

## Relazione – Attività di ricerca presso “Sequentia Biotech” (Center for Research in Agricultural Genomics – CRAG – Barcellona)

### Titolo del programma: Analisi bioinformatica di dati genomici e trascrittomici in grano duro

L'attività è stata impostata come continuazione del lavoro incentrato sullo studio della risposta del frumento duro a condizioni di stress azotato iniziato il precedente anno presso la medesima struttura.

La prima parte di questa ricerca ha riguardato l'analisi di dati di trascrittomica prodotti da un esperimento RNA-Seq condotto su campioni controllo e stressati prelevati in fase di riempimento della cariosside da quattro tessuti principali della varietà Svevo: radici, foglie e fusti, foglia a bandiera e spighe.

L'analisi differenziale ha portato all'individuazione di un totale di 4626 geni differenzialmente espressi (DE).

Lo scopo della seconda parte del lavoro è stata la costruzione dei network relativi ai geni codificanti per fattori di trascrizione (TF) e proteina-chinasi (PK) DE in questo nuovo dataset. Questa strategia è stata utile ad identificare nell'ampio dataset trascrittomico le interconnessioni tra il metabolismo dell'azoto e la risposta della pianta allo stress da carenza di azoto.

A questo scopo la correlazione di Pearson è stata utilizzata per costruire un grande network a livello trascrittomico di tutti i geni la cui espressione è stata rilevabile in almeno due tessuti.

I valori assoluti di correlazione sono stati filtrati in maniera molto stringente in modo da focalizzarsi sui geni fortemente interconnessi (Pearson correlation value  $\Rightarrow$   $|0.9|$ ).

Il genoma di *Triticum turgidum* subsp. *durum* non è ancora stato pubblicato. Nel database di EnsemblPlants è però disponibile la sequenza di *T. aestivum* sottoforma di pseudomolecole associate al relativo file di annotazione. Queste informazioni sono state utili ad identificare i geni all'interno del dataset di quelli differenzialmente espressi estraendo dal grande network i sub-network corrispondenti ai geni codificanti per DE-TF e DE-PK identificati ed annotati nello studio di trascrittomica.

In totale sono stati identificati 272 network corrispondenti a 109 TF e 163 PK. Per ogni network è stata condotta un'analisi di arricchimento funzionale, che ha consentito di avere un panorama dei processi biologici coinvolti nella risposta allo stress qui studiato. Il filtraggio per *enrichment score* ha permesso di evidenziare le categorie gene ontology significativamente arricchite, ed in particolare quelle direttamente associate ai *pathway* metabolici coinvolti nell'utilizzo dell'azoto. Inoltre i network sono stati filtrati focalizzandosi su quelli più arricchiti in geni DE e codificanti per TF, PK, includendo anche quelli facenti parte del metabolismo primario dell'azoto.

Queste analisi saranno fondamentali per identificare all'interno dei network i geni di interesse maggiormente implicati nell'uso dell'azoto e potenzialmente utili per il miglioramento genetico del frumento duro.

Bari,  
13/12/2016

Il fruitore,

Pasquale Luca Curci