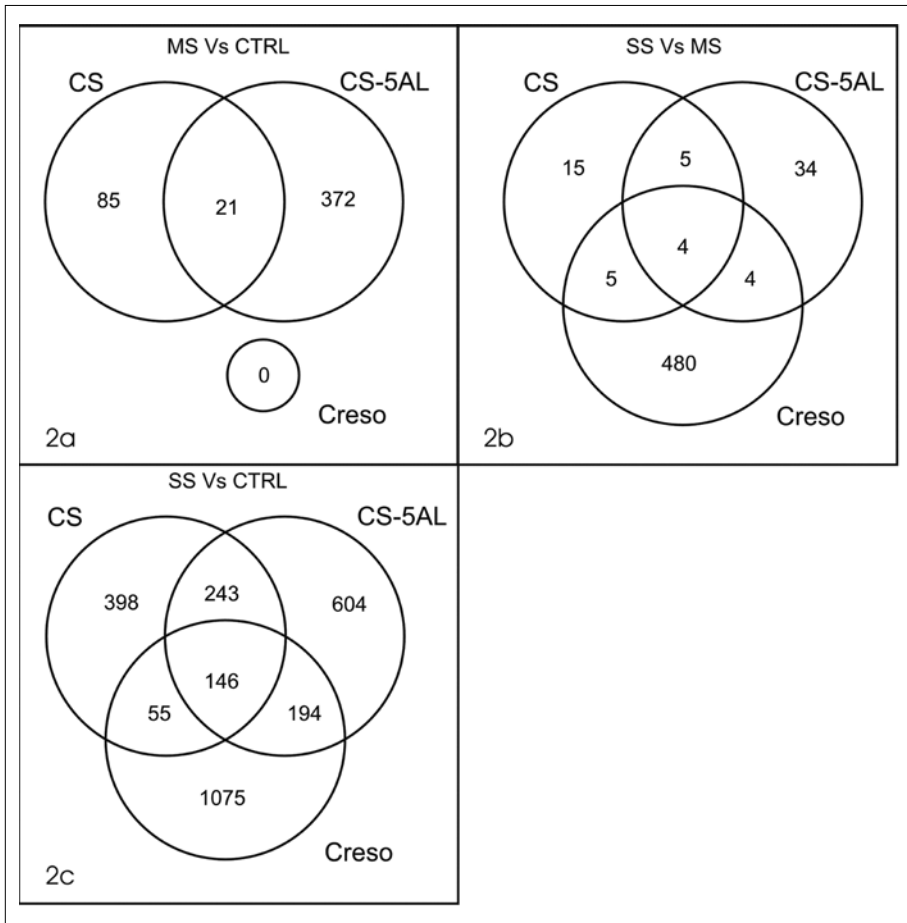


Fig. 2 Diagramma di Venn sulle liste di geni differenzialmente espressi nei confronti fra i diversi stress imposti. 2a) Stress moderato (MS) contro controllo (CTRL). 2b) Stress intenso (SS) contro moderato. 2c) Stress intenso contro controllo

Lo stress moderato comporta l'induzione dell'espressione di 106 probe set in CS e 393 in CS-5AL, mentre in Creso non si osservano differenze tra il controllo e lo stress moderato. Questo dato era già stato intuito dall'analisi delle componenti principali (fig. 1) e sottolinea come Creso necessiti di stress idrici maggiori per attivare meccanismi di regolazione dell'espressione genica rispetto al frumento tenero. Tale fenomeno può essere interpretato in due maniere differenti: a) Creso non attiva risposte trascrizionali perché lo stress subito non è tale da creare danni alla pianta; b) La risposta di Creso è tardiva rispetto a quella del frumento tenero.

Situazione opposta si verifica quando il trascrittoma in condizioni di stress moderato viene confrontato con quello in condizioni di stress severo. I frumenti teneri mostrano solo 29, in CS, e 47, in CS-5AL, geni differenzialmente espressi, mentre in Creso si osserva la differente espressione di 493 probe set (tab. 2).

I confronti fra lo stress intenso e il controllo mostrano un ampio numero di geni differenzialmente espressi, più in Creso (1470) che in CS (842).

Per identificare geni che si attivano o reprimono in seguito a carenza d'acqua in tutti e tre i genotipi, in uno solo di essi, o fra le diverse combinazioni sono stati prodotti dei diagrammi di Venn (fig. 2). In tutte e tre le rappresentazioni, corrispondenti ai tre confronti possibili, sono stati identificati pochissimi probe set comuni fra i tre genotipi. Solo 21 su 478 sono comuni fra CS e CS-5AL nel confronto tra stress moderato e controllo (fig. 2a).

Solo 4 su 547 nel confronto tra stress intenso e moderato (fig. 2b). 146 su 2715 nel confronto tra stress intenso e controllo. Mentre i dati relativi ai primi due confronti non sono molto informativi, causa il ridotto numero di geni provenienti ora da Creso (MS Vs CTRL), ora dai frumenti teneri (SS Vs MS), risulta molto sorprendente il dato proveniente dall'ultimo confronto. Solo il 5% dei geni totali che vengono attivati o repressi nella transizione dal controllo allo stress intenso risultano in comune a tutti i genotipi.

Data mining

Complessivamente 3056 probe set sono stati identificati come differenzialmente espressi in seguito a stress idrico. Tentare di estrapolare dati significativi da un così elevato numero di geni è impresa assai difficile se non si prova a frammentare il dato in porzioni più piccole e facili da analizzare.

Il data mining è, appunto, quel processo che permette di ottenere informazioni biologicamente rilevanti da una serie di liste di geni e valori di espressione. A tal scopo si è deciso di creare dei sottogruppi di geni ("cluster") accomunati da comportamenti di espressione simile. Si è opzionato l'utilizzo di un algoritmo di clustering di tipo "unsupervised" poiché non si disponeva di alcuna indicazione preliminare sul numero di gruppi attesi. Il QT-clustering (Hayer et al., 1999) permette di generare n gruppi di geni sulla base di due parametri soglia definiti dall'utente. I parametri definibili

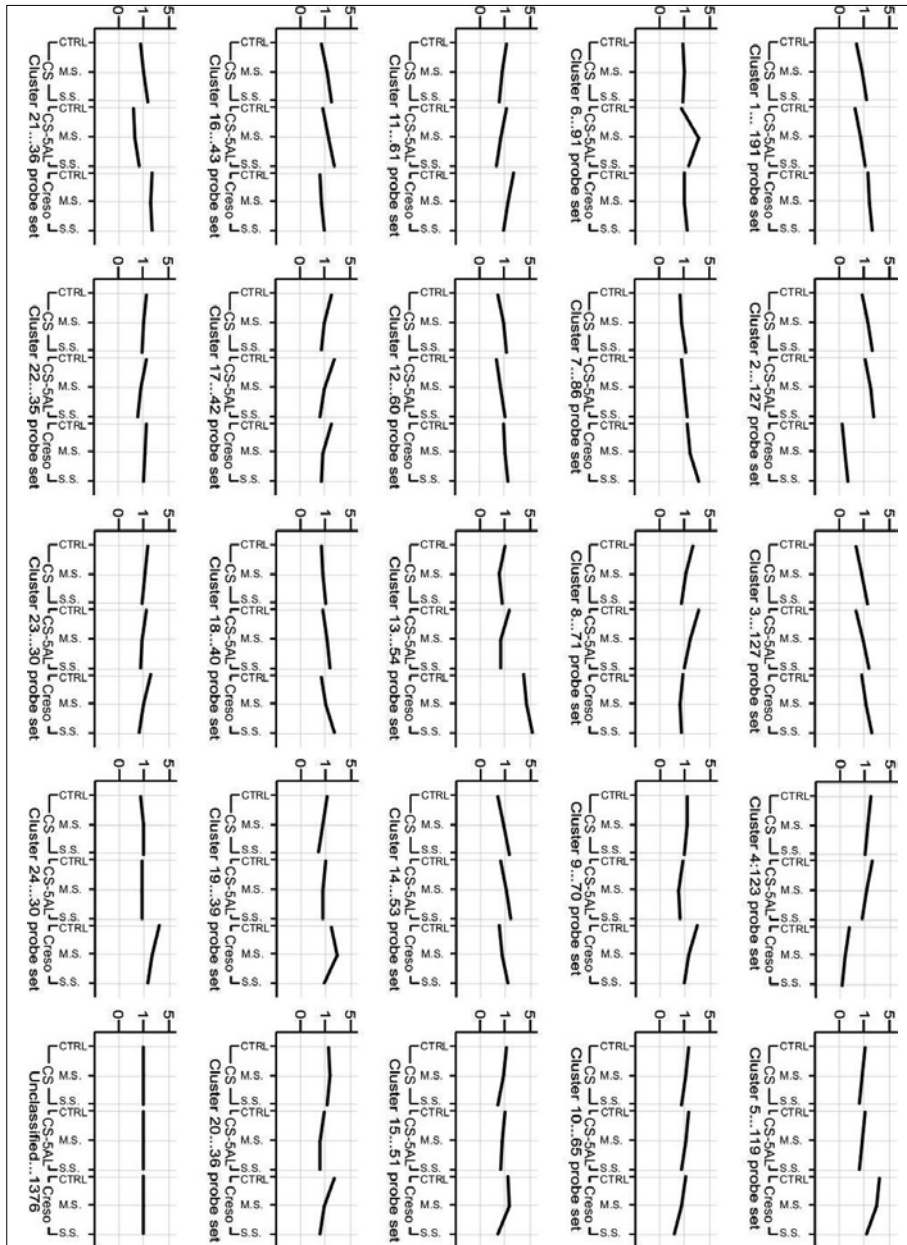


Fig. 3 Rappresentazione dei 24 cluster ottenuti con algoritmo QT-clustering. Sull'asse x sono rappresentati i tre genotipi, che raggruppano i relativi CTRL, MS e SS. Sull'asse y sono rappresentati i valori di espressione inteso come rapporto tra il valore d'espressione del gene e la mediana su tutti i campioni. L'andamento del cluster è dato dal valore di espressione medio di tutti i geni costituenti il cluster

sono la correlazione minima (correlazione di Pearson) tra gli elementi del cluster e il numero minimo di elementi per generare un cluster. Diminuire la soglia di correlazione comporta l'ottenimento di un numero minore di cluster totali, ma di dimensioni maggiori; mentre il numero di geni che risulteranno non raggruppati in alcun cluster diminuisce. Tutto questo a discapito della stringenza dell'analisi: cioè anche geni con un comportamento d'espressione non perfettamente identico potranno essere raggruppati insieme. Il parametro "numero minimo di geni per cluster" evita che vengano creati cluster di dimensioni troppo piccole e quindi poco significativi nel quadro generale. Aumentare questo parametro comporta l'eliminazione di cluster con pochi geni che di conseguenza arricchiranno il gruppo dei geni non raggruppati.

Sono stati testati differenti valori dei parametri soglia: il miglior compromesso tra numero di cluster ottenuti, geni raggruppati e geni non raggruppati è stato ottenuto impostando la correlazione di Pearson a 0.75 e il numero minimo di geni a 30.

Sono stati così individuati 24 gruppi di geni, più 1376 probe set che non sono stati raggruppati in alcun cluster (fig. 3).

In base all'andamento del livello di espressione ogni singolo cluster può essere messo in relazione con una particolare risposta della pianta allo stress idrico. Vi sono, per esempio, geni che vengono attivati allo stesso modo in tutti e tre i genotipi (come il cluster 14) e che quindi rappresentano i geni della risposta comune tra frumento tenero e duro. Analizzando inoltre le classi funzionali maggiormente rappresentate in questi cluster è possibile risalire alle funzioni geniche conservate tra le due specie in risposta a stress idrico. Di seguito sono descritti i cluster ottenuti e i possibili processi biologici cui partecipano.

Geni modulati similmente nei tre genotipi

L'unico cluster che mostra un chiaro andamento di questo tipo, cioè con geni indotti allo stesso modo nei tre genotipi, è il cluster 14. Anche il cluster 3 mostra un andamento paragonabile a quello del cluster 14, tuttavia i livelli d'espressione dei geni di frumento duro sono mediamente maggiori rispetto a quelli dei frumenti teneri (fig. 4). Per semplicità tratteremo questi due cluster come un cluster unico.

I geni del cluster 10 sono anche essi modulati allo stesso modo nei tre genotipi e repressi all'aumentare dello stress idrico (fig. 5).

L'analisi delle categorie funzionali sui 180 geni di questi due cluster ha messo in evidenza particolari classi funzionali sovra-rappresentate rispetto alla

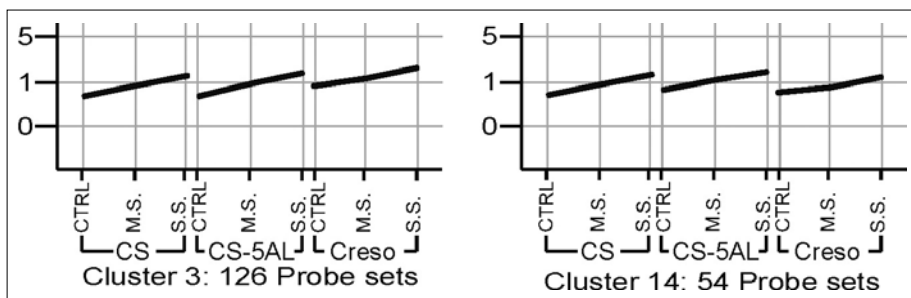


Fig. 4 Cluster 3 e 14

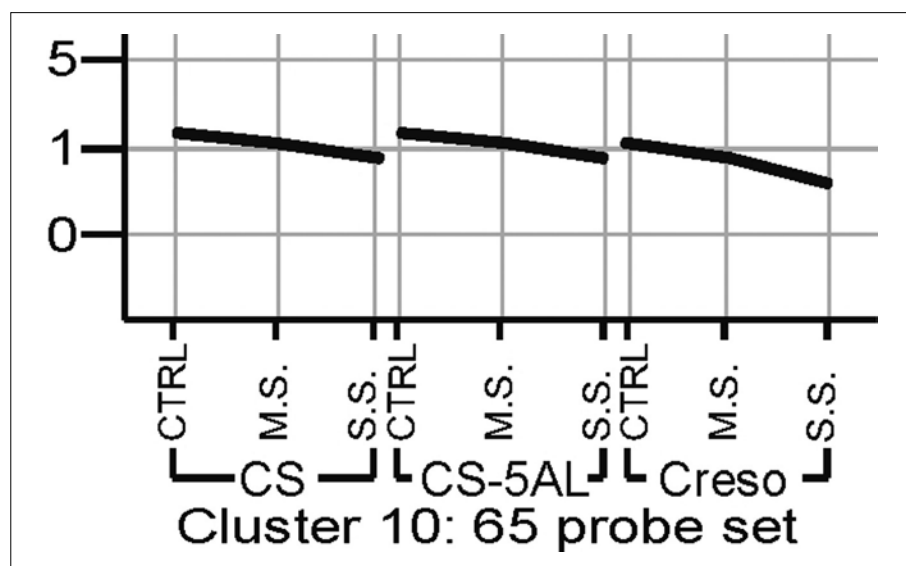


Fig. 5 Cluster 10

frequenza media osservata nel genoma. Alcune di queste classi sono coinvolte nel metabolismo primario, produzione di energia, trasporto cellulare e interazione con l'ambiente. Fra le sottocategorie di quest'ultima, vi sono la categoria "percezione e risposta allo stress idrico", "risposta all'acido abscissico" e "metabolismo della prolina".

Nello specifico appartengono a questo cluster molte delle componenti molecolari già note in letteratura per il loro ruolo nella risposta allo stress idrico: vi sono due probe set (Ta.1981.1.A1_a_at = chinasi, TaAffx.77541.1.S1_at) che rappresentano geni i cui prodotti sono delle chinasi coinvolte nella trasduzione del segnale. Vi sono, inoltre, ben quattro probe set (Ta.12813.1.S1_x_at,

Ta.12813.2.S1_x_at, TaAffx.13292.1.S1_at, TaAffx.13292.1.S1_s_at) con elevata similarità di sequenza con il gene per la NCED (9-cisepossicarotenoido diossigenasi), l'enzima chiave del pathway di biosintesi dell'ormone ABA (Seo and Koshiba 2002). Di questo enzima si parlerà nel paragrafo dedicato al pathway di sintesi dell'ABA.

Appartiene a questo cluster anche il probe set TaAffx.108538.1.S1_at codificante per il fattore di trascrizione HB-7 (homeobox-leucine zipper). L'espressione genica di HB-7, in *Arabidopsis*, è indotta da stress idrico e da trattamenti con ABA esogeno e rappresenta un fattore di trascrizione stress idrico-inducibile ABA dipendente (Soderman et al., 1996; Shinozaki and Yamaguchi-Shinozaki, 2007). HB-7 si colloca a monte della trascrizione di RD20 (Resistant to Dehydration), una calcium binding protein che è indotta da stress idrico, salino e ABA esogeno (Takahashi et al., 2000). Un probe set (Ta.9830.2.S1_at) appartenente al cluster 3-14 ha elevata similarità con la sequenza di RD20 di *Arabidopsis*. Questi tre geni (NCED (ABA) → HB-7 → RD20) identificati nei cluster 3 e 14 hanno permesso di ricostruire in frumento una delle risposte allo stress idrico/salino ABA dipendente già descritta in *Arabidopsis*.

Altri geni codificanti fattori di trascrizione noti per il loro coinvolgimento nella risposta allo stress idrico sono stati identificati nel cluster 3-14: otto probe set di geni MYB, un WRKY, un DOF (DNA binding with One Finger) e tre probe set per fattori di trascrizione della famiglia con motivo Zinc Finger.

Infine vi sono diversi probe set con similarità di sequenza con geni effettori della risposta allo stress idrico: uno per il gene della sorbitolo deidrogenasi (Ta.21428.1.S1_x_at), enzima che catalizza la reazione che converte il glucosio-6P in sorbitolo-6P, passaggio chiave della sintesi del sorbitolo (Duanggraisai et al., 2007). Il sorbitolo è un polialcole il cui accumulo in seguito a stress idrico è stato osservato in molte piante e la sua funzione, oltre che come osmolita compatibile, sembra essere quella di difesa dalle specie reattive dell'ossigeno (ROS) (Lo Bianco et al., 2000). Anche il gene chiave che porta all'accumulo di prolina a partire dal glutammato (P5CS, Δ -pirrolina 5 carbossilato sintasi = Ta.7091.1.S1_at) è stato identificato nel cluster 3-14. La sua discussione è rimandata al paragrafo appositamente dedicato al pathway della prolina.

Altri geni noti nella risposta allo stress idrico e appartenenti a questo gruppo di geni sono l'aldeide deidrogenasi (Ta.18775.1.S1_at, Ta.25596.3.A1_a_at) e la betaina-aldeide deidrogenasi (Ta.435.1.S1_at). Gli enzimi della classe delle aldeidi deidrogenasi sono state caratterizzate in uomo e sono coinvolte nella detossificazione delle aldeidi. In pianta sono state identificate diverse aldeidi deidrogenasi (ALDH) indotte da stress idrico; in particolare l'ALDH substrato specifica per la betaina-aldeide catalizza l'ossidazione verso l'osmoli-

ta compatibile glicina-betaina (Chen e Murata 2002). Una ALDH inducibile con ABA, nonché da stress idrico, è stata isolata dalla “resurrection plant” *C. plantagineum* ed è probabilmente coinvolta nella detossificazione delle aldeidi derivanti dalla perossidazione dei lipidi (Kirch et al., 2005).

Infine appartengono al cluster 3-14 anche due probe set relativi a proteine “early-responsive to dehydration” (ERD).

Il quadro complessivo che prende forma dallo studio dei cluster 3 e 14 è che questi 180 probe set, sia per il loro livello di espressione che viene indotto all’aumentare dello stress idrico, sia per il fatto che molti di loro hanno una funzione nota nella risposta allo stress idrico già descritte in altre specie, possono essere definiti come geni che costituiscono la base conservata della risposta a stress idrico in molte piante. Inoltre visto che tutti questi geni sono coregolati con l’enzima chiave della biosintesi dell’ABA è probabile che molti di loro siano a valle della produzione di questo ormone.

Tra i geni dei cluster 3 e 14 vi sono 37 probe set non ancora annotati per i quali questo lavoro fornisce una prima importante annotazione funzionale,

Il cluster 10, invece, mostra lo stesso andamento nei tre genotipi, ma il livello d’espressione diminuisce all’aumentare dello stress (fig. 5). Nessuna categoria funzionale è sovra rappresentata in questo cluster. Anche l’indagine sui singoli trascritti del cluster 10 non ha messo in evidenza geni noti nel coinvolgimento nella risposta a stress idrico, se non geni per la sintesi di proteine coinvolte in generali pathway di trasduzione del segnale (chinasi, calmoduline).

In totale i geni dei cluster che mostrano stesso andamento in risposta allo stress idrico nei tre genotipi sono 245, su un totale di 3056 geni modulati. Quindi circa l’8% dei geni che si attivano in almeno un confronto mostrano lo stesso andamento nei tre genotipi. Questo dato, evidenziato anche con l’uso dei diagrammi di Venn, sottolinea che i meccanismi messi in atto da frumento tenero e duro sono, in termini di sequenze indotte, molto differenti tra loro.

Molti dei geni osservati in questi cluster sono geni tipici della risposta allo stress idrico e molti di essi sono stati descritti in specie filogeneticamente distanti come *Arabidopsis* indicando che queste risposte sono conservate dal punto di vista evolutivo.

Probe set la cui espressione è costitutivamente più elevata in Creso

Il Cluster 1 (fig. 6) è il più grande fra i 24 cluster individuati ed è composto di 191 probe set. I geni del cluster 1 presentano un andamento differente

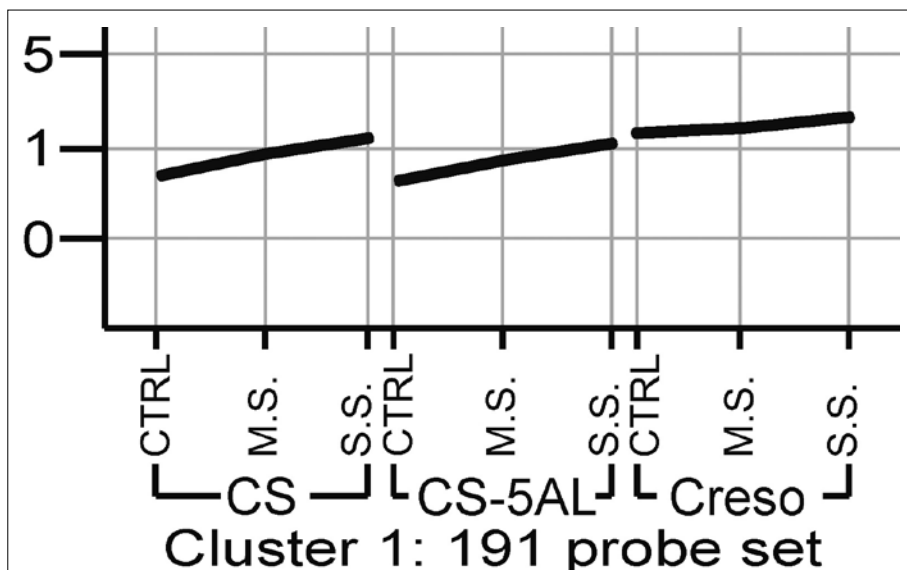


Fig. 6 Cluster 1

tra i campioni di frumento tenero e quelli di frumento duro. Essi sono infatti caratterizzati da un aumento nel livello di espressione all'aumentare dello stress idrico imposto e da un livello di espressione più elevato in Creso rispetto a CS. Inoltre nei frumenti teneri il livello d'espressione è generalmente inferiore (in particolare nei controlli), ma il livello di induzione è più elevato.

Le categorie sovra rappresentate in questo cluster riguardano il metabolismo primario e la produzione d'energia, in particolar modo i meccanismi di tipo fotosintetico.

In *Arabidopsis* è stato osservato che stress idrico e piante transgeniche per la sovra espressione di SRK2C (una proteina chinasi coinvolta nella tolleranza allo stress idrico) portano all'aumento dell'attività trascrizionale per geni codificanti proteine coinvolte nei sistemi fotosintetici, soprattutto proteine che si legano alla clorofilla (Umezawa et al., 2004). In questo cluster vi sono 24 probe set relativi a proteine cloroplastiche, di cui 16 sono annotate come "chlorophyll A-B binding protein". Il complesso che cattura la luce (LHC, Light-harvesting complex) è composto da clorofille di tipo A e B e da proteine che si legano a queste clorofille. I sistemi LHC catturano la luce e trasportano l'energia di eccitazione ai fotosistemi I e II ai quali sono strettamente associati. In seguito a cambiamenti delle condizioni di luminosità, la fosforilazione reversibile delle "chlorophyll A-B binding protein" dell'LHC

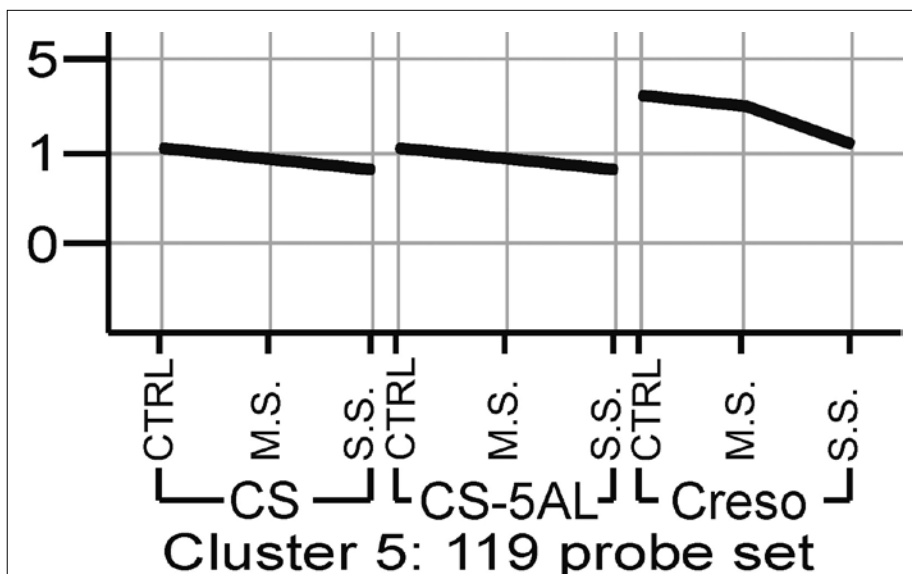


Fig. 7 Cluster 5

permette la mobilitazione del sistema antenna per il bilanciamento dell'energia fra i due fotosistemi (Liu e Shen, 2004). Un loro aumento nei livelli di espressione potrebbe essere quindi necessario per la stabilizzazione dei fotosistemi I e II che, a causa della ridotta disponibilità d'acqua, potrebbero subire delle interruzioni nel processo di accumulo dell'energia e quindi andare incontro a stress ossidativo.

Anche fattori di trascrizione di tipo MYB, bZIP e zinc finger appartengono a questo cluster. Solo quattro probe set Ta.2638.1.S1_at, Ta.29352.1.S1_at (RAB18, Responsive to ABA), Ta.28209.3.S1_at (RD22) e Ta.6973.1.A1_at (ERD, Early Responsive to Dehydration) sono noti per essere correlati con la risposta allo stress idrico. Nonostante ciò il tipo di risposta caratterizzante questo cluster, più che una risposta diretta alla carenza d'acqua, sembra essere una riorganizzazione del metabolismo dei carboidrati e della fotosintesi. Inoltre sembrerebbe che la risposta mediata dall'induzione dell'espressione di questi geni non sia necessaria nel frumento duro poiché il livello d'espressione è costitutivamente più elevato.

Anche i geni del cluster 5 risultano costitutivamente più espressi nel Creso rispetto ai frumenti teneri. In questo caso, però, il livello di espressione di Creso diminuisce all'aumentare dello stress idrico sino ad arrivare ai valori d'espressione riscontrati nei frumenti teneri nei tre trattamenti (fig. 7). In questo cluster solo categorie appartenenti al metabolismo primario sono

sopra rappresentate e anche la ricerca manuale per proteine note nel coinvolgimento nella risposta allo stress idrico non ha prodotto risultati di rilievo.

Probe set la cui espressione è costitutivamente più elevata nei frumenti teneri

Al cluster 2 e al cluster 4 appartengono quei probe set il cui livello di espressione in Creso è molto minore, quando non addirittura assente, rispetto ai frumenti teneri (fig. 8). Mentre nel cluster 2 i geni mostrano una induzione in risposta allo stress idrico, quelli del cluster 4 sono repressi.

L'analisi delle classi funzionali sul cluster 2 ha messo in evidenza che classi coinvolte nella regolazione osmotica e nella percezione dello stress idrico sono rappresentate in modo più consistente rispetto alla frequenza con cui si manifestano nel genoma intero. In questo cluster si ritrovano probe set relativi alla deidrina RAB18, a una LEA (Late Embryogenesis Abundant), una USP (Universal Stress Protein) e una "low temperature and salt responsive protein", quindi geni codificanti proteine a valle della risposta agli stress di tipo osmotico. Il probe set Ta.10389.1.S1_at codifica per una aldoso reductasi. Questo enzima in orzo è putativamente coinvolto nella conversione del glucosio in sorbitolo (pathway già incontrato nei cluster 3 e 14) e nella detossificazione da aldeidi e ha una funzione attiva nella tolleranza allo stress idrico (Roncarati et al., 1995). Questo probe set, espresso solo in frumento tenero e non in Creso potrebbe essere regolato da elementi del genoma D. Se anche l'attività enzimatica fosse confermata essere presente solo nei frumenti teneri, si potrebbe concludere che la risposta a stress idrico ottenuta attraverso l'accumulo di sorbitolo ha due vie preferenziali nei frumenti teneri.

Restando nella categoria produzione di soluti compatibili, in questo cluster è presente il probe set TaAffx.3441.1.S1_at che ha un'alta similarità di sequenza

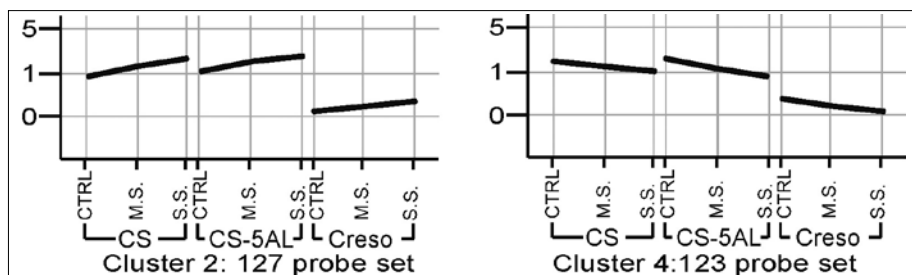


Fig. 8 Cluster 2 e cluster 4

con il gene batterico per l'ornitina ciclodeaminasi, enzima non ancora caratterizzato nelle piante. Questo enzima catalizza la produzione di prolina utilizzando l'aminoacido ornitina come precursore (Trovato et al., 2001). Si rimanda al paragrafo dedicato alla sintesi della prolina per la discussione di questo risultato.

Relativamente ai fattori di trascrizione, un probe set codificante per una proteina della famiglia MYB è presente in questo cluster 2. Un altro TF, HBP-1b (famiglia bZIP) appartiene al cluster 2. Questo fattore di trascrizione agisce legandosi ai promotori dei geni codificanti istoni (Tabata et al., 1991) ed è coinvolto nella trascrizione di geni indotti dal trattamento con acido salicilico (SA, ormone prodotto nella risposta a stress biotici). Tuttavia studi in frumento hanno dimostrato che l'applicazione di SA esogeno porta alla produzione di ABA (Shakirova et al., 2003) e, in orzo, all'accumulo di osmoliti come la prolina (El-Tayeb, 2005). Siccome HBP-1b non è espresso in Creso, probabilmente non è presente sui genomi A e B. Di conseguenza un approccio di genomica funzionale volto al "knock-out" di questo gene sarebbe possibile perché non risentirebbe dei problemi dovuti alla ridondanza genica dei frumenti e potrebbe portare a identificare l'effettivo ruolo di questo TF in frumento tenero.

Infine, in questo cluster è presente il probe set Ta.485.1.A1_at con alta similarità di sequenza con un gene per la lipossigenasi (LOX) enzima coinvolto nella sintesi dell'acido jasmonico (JA). Anche se non indotto da stress idrico l'JA può indurre l'attivazione di fattori di trascrizione MYB a monte della risposta allo stress idrico (Shinozaki and Yamaguchi-Shinozaki, 2007).

I geni del cluster 2 si possono definire, dato il loro livello di espressione, geni indotti da stress idrico e, in base alla funzione di molti geni appartenenti a esso, coinvolti nella risposta di tolleranza. Inoltre, questi geni sono tipici della risposta del frumento tenero, mentre in Creso vengono leggermente indotti o in molti casi non sono espressi affatto. Come detto per i cluster 3 e 14, anche in questo cluster possono essere ricercati geni non ancora descritti come coinvolti nella risposta allo stress idrico.

I geni del cluster 4 sono, di contro, geni che vengono repressi dallo stress idrico e sono coinvolti soprattutto nel trasporto cellulare e nel metabolismo primario che, probabilmente, viene rallentato e riorganizzato in seguito allo stress idrico.

Probe set con andamento caratteristico in Creso

I geni appartenenti ai cluster 15 e 19 hanno un andamento molto particolare: mentre in CS il livello di espressione diminuisce all'aumentare dello stress, in Creso il livello d'espressione aumenta o rimane costante in

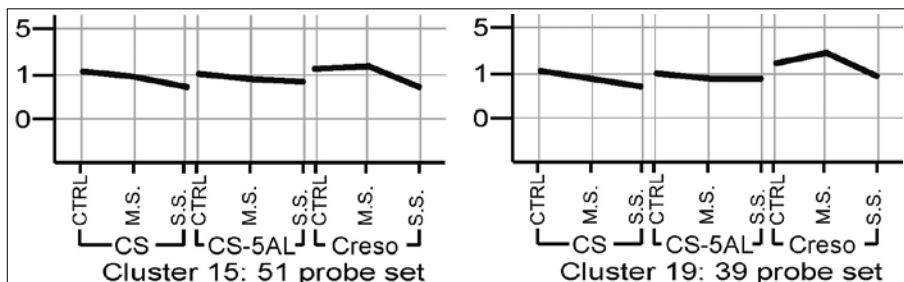


Fig. 9 Cluster 15 e cluster 19

seguito a uno stress idrico moderato, ma diminuisce rispetto al controllo se lo stress è intenso. In CS-5AL rimane pressoché uguale nei tre livelli di stress (fig. 9).

Fra le classi funzionali sopra rappresentate in questi due cluster vi sono microtubuli, elementi del citoscheletro, parete cellulare, risposta agli stress. La sintesi di elementi del citoscheletro ed enzimi per la degradazione/sintesi della parete cellulare viene di norma attivata per favorire la distensione e divisione cellulare. Una delle risposte comuni in tutte le piante è proprio il blocco della distensione e crescita in modo da ridurre le dimensioni delle superfici traspiranti.

Da questi dati si può dedurre che questo meccanismo è evidente anche nei frumenti e che mentre CS inizia a ridurre il livello di espressione di questi geni anche in seguito a uno stress moderato, Creso necessita di uno stress maggiore affinché attivi questo tipo di risposte.

Probe set con andamento caratteristico in CS-5AL

Il cluster 6 rappresenta un gruppo di geni dall'andamento estremamente anormale (fig. 10).

Il livello di espressione dei 91 geni che costituiscono il cluster è costante sia in CS che in Creso. In CS-5AL il livello d'espressione nella condizione di controllo è simile ai livelli dei controlli di CS e Creso. Nello stress moderato i livelli d'espressione aumentano notevolmente, per poi ritornare quasi ai livelli osservati nel controllo quando la pianta è sottoposta a uno stress intenso. I 91 geni di questo cluster sono quasi tutti privi di annotazione, tranne 13 che sono associati a sequenze di trasposoni e retrotrasposoni.

Esistono numerosi studi a supportare l'idea che i trasposoni e i retrotrasposoni siano elementi mobili del genoma la cui trasposizione è attivata come sistema di regolazione epigenetica all'insorgere di stress di diversa natura (biotici e abiotici) come fonte di variabilità genetica (riarrangiamento del genoma) (Rapp e Wendel, 2005). L'esempio più studiato di questo tipo di risposta è l'attività trascrizionale degli elementi *Tam* (elementi trasponibili di *Antirrhinum majus*), trasposoni della classe II sensibili alla temperatura. Coen et al. (1986) hanno caratterizzato tre elementi di petunia (*Tam1*, *Tam2* e *Tam3*), responsabili delle mutazioni del colore dei fiori. In *Medicago sativa* è stata osservata attivazione trascrizionale di copie multiple di un retrotrasposone indotta da freddo (Ivashuta et al., 2002) (Madlung e Comai, 2004). Recentemente è stato proposto che questi elementi si attivino anche in presenza di un cosiddetto "stress genetico". Stress interni come l'alloploidizzazione e l'ibridazione sono un esempio di stress genetico e causano l'attivazione di elementi trasponibili. Ad esempio Kashkush et al. (2003) hanno osservato attivazione trascrizionale di retrotrasposoni in frumento poliploide e l'attivazione di retrotrasposoni è stata individuata in ibridi di riso (Liu e Wendel, 2000). I risultati di questo lavoro potrebbero rafforzare questo tipo di ipotesi dato che sono derivati dall'analisi di una linea deleta (la delezione è da considerare uno stress genetico) in condizione di stress idrico.

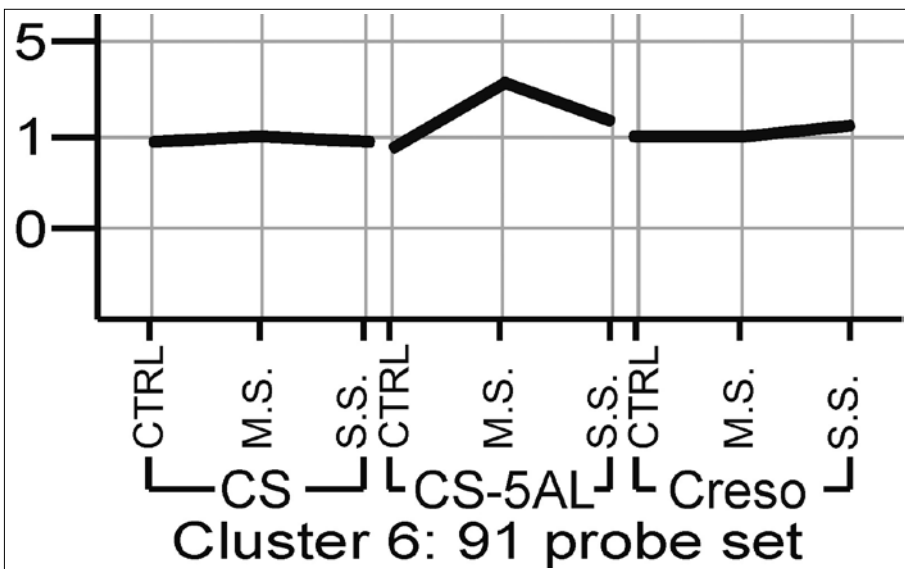


Fig. 10 Cluster 6

I risultati di questo lavoro indicano che c'è attivazione trascrizionale maggiore di questi elementi trasponibili nei campioni sottoposti a stress moderato rispetto ai campioni severamente trattati. Ciò fa supporre che un tipo di regolazione epigenetica basata sull'attivazione degli elementi trasponibili può avvenire in una fase di acclimatamento a stress abiotico.

I restanti cluster sono tutti piccole variazioni di quelli già descritti sinora.

L'analisi cluster unita alla ricerca di categorie sovra rappresentate si è rivelata essere efficace nell'identificazione di geni coinvolti nello stress idrico e ha permesso di fare luce sulle differenze fra i profili trascrizionali di frumento tenero e duro. Un'altra strategia perseguibile di data mining è la ricostruzione di pathway di interesse. Il concetto sottostante questa analisi è molto semplice e consiste nell'osservare i livelli di espressione di tutti i geni codificanti gli enzimi di una via biosintetica.

Geni codificanti per il pathway di biosintesi dell'ormone ABA

L'acido abscissico è l'ormone che regola la risposta alla carenza d'acqua nelle piante. Il pathway di biosintesi dell'ABA e la sua regolazione sono stati ampiamente descritti (Seo and Koshiba, 2002). L'utilizzo di esperimenti microarray consente di dissezionare punto per punto vie biosintetiche completamente descritte e di ricostruire l'andamento d'espressione dei geni coinvolti nelle condizioni e nei genotipi di interesse.

I passaggi iniziali del pathway coincidono con il pathway di biosintesi dei carotenoidi poiché il β -carotene è il principale precursore dell'ABA. Dato che la fitoene sintasi (PSY) è stata descritta essere il primo enzima che regola la sintesi dei carotenoidi, il controllo dei livelli d'espressione è iniziato proprio dalla PSY, per poi ripercorrere tutti i passaggi che portano alla sintesi dell'ABA.

La biosintesi dell'ABA, a partire dal geranilgeranil-PP, necessita di 10 reazioni chimiche di cui 9 catalizzate da enzimi (fig. 11).

Sul GeneChip[®]Wheat Genome Array sono stati identificati 118 probe set corrispondenti a sequenze codificanti per i 9 enzimi del pathway.

Il primo enzima del pathway è la fitoene sintasi (2.5.1.32) ed è rappresentato da tre probe set sul microarray. Uno dei tre probe set mostra livelli di espressione nulli (o comunque inferiori al livello del background medio dell'array), un altro non è differenzialmente espresso e mostra livelli d'espressione poco maggiori del valore del background, mentre il terzo (Ta.20776.1.S1_at) risulta differenzialmente espresso e il comportamento d'espressione appartiene alla tipologia descritta nel cluster 1.

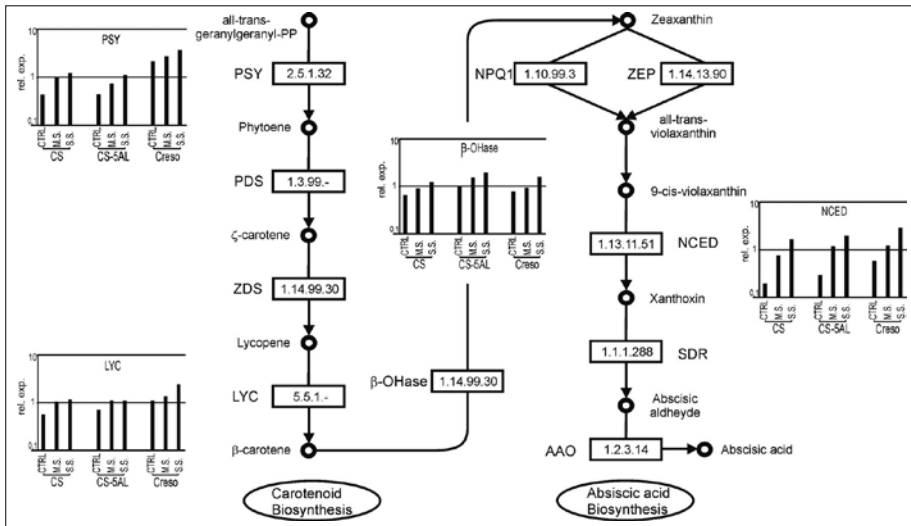


Fig. 11 *Pathway di biosintesi dell'acido abscissico. Sulla sinistra la porzione di via metabolica che porta alla formazione del β -carotene. Sulla destra quella specifica per l'acido abscissico. I quattro istogrammi rappresentano i livelli di espressione dei probe set differenzialmente espressi di PSY, LYC, β -OHase e NCED. In ascissa vi sono i campioni raggruppati per genotipi e stress crescente (CTRL, MS e SS). In ordinata i valori di espressione del gene normalizzato in base al valore della mediana in tutti i campioni. Il valore 1 significa che il valore d'espressione del gene è uguale a quello della mediana su tutti i campioni. 2.5.1.32=Fitoene sintasi; 1.3.99.-=Fitoene desaturasi; 1.14.99.30= ζ -carotene desaturasi; 5.5.1.-=licopene β -ciclasi; 1.14.99.-= β -carotene idrossilasi; 1.14.13.90=Zeaxantina epossidasi; 1.13.11.51=9-cis-epossicarotenoide diossigenasi; 1.1.1.288=SDR; 1.2.3.14=AAO3*

Il secondo enzima è la fitoene desaturasi (PDS, 1.14.99.-). Vi sono due probe set sul chip che hanno elevata similarità di sequenza con il gene codificante questo enzima. Uno non è espresso, mentre l'altro mostra livelli d'espressione medio-bassi (2 volte il valore del background) e costanti in tutti i campioni. Si osserva la stessa situazione anche per l'enzima ζ -carotene desaturasi (ZDS, 1.14.99.30): due probe set di cui uno non espresso e l'altro che mostra livelli d'espressione elevati ma costanti in tutti i campioni.

Lo step successivo è regolato dalla licopene β -ciclasi (LYC, 5.5.1.-). Sul microarray sono presenti tre probe set relativi al gene codificante la LYC. Due di questi hanno livelli d'espressione costanti. Uno (TaAffx.95060.1.S1_at), invece, risulta differenzialmente espresso nel confronto tra controllo e campione sottoposto a stress intenso nel genotipo Cres. Tuttavia il livello d'espressione di quest'ultimo probe set non è molto elevato. Con questa reazione si conclude il pathway di biosintesi del β -carotene. Le reazione che

verranno discusse in seguito sono, invece, specifiche della via biosintetica dell'acido abscisico.

Il primo passaggio prevede la conversione del β -carotene a zeaxantina, una reazione catalizzata dalla β -carotene idrossilasi (β -OHase, 1.14.99.-). Su tre probe set corrispondenti a questo enzima, due mostrano livello d'espressione nullo, mentre uno (Ta.8484.1.A1_at) risulta indotto all'aumentare dello stress idrico nei tre genotipi.

La zeaxantina epossidasi (ZEP, 1.14.13.90) è rappresentata da due probe set. Uno non è espresso, mentre l'altro ha livelli di espressione medi (8 volte il background) e risulta essere indotto nei tre genotipi all'aumentare dello stress idrico. Tuttavia il livello di induzione non è così elevato da permettere, in base alla definizione riportata nei materiali e metodi, di essere considerato un gene differenzialmente espresso.

Il passaggio successivo è la scissione ossidativa della 9-cis-violaxantina in xantoxina catalizzata dall'enzima 9-cis-epossicarotenoide diossigenasi (NCED, 1.13.11.51). In *Phaseolus vulgaris* è stato osservato che, in condizioni di stress idrico, il livello di mRNA per questo enzima aumenta notevolmente, così come i livelli di proteina e di ABA suggerendo il ruolo di regolazione di questo gene nella biosintesi (Qin and Zeevart 1999). Questi dati sono stati confermati anche in altre specie di interesse agrario come in avocado, pomodoro e fagiolo. Anche in *Arabidopsis* è stato osservato come l'sovra espressione di AtNCED sia indotta da stress idrico e determini l'accumulo di ABA (Iuchi et al., 2001). Sul GeneChip[®]Wheat Genome Array esistono 13 probe set associati al gene codificante la NCED. Di questi, cinque hanno livelli di espressione inferiori a quelli del background, quattro hanno livelli di espressione medio-alti ma costanti, e i restanti quattro (Ta.12813.1.S1_x_at, Ta.12813.2.S1_x_at, TaAffx.13292.1.S1_at, TaAffx.13292.1.S1_s_at) sono differenzialmente espressi e sono stati raggruppati nel cluster 3, cioè il cluster dei geni di risposta allo stress idrico.

L'enzima SDR (Short Chain Dehydrogenase/reductase) è rappresentato da un numero elevatissimo di probe set (90) e solo tre di essi risultano differenzialmente espressi. Un unico probe set rappresenta l'ultimo enzima del pathway e risulta non differenzialmente espresso.

In conclusione dallo studio dettagliato dei geni codificanti gli enzimi principali della via biosintetica per la produzione di ABA è risultato che alcuni trascritti (PSY, LYC, β -OHase, e SDR) sono blandamente indotti dalla riduzione di acqua disponibile, mentre il trascritto relativo alla NCED risulta fortemente indotte. Questi dati si allineano perfettamente a quelli riscontrati

in letteratura dove l'enzima per la NCED è descritto essere l'enzima chiave del pathway; risulta altresì evidente che la sua regolazione avviene a livello trascrizionale.

Geni codificanti per il pathway di biosintesi della prolina

La prolina è un amminoacido apolare che funziona anche come osmo-protettore nella risposta allo stress idrico nelle piante. Piante transgeniche in grado di accumulare alte concentrazioni di questo osmolita mostrano una maggiore tolleranza a condizioni di carenza idrica (Kavi Kishor et al., 1995).

Sul GeneChip[®] Wheat Genome Array sono stati identificati 28 probe set corrispondenti a sequenze codificanti per 9 enzimi coinvolti nel pathway di biosintesi e degradazione della prolina.

Nelle piante questo amminoacido è sintetizzato attraverso due vie differenti: una via ha come precursore il glutammato mentre l'altra l'ornitina (De-launey et al., 1993) ed è stato dimostrato che il pathway che ha inizio con il glutammato è catalizzato dalla P5CS (Δ -Pyrroline-5-carboxylate synthetase), un enzima bifunzionale dal quale dipende la velocità di sintesi. La sua prima attività, la γ -glutammil sintasi, catalizza la fosforilazione ATP-dipendente dell'L-glutammato. Il γ -glutammil fosfato così prodotto è poi convertito a glutammico γ -semialdeide (GSA) a opera della GSA riduttasi NADPH-dipendente (Hu et al., 1992; Yoshiba et al., 1995). Il primo enzima analizzato, dunque, è stata la P5CS che è rappresentata sull'array da un solo probe set (Ta.7091.1.S1_at). Questo probe set risulta indotto nella condizione di stress ed è un gene raggruppato nel cluster 3. L'enzima P5CS è, infatti, subisce un'induzione proporzionale alla severità dello stress nei tre genotipi di frumento, con l'eccezione che i livelli di espressione dei geni di frumento duro sono mediamente maggiori rispetto a quelli dei frumenti teneri (fig. 12).

Dato che in condizioni di stress osmotico P5CS svolge un ruolo chiave nella biosintesi della prolina in quanto è coinvolto nel passaggio maggiormente regolato del processo biosintetico (Kiyosue et al., 1996) possiamo pensare che le varietà di frumento utilizzate in questo lavoro attuino una risposta di aggiustamento osmotico mediato dall'accumulo di prolina come difesa allo stress idrico.

Il passaggio successivo è la ciclizzazione spontanea dell'intermedio L-glutammato-semialdeide a D-pirrolina-5-carbossilato, che è convertita a prolina a opera dell'enzima NADPH-dipendente D-pirrolina-5-carbossilato reductasi (P5C1, 1.5.1.2). Questo enzima sull'array è rappresentato da tre Probe set

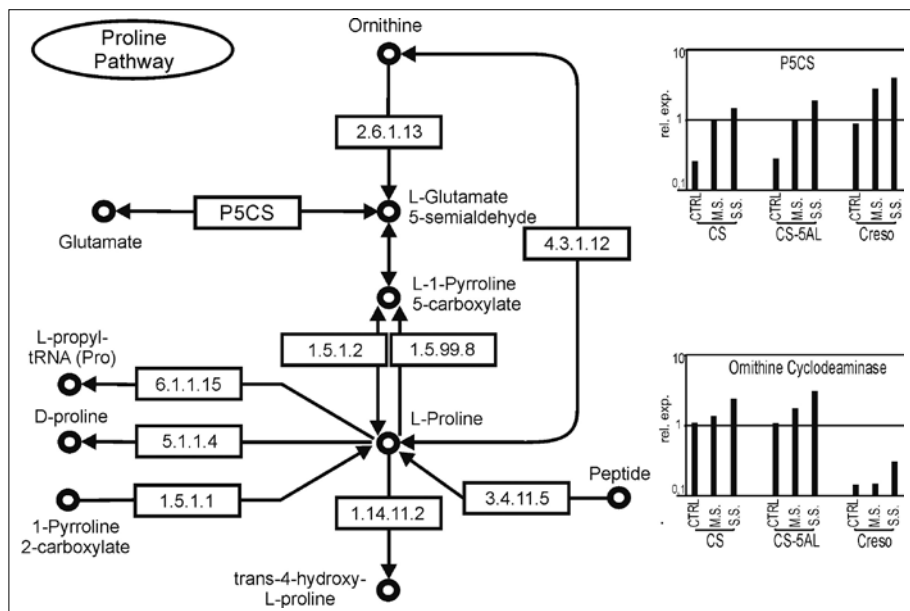


Fig. 12 Pathway di biosintesi della prolina. I due istogrammi rappresentano i livelli di espressione dei probe set differenzialmente espressi: P5CS e Ornitina ciclodeaminasi. In ascissa vi sono i campioni raggruppati per genotipi e stress crescente (CTRL, MS e SS). In ordinata i valori di espressione del gene normalizzati in base al valore della mediana in tutti i campioni. Il valore 1 significa che il valore d'espressione del gene è uguale a quello della mediana su tutti i campioni. 2.6.1.13=ornithine aminotransferase; P5CS= Δ -pyrroline-5-5carboxylate synthase; 4.3.1.12= ornithine cyclodeaminase; 1.5.1.2= Δ -pyrroline-5-5carboxylate reductase; 1.5.99.8=proline dehydrogenase; 6.1.1.15=prolyl-tRNA synthase; 5.1.1.4=proline racemase; 1.14.11.2=prolyl hidrosilase; 3.4.11.5=prolyl aminotransferase; 1.5.1.1=pyrroline-2-carboxylate reductase

nessuno dei quali è risultato differenzialmente espresso. Ciò conferma che la biosintesi della prolina è regolata a livello dell'enzima P5CS.

La seconda via di biosintesi della prolina, converte l'ornitina a L-glutammato-semialdeide con l'intervento di catalisi dell'ornitina- δ -amminotransferasi (2.6.1.13). Sul GeneChip[®]Wheat Genome Array vi sono 5 Probe sets rappresentativi per questo enzima ma nessuno di questi presenta variazioni del livello di espressione in funzione dello stress imposto. L'ornitina- δ -amminotransferasi e l'enzima della via derivata dal glutammato, pirrolina-5-carbossilato sintetasi, appaiono essere regolati in modo differente a livello trascrizionale. Da qui le due vie convergono fino alla sintesi della prolina.

La produzione della prolina a partire da ornitina può avvenire anche in un unico passaggio catalizzato dalla ornitina ciclodeaminasi (4.3.1.12) (Trovato

et al., 2001). Nei microarray per questo enzima è presente un solo probe set e il suo livello di espressione aumenta in seguito a stress idrico nei frumenti teneri, mentre in Creso il livello di espressione è nullo.

Sono stati indagati altri enzimi che catalizzano la produzione di prolina a partire da altri precursori, come la prolyl aminopeptidasi (3.4.11.5), per la quale esistono 4 Probe set sull'array e altri 5 Probe set rappresentativi per la famiglia delle aminopeptidasi ma nessuno di questi ha mostrato modulazione del livello di espressione in condizione di stress idrico o la pyrroline-2-carboxylate reductase (1.5.1.1) che, però, non presenta Probe sets sull'array.

La degradazione della prolina può avvenire in diversi modi. Uno degli enzimi coinvolti nel catabolismo di questo amminoacido è la prolina deidrogenasi (1.5.99.8), che è rappresentata sul chip da 4 Probe sets. Tra questi c'è pure la prolyl-tRNA synthase, che conta sul chip 4 Probe sets. Nessuno tra questi enzimi di degradazione della prolina risulta differenzialmente espresso. Gli enzimi proline racemase (5.1.1.4) e prolyl idrossilasi (1.14.112) non sono rappresentati da alcun Probe set sul microarray.

L'analisi della prolina deidrogenasi di *Arabidopsis* (PDH), che catalizza la conversione della prolina di nuovo a pirrolina-5-carbossilato, la prima delle reazioni degradative in lievito, suggerisce che il catabolismo giochi un ruolo importante nel determinare le concentrazioni di prolina libera nelle piante (Nanjo et al., 1999). L'mRNA della prolina deidrogenasi è regolato in modo reciproco all'enzima biosintetico pirrolina-5-carbossilato sintetasi sia durante la disidratazione/reidratazione sia durante lo stress da NaCl e il recupero. Questi risultati suggeriscono che la via degradativa è soppressa durante i periodi di sintesi *de novo*, per impedire un ciclo futile, e successivamente è attivata per riportare le concentrazioni di prolina alle condizioni precedenti lo stress. Il risultato di questo lavoro confermano questi studi, dato che, i livelli di espressione del gene della prolina deidrogenasi sono molto bassi o addirittura nulli nei tre diversi genotipi e alle diverse condizioni di idratazione.

Ciò che si può concludere da queste osservazioni è che la via biosintetica della prolina dipendente da glutammato è attiva in tutti e tre i genotipi. L'espressione genica dell'enzima regolatore della via biosintetica della prolina a partire da acido glutammico risulta indotta da stress idrico. Tuttavia, i livelli di espressione della P5CS risultano maggiori in Creso. Per contro, nei frumenti teneri è stato individuato un trascritto con alta similarità di sequenza con una ornitina ciclodeaminasi, tale trascritto risulta essere indotto da stress idrico. Mentre per la P5CS è nota la funzione in piante e la relativa regolazione di tipo trascrizionale, per l'ornitina ciclodeaminasi sono necessari ulteriori studi per verificare

l'effettiva presenza della proteina e la putativa funzione. Nel caso in cui fosse confermato il ruolo osservato nei batteri potremmo concludere che, mentre Creso produce prolina preferenzialmente a partire da glutammato, i frumenti teneri sono in grado di convertire sia glutammato che ornitina a prolina.

Validazione dei dati microarray

Per confermare che il dato ottenuto attraverso l'analisi microarray rispecchi realmente la quantità di trascritto presente nei campioni, sono stati selezionati 21 geni al fine di controllare il loro livello d'espressione con la tecnologia a oggi ritenuta la più sensibile fra quelle disponibili: la qRT-PCR.

Delle 21 coppie di primer disegnate, 15 si sono rivelate idonee per l'analisi in real time PCR dato che presentavano un solo picco nella curva di dissociazione. Di queste 15 coppie, circa il 90% hanno mostrato avere un livello di espressione simile a quello osservato tramite microarray nei 9 campioni oggetto di questo studio. Anche in altri lavori sono state riportate percentuali simili di sovrapposizioni tra dati array e dati qRT-PCR.

Anche il reference gene selezionato sfruttando i dati array si è dimostrato idoneo allo scopo.

Mappaggio di geni sulla base del livello di espressione

Sulla base della procedura descritta nei materiali e metodi, è possibile identificare geni il cui livello di espressione sia nullo in Creso o in CS-5AL e allo stesso tempo che sia medio-alto in CS. Tale approccio è il primo esempio di identificazione della posizione cromosomica di geni di frumento sulla base dell'espressione genica ottenuta da dati microarray. Per questo motivo sono stati applicati criteri molto stringenti. Si sono così identificati 197 geni putativamente collocati sul genoma D e 25 sul braccio lungo del cromosoma 5A.

Dato l'alta tasso di ridondanza genica nei frumenti sono stati necessari alcuni accorgimenti nel disegno dei primer al fine di escludere falsi positivi. A tal scopo per ogni gene candidato per una determinata regione cromosomica, sono state disegnate più coppie di primer. Una coppia di primer definita "aspecifica" poiché in grado di amplificare su tutti i geni omologhi dei tre genomi di frumento tenero. Tale coppia di primer funge da controllo positivo. Successivamente è stata disegnata almeno un'altra coppia di primer definita "specificata" dato che il loro target è esclusivamente il gene da map-

pare. Se i primer “aspecifici” amplificano la sequenza prevista in frumento tenero e duro e si osserva assenza di amplificazione da parte dei primer “specifici” in frumento duro allora il gene può essere ragionevolmente localizzato sul genoma D.

La validazione su DNA genomico ha evidenziato che circa il 50% delle 70 coppie di primer per i geni testati per putativa posizione sul genoma D, effettivamente non danno alcun prodotto di amplificazione in Creso e in Trinakria.. Dato che circa il 33% dei geni di frumento tenero sono attesi essere posizionati sul genoma D, il test del χ^2 , che permette di confrontare la percentuale attesa con quella effettiva, porta alla conclusione che l'approccio utilizzato non è significativo nell'individuazione di geni posizionati sul genoma D. In futuro si pensa di aumentare il numero di geni testati.

Stessa percentuale di conferma si è osservata per i putativi geni del braccio lungo del cromosoma 5A. Su 23 coppie di primer testate, 12 (52%) non amplificano alcun prodotto genico nel genotipo CS-5AL.

In questo caso, dove la frequenza attesa di geni appartenenti al braccio lungo del cromosoma 5A è di circa il 5%, il test del χ^2 porta alla conclusione che la strategia utilizzata permette un consistente arricchimento di liste di geni appartenenti al 5AL.

In conclusione la medesima strategia nella ricerca di geni appartenenti al genoma D o a una porzione del braccio lungo del cromosoma 5A ha portato a ottenere liste arricchite nella presenza di tali geni. Tuttavia mentre nel secondo caso la percentuale di arricchimento è molto elevata, nel caso dei geni del genoma D l'arricchimento è meno evidente. Questo fenomeno potrebbe derivare dal disegno di primer non effettivamente specifici per le sequenze di interesse. Infatti a causa della limitata conoscenza del genoma del frumento potrebbero esistere sequenze non presenti in banche dati, ma omologhe a quelle individuate. Per questo motivo le sequenze che non sono state confermate tramite PCR per essere posizionate sul genoma D non necessariamente sono collocate sul genoma A o B.

CONCLUSIONI

Il trascrittoma di CS (frumento esaploide) in condizioni di irrigazione è risultato profondamente differente da quello della varietà Creso (frumento tetraploide) mostrando oltre 6000 geni espressi in modo differente tra i due genotipi. Considerando che, mediamente, sono stati individuati circa 30mila probe set espressi, l'assenza del genoma D nel frumento duro comporta una

ampia riorganizzazione del trascrittoma attraverso la modulazione di 1/5 della totalità dei geni espressi. Questa situazione influenza anche la risposta allo stress idrico poiché pochi geni hanno mostrato avere comportamenti di espressione simili in risposta a stress idrico a causa di livelli trascrizionali di partenza considerevolmente differenti.

Nel confronto tra CS e CS-5AL in condizioni di irrigazione sono stati osservati circa 600 geni con differenti livelli di espressione causati, quindi, dalla mutazione cromosomica a carico del braccio lungo del cromosoma 5A. Tuttavia questa mutazione cromosomica non ha determinato notevoli cambiamenti nella risposta allo stress idrico tra CS e CS-5AL.

L'analisi dei profili di espressioni di frumento tenero e duro ha messo in evidenza, innanzitutto, che oltre 3000 geni differenti possono essere attivati come risposta allo stress idrico. Solo una piccola porzione è attivata o repressa comunemente dai tre genotipi studiati, mettendo in evidenza una netta differenza tra la risposta molecolare del frumento tenero da quella del duro. Tuttavia, fra i geni indotti in egual modo fra frumento tenero e duro ritroviamo geni codificanti per enzimi con ruoli chiave nella risposta allo stress idrico, come quelli coinvolti nella sintesi dell'acido abscissico, l'ormone che regola la risposta allo stress idrico nei vegetali.

Un'altra risposta molto importante che le piante mettono in atto è la produzione di osmoliti in grado di contrastare gli effetti della carenza idrica. Dall'analisi dei trascrittomi emergono alcune situazioni interessanti. Il sorbitolo, per esempio, si ottiene dalla conversione del glucosio a opera dell'enzima aldoso reductasi. Sul microarray utilizzato vi sono due interessanti probe set per questo gene e tutti hanno un differente comportamento d'espressione: uno ha un andamento comune nei tre genotipi, mentre l'altro è frumento-tenero specifico.

Situazione simile anche per un altro osmolita come la prolina. Il pathway di biosintesi di questo aminoacido ha due possibili precursori: il glutammato e l'ornitina. L'enzima chiave del pathway glutammato dipendente è la P5CS (Delta 1 Pirrolina-5-carbossilato sintasi), mentre la ornitina ciclodeaminasi catalizza la reazione da ornitina a prolina. I livelli di espressione del gene per la P5CS aumentano sia nei frumenti teneri che duri all'aumentare dello stress idrico, tuttavia il livello di espressione in Creso è molto più elevato rispetto ai frumenti teneri. Per contro il gene per l'ornitina ciclodeaminasi viene attivato solo nei frumenti teneri, mentre il livello di espressione in Creso è nullo in tutte le condizioni.

Da questi dati si può riassumere che i trascritti attivati da frumento tenero e frumento duro sono molto differenti tra loro, così come l'intero profilo

d'espressione. Tuttavia le risposte che vengono attivate sono spesso identiche anche se vengono ottenute attraverso l'induzione di geni differenti.

Un altro punto critico che si può evidenziare nell'analisi dei profili d'espressione è che Creso necessita di uno stress maggiore affinché attivi una risposta di tipo trascrizionale. Infatti mentre CS, nella transizione tra controllo e stress moderato, riesce a modulare l'espressione di circa 100 geni, Creso non ne modula nemmeno uno. Questo fenomeno si riscontra anche nei cluster 15 e 19. Come spiegato in precedenza questi cluster contengono molti geni codificanti microtubuli ed enzimi deputati alla sintesi/degradazione della parete cellulare. Per questo motivo una riduzione del loro livello di espressione comporta una riduzione della divisione e distensione cellulare. Il fatto che il loro livello di espressione sia ridotto in seguito a uno stress moderato in CS, mentre per Creso sia necessario uno stress intenso, fa supporre che CS sia maggiormente sensibile nella percezione della carenza d'acqua rispetto a Creso.

Inoltre lo studio dei trascrittomi ha permesso di estrapolare informazioni non strettamente inerenti con lo stress idrico:

- 1) è stato identificato un ottimo reference gene il cui livello di espressione in condizioni di stress idrico rimane immutato. Il suo utilizzo si potrà estendere anche a situazioni e genotipi differenti.
- 2) sono stati identificati 91 putative sequenze di trasposoni e retrotrasposoni con elevata attività trascrizionale in situazioni di particolari combinazioni di stress (mutazione cromosomica e moderato stress idrico).
- 3) in base al livello d'espressione sono stati collocati geni sul genoma D e sul braccio lungo del cromosoma 5A.

RIASSUNTO

Lo stress idrico durante la fase di riempimento delle cariossidi ha un marcato effetto sulla produzione poiché determina una riduzione del numero delle cellule dell'endosperma e, di conseguenza, della capacità di accumulare sostanza secca. In questo lavoro per osservare la risposta molecolare in frumento e la sua correlazione con la presenza o assenza di determinate regioni cromosomiche, è stata condotta una analisi dei profili trascrizionali su tre genotipi: un frumento tenero (Chinese Spring – CS), una sua linea di delezione (CS_5AL-10) e un frumento duro (Creso).

L'analisi del trascrittoma fra le due specie ha messo in evidenza 8552 probe sets con livelli di espressione significativamente differenti. Lo stress idrico, invece, ha causato la variazione di espressione di 3056 probe sets. Fra questi è stato possibile individuare gruppi di geni con andamenti di espressione simili fra i tre genotipi. Tuttavia altri gruppi hanno mostrato comportamenti differenti a seconda della struttura genomica di Creso, CS e CS_5AL-10. Inoltre l'approccio dell'analisi di espressione fra genotipi con struttura ge-

nomica differente ha permesso di identificare geni posizionati sul genoma D e sul braccio lungo del cromosoma 5A, successivamente verificati tramite PCR. L'analisi funzionale dei geni differenzialmente espressi in condizione di stress idrico ha evidenziato che geni appartenenti alle vie metaboliche dell'ABA, prolina, glicina-betaina e del sorbitolo sono indotti da carenza idrica. Allo stesso modo sono stati identificati una serie di sequenze correlate con trasposoni e retrotrasposoni che si attivano solo nella linea di delezione e in seguito a moderati stress idrici.

ABSTRACT

Water stress during grain filling has a marked effect on grain yield, leading to a reduced endosperm cell number and thus sink capacity to accumulate dry matter. The bread wheat cultivar Chinese Spring (CS), a Chinese Spring terminal deletion line (CS_5AL-10) and the durum wheat cultivar Creso were subjected to transcriptional profiling after exposure to mild and severe drought stress at the grain filling stage to find evidences of differential stress responses associated to different wheat genome regions.

The transcriptome analysis of Creso, CS and its deletion line revealed 8,552 non redundant probe sets with different expression levels, mainly due to the comparisons between the two species. The drought treatments modified the expression of 3,056 probe sets. Besides a set of genes showing a similar drought response in Creso and CS, cluster analysis revealed several drought response features that can be associated to the different genomic structure of Creso, CS and CS_5AL-10. Some drought-related genes were expressed at lower level (or not expressed) in Creso (which lacks the D genome) or in the CS_5AL-10 deletion line compared to CS. The chromosome location of a set of these genes was confirmed by PCR-based mapping on the D genome (or the 5AL-10 region). Many clusters were characterized by different level of expression in Creso, CS and CS_5AL-10, suggesting that the different genome organization of the three genotypes may affect plant adaptation to stress. Clusters with similar expression trend were grouped and functional classified to mine the biological mean of their activation or repression. Genes involved in ABA, proline, glycine-betaine and sorbitol pathways were found up-regulated by drought stress. Furthermore, the enhanced expression of a set of transposons and retrotransposons was detected in CS_5AL-10.

Bread and durum wheat genotypes were characterized by a different physiological reaction to water stress and by a substantially different molecular response. The genome organization accounted for differences in the expression level of hundreds of genes located on the D genome or controlled by regulators located on the D genome. When a genomic stress (deletion of a chromosomal region) was combined with low water availability, a molecular response based on the activation of transposons and retrotransposons was observed.

BIBLIOGRAFIA

ATTIENZA S.G., FACCIOLI P., PERROTTA G., DALFINO G., ZSCHIESCHE W., HUMBECK K., STANCA A.M., CATTIVELLI L. (2004): *Large scale analysis of transcripts abundance in barley subjected to several single and combined abiotic stress conditions*, «Plant Science», 167, pp. 1359-1365.

- BENJAMINI Y., HOCBERG Y. (1995): *Controlling the false discovery rate: a practical and powerful approach to multiple testing*, «Journal of the Royal Statistical Society», 57, pp. 289-300.
- CATTIVELLI L., RIZZA F., BADECK FW., MAZZUCOTELLI E., MASTRANGELO A.M., FRANCA E., MARÈ C., TONDELLI A., STANCA A.M. (2008): *Drought tolerance improvement in crop plants: An integrated view from breeding to genomics*, «Field Crops Research», 4, pp. 1-14.
- CATTIVELLI L., BALDI P., CROSATTI C., DI FONZO N., FACCIOLI P., GROSSI M., MASTRANGELO A.M., PECCHIONI N., STANCA A.M. (2002): *Chromosome regions and stress-related sequences involved in resistance to abiotic stress in Triticeae*, «Plant Molecular Biology», 48, pp. 649-665.
- CHEN T.H., MURATA N. (2002): *Enhancement of tolerance of abiotic stress by metabolic engineering of betaines and other compatible solutes*, «Current Opinion in Plant Biology», 5, pp. 250-257.
- CHEN J.J., WANG S-J., TSAI C-A., LIN C-J. (2007): *Selection of differentially expressed genes in microarray data analysis*, «The Pharmacogenomics Journal», 7, pp. 212-220.
- CHOI H., HONG J.H., HA J., KANG J.Y., KIM S.Y. (2000): *ABFs, a family of ABA-responsive elements binding factors*, «Journal of Biological Chemistry», 275, pp. 1723-1730.
- CHOMCZYNSKI P., SACCHI N. (1987): *Signal-Step Method of RNA Isolation by Acid Guanidinium Thiocyanate-Phenol-Chloroform Extraction*, «Analytical Biochemistry», 162, pp. 156-159.
- COEN E.S., CARPENTER R., ARTIN C. (1986): *Transposable elements generate novel spatial patterns of gene expression in Antirrhinum majus*, «Cell», 47, pp. 285-296.
- CRAMER G.R., ERGÜL A., GRIMPLET J., TILLET R.L., TATTERSALL E.A.R., BOHLMAN M.C., VINCENT D., SONDEREGGER J., EVANS J., OSBORNE C., QUILICI D., SCHLAUCH K.A., SCHOOLEY D.A., CUSHMAN J.C. (2007): *Water and salinity stress in grapevines: early and late changes in transcript and metabolite profiles*, «Functional and Integrative Genomics», 7, pp. 111-134.
- CZECHOWSKI T., STITT M., ALTMANN T., UDVARDI M.K., SCHEIBLE W-R. (2005): *Genome-Wide Identification and Testing of Superior Reference Genes for Transcript Normalization in Arabidopsis*, «Plant Physiology», 139, pp. 5-17.
- DELAUNEY A.J., HU C.A., KISHOR B.P., VERMA D.P. (1993): *Cloning of Ornithine delta-aminotransferase cDNA from Vigna aconitifolia by trans-complementation in Escherichia coli and regulation of proline biosynthesis*, «Journal of Biological Chemistry», 268, pp. 18673-18678.
- DUANGSRISAI S., YAMADA K., BANTOG N.A., SHIRATAKE K., KANAYAMA Y., YAMAKI S. (2007): *Presence and expression of NAD⁺-dependent sorbitol dehydrogenase and sorbitol-6-phosphate dehydrogenase genes in strawberry*, «Journal of Horticulture Science and Biotechnology», 82, pp. 191-198.
- EL-TAYEB M.A. (2005): *Response of barley grains to the interactive effect of salinity and salicylic acid*, «Plant Growth Regulation», 45, pp. 215-224.
- GAUTIER L., COPE L., BOLSTAD B.M., IRIZARRY R.A. (2004): *Affy - an R package for the analysis of affymetrix genechip data at the probe level*, «Bioinformatics», 20, pp. 307-315.
- FACCIOLI P., CICERI G.P., PROVERO P., STANCA A.M., MORCIA C., TERZI V. (2007): *A combined strategy of "in silico" transcriptome analysis and web search engine optimization*

- allows an agile identification of reference genes suitable for normalization in gene expression studies, «Plant Molecular Biology», 63, pp. 679-688.
- HAAKE V., COOK D., RIECHMANN J.L., PINEDA O., THOMASHOW M.F., ZHANG J.Z. (2002): *Transcription factor CBF4 is a regulator of drought adaptation in Arabidopsis*, «Plant Physiology», 130, pp. 639-648.
- HEYER L.J., KRUGLYAK S., YOSEPH S. (1999): *Exploring expression data: identification and analysis of coexpressed genes*, «Genome Research», 9, pp. 1106-1115.
- HU C.A., DELAUNEY A.J., VERMA D.P.S. (1992): *A bifunctional enzyme (Δ^1 -pyrroline-5-carboxylate synthase) catalyzes the first two steps in proline biosynthesis in plants*, «Proceedings of the National Academy of Sciences USA», 89, pp. 9354-9358.
- IRIZARRY R.A., HOBBS B., COLIN F., BEAZER-BARCLAY Y.D., ANTONELLIS K., SCHERF U., SPEED T.P. (2003): *Exploration, normalization and summaries of high density oligonucleotide array probe level data*, «Biostatistics», 4, pp. 249-264.
- IRIZARRY R.A., GAUTIER L., BOLSTAD B.M., MILLER C. (2006): *Affy: Methods for Affymetrix Oligonucleotide Arrays*, «R package version 1.12.1».
- IVASHUTA S., NAUMKINA M., GAU M., UCHIYAMA K., ISOBE S., MIZUKAMI Y., SHIMAMOTO Y. (2002): *Genotype-dependent transcriptional activation of novel repetitive elements during cold acclimation of alfalfa (Medicago sativa)*, «Plant Journal», 31, pp. 615-627.
- IUCHI S., KOBAYASHI M., TAJI T., NARAMOTO M., SEKI M., KATO T., TABATA S., KAKUBARI Y., YAMAGUCHI-SHINOZAKI K., SHINOZAKI K. (2001): *Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in Arabidopsis*, «The Plant Journal», 27, pp. 325-333.
- KASHKUSH K., FELDMAN M., LEVY A.A. (2003): *Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat*, «Nature Genetics», 33, pp. 102-106.
- KAVI KISHOR P.B., HONG Z., MIAO G.H., HU C.A.A., VERMA D.P.S. (1995): *Overexpression of Δ^1 -pyrroline-5-carboxylate synthase increases proline production and confers osmotolerance in transgenic plants*, «Plant Physiology», 108, pp. 1387-1394.
- KIRCH H.H., SCHLINGENSIEPEN S., KOTCHONI S., SUNKAR R., BARTELS D. (2005): *Detailed expression analysis of selected genes of the aldehyde dehydrogenase (ALDH) gene superfamily in Arabidopsis thaliana*, «Plant Molecular Biology», 57, pp. 315-332.
- KIYOSUE T., YOSHIDA Y., YAMAGUCHI-SHINOZAKI K., SHINOZAKI K. (1996): *A gene encoding mitochondrial proline dehydrogenase, an enzyme involved in proline metabolism, is up-regulated by proline but down-regulated by dehydration in Arabidopsis*, «Plant Cell», 8, pp. 1323-1335.
- KNIGHT H. (2000): *Calcium signaling during abiotic stress in plant*, «International Review of Cytology», 195, pp. 269-324.
- LEE P.D., SLADEK R., GREENWOOD C.M.T., HUDSON T.J. (2001): *Control genes and variability: absence of ubiquitous reference transcripts in diverse mammalian expression studies*, «Genome Research», 12, pp. 292-297.
- LIU X.D., SHEN Y.G. (2004): *NaCl-induced phosphorylation of light harvesting chlorophyll a/b proteins in thylakoid membranes from the halotolerant green alga, Dunaliella salina*, «FEBS Letters», 569, pp. 337-340.
- LIU B., WENDEL J.F. (2000): *Retrotransposon activation followed by rapid repression in introgressed rice plants*, «Genome», 43, pp. 874-880.
- LO BIANCO R., RIEGER M., SUNG S.J.S. (2000): *Effect of drought on sorbitol and sucrose metabolism in sinks and sources of peach*, «Physiologia Plantarum», 108, pp. 71-78.

- MADLUNG A., COMAI L. (2004): *The effect of stress on genome regulation and structure*, «Annals of Botany», 94, pp. 481-495.
- MARÈ C., MAZZUCOTELLI E., CROSATTI C., FRANCA E., STANCA AM., CATTIVELLI L. (2004): *Hv-WRKY38: a new transcription factor invlved in cold- and drought-response in barley*, «Plant Molecular biology», 55, pp. 399-416.
- MITTLER R. (2002): *Oxidative stress, antioxidants and stress tolerance*, «Trends Plant Sciences», 7, pp. 405-410.
- NANJIO T., KOBAYASHI M., YOSHIDA Y., KAKUBARI Y., YAMAGUCHI-SHINOZAKI K., SHINOZAKI K. (1999): *Antisense suppression of proline degradation improves tolerance to freezing and salinity in Arabidopsis thaliana*, «FEBS letters», 461, pp. 205-210.
- QIN X., ZEEVART J.A.D. (1999): *The 9-cis-epoxycarotenoid cleavage reaction is the key regulatory step of abscisic acid biosynthesis in water stressed bean*, «Proceeding of the National Academy of Sciences of the USA», 96, pp. 15354-15361.
- RAPP R.A., WENDEL J.F. (2005): *Epigenetics and plant evolution*, «New Phytologist», 168, pp. 81-91.
- RIZZA F., BADECK FW., CATTIVELLI L., LI DESTRI O., DIFONZO N., STANCA AM. (2004): *Use of water stress index to identify barley genotypes adapted to rainfed and irrigated conditions*, «Crop Science», 44, pp. 2127-2137.
- RONCARATI R., SALAMINI F., BARTELS D. (1995): *An aldose reductase homologous gene from barley: regulation and function*, «Plant Journal», 7, pp. 809-822.
- RUEPP A., ZOLLNER A., MAIER D., ALBERMANN K., HANI J., MOKREJS M., TETKO I., GULDENER U., MANNHAUPT G., MUNSTERKOTTER M., MEWES H.W. (2004): *The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes*, «Nucleic Acids Research», 32, pp. 5539-5545.
- SAKUMA Y., MARUYAMA K., OSAKABE Y., QIN F., SEKI M., SHINOZAKI K., YAMAGUCHI-SHINOZAKI K. (2006): *Functional Analysis of an Arabidopsis Transcription Factor, DREB2A, Involved in Drought-Responsive Gene Expression*, «Plant Cell», 18, pp. 1292-1309.
- SAKUMA Y., LIU Q., DUBOUZET J.G., ABE H., SHINOZAKI K., YAMAGUCHI-SHINOZAKI K. (2002): *DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression*, «Biochemical and Biophysical Research Communications», 290, pp. 998-1009.
- SANDERS D., BROWNLEE C., HARPER J.F. (1999): *Communicating with calcium*, «Plant Cell», 11, pp. 691-706.
- SEKI M., NARUSAKA M., ISHIDA J. ET AL. (2002a): *Monitoring the expression profiles of ca. 7000 Arabidopsis genes under drought, cold, and high-salinity stresses using a full-length cDNA microarray*, «The Plant Journal», 31, pp. 279-292.
- SEKI M., ISHIDA J., NARUSAKA M. ET AL. (2002b): *Monitoring the expression pattern of ca. 7000 Arabidopsis genes under ABA treatments using a full-length cDNA microarray*, «Functional and Integrative Genomics», 2, pp. 282-291.
- SEO M., KOSHIBA T. (2002): *Complex regulation of ABA biosynthesis in plants*, «Trends in Plant Sciences», 7, pp. 41-48.
- SHAKIROVA F.M., SAKHABUTDINOVA A.R., BEZRUKOVA M.V., FATKHUTDINOVA R.A., FATKHUTDINOVA D.R. (2003): *Changes in the hormonal status of wheat seedlings induced by salicylic acid and salinity*, «Plant Science», 164, pp. 317-322.
- SHINOZAKI K., YAMAGUCHI-SHINOZAKI K. (2007): *Gene networks involved in drought stress response and tolerance*, «Journal of Experimental Botany», 58, pp. 221-227.

- SODERMAN E., MATTSSON J., ENGSTROM P. (1996): *The Arabidopsis homeobox gene ATHB-7 is induced by water deficit and by abscisic acid*, «Plant Journal», 10, pp. 375-381.
- SUZUKI T., HIGGINS P.J., CRAWFORD D.R. (2000): *Control selection for RNA quantitation*, «Biotechniques», 29, pp. 332-337.
- TABATA T., NAKAYAMA T., MIKAMI K., IWABUCHI M. (1991): *HBP-1a and HBP-1b: leucine zipper-type transcription factors of wheat*, «EMBO Journal», 10, pp. 1459-1467.
- TAKAHASHI S., KATAGIRI T., YAMAGUCHI-SHINOZAKI K., SHINOZAKI K. (2000): *An Arabidopsis Gene Encoding a Ca^{2+} -Binding Protein is Induced by Abscisic Acid during Dehydration*, «Plant and Cell Physiology», 41, pp. 898-903.
- THELLIN O., ZORZI W., LAKAYE B., DE BORMAN B., COUMANS B., HENNE G., GRISAR T., IGOUT A., HEINEN E. (1999): *Housekeeping genes as internal standards: use and limits*, «Journal of Biotechnology», 75, pp. 197-200.
- TROVATO M., MARAS B., LINHARES F., COSTANTINO P. (2001): *The plant oncogene rolD encodes a functional ornithine cyclodeaminase*, «Proceedings of the National Academy USA», 98, pp. 13449-13453.
- UMEZAWA T., YOSHIDA R., MARUYAMA K., YAMAGUCHI-SHINOZAKI K., SHINOZAKI K. (2004): *SRK2C, a SNF1-related protein kinase 2, improves drought tolerance by controlling stress-responsive gene expression in Arabidopsis thaliana*, «Proceedings of the National Academy USA», 101, pp. 17306-17311.
- UMEZAWA T., FUJITA M., FUJITA Y., YAMAGUCHI-SHINOZAKI K., SHINOZAKI K. (2006a): *Engineering drought tolerance in plants: discovering and tailoring genes unlock the future*, «Current Opinion in Biotechnology», 17, pp. 113-122.
- UMEZAWA T., OKAMOTO M., KUSHIRO T., NAMBARA E., OONO Y., SEKI M., KOBAYASHI M., KOSHIBA T., KAMIYA Y., SHINOZAKI K. (2006b): *CYP707A3, a major ABA 8#-hydroxylase involved in dehydration and rehydration response in Arabidopsis thaliana*, «The Plant Journal», 46, pp. 171-182.
- YEUNG K.Y., RUZZO W.L. (2001): *Principal component analysis for clustering gene expression data*, «Bioinformatics», 17, pp. 763-774.
- YOSHIBA Y., KIYOSUE T., KATAGIRI T., UEDA H., MIZOGUCHI T., YAMAGUCHI-SHINOZAKI K., WADA K., HARADA Y., SHINOZAKI K. (1995): *Correlation between the induction of a gene for delta1-pyrroline-5-carboxylate synthetase and the accumulation of proline in Arabidopsis thaliana under osmotic stress*, «Plant Journal», 7, pp. 751-760.
- ZHANG J.Z., CREELMAN R.A., ZHU J.K. (2004): *From laboratory to field. Using information from Arabidopsis to engineer salt, cold, and drought tolerance in crops*, «Plant Physiology», 135, pp. 615-621.