

 **VITIS**
VALUTAZIONE DELL'ATTITUDINE DELLE UVE FALANGHINA ED AGLIANICO
COLTIVATE NEL SANNIO ALLA PRODUZIONE DI VINO SPUMANTE DI QUALITÀ

CUP: B73D09000380006

DRD n° 195 del 22/11/2013

PSR Campania 2007-2013 mis. 124 (ambito PIF)

**STUDIO DEL GENOMA, DEL TRASCRITTOMA E DEL
METABOLOMA DELL'AGLIANICO E DELLA FALANGHINA
BENEVENTANA**

Clizia Villano, Luigi Frusciante*, Domenico Carputo*, Riccardo Aversano***

*Dipartimento di Agraria, Università degli studi di Napoli Federico II, Via Università 100, 80055
Portici (NA). * Sezione di genetica e biotecnologie vegetale; ** Sezione della Vigna e del Vino.*

1. Introduzione.....	3
2. Theoretical background	5
3. Obiettivi della ricerca	7
4. Metodologia.....	7
5. Risultati e discussione.....	10
Bibliografia.....	15

1. Introduzione

La regione Campania è caratterizzata da un elevato numero di ecotipi e vitigni antichi (ultracentenari) geneticamente ed enologicamente diversi. Questa peculiare condizione è la conseguenza di elementi storici, sociali, geografici e culturali, come l'orografia della regione, la frammentazione dei terreni, l'origine vulcanica del suolo e la tradizionale propagazione asessuata tramite talee accoppiata con la coltivazione di diverse varietà in un unico vigneto. Questi aspetti rendono le uve campane capaci di dar vita ad alcuni dei migliori vini del mondo, tra cui l'Aglianico e la Falanghina. L'Aglianico è un'uva caratterizzata da un elevato contenuto di resveratrolo, polifenoli e antociani. La forte attività antiossidante e l'alto contenuto di resveratrolo conferiscono a questo vino elevate proprietà nutriceutiche positive (De Nisco et al, 2013). Inoltre, l'elevato contenuto di polifenoli, principalmente tannini, conferisce al vino un carattere aspro e astringente. Rinaldi et al. (2014) hanno effettuato una caratterizzazione preliminare delle proantocianidine di semi e bucce di Aglianico e lo hanno confrontato con cultivar internazionali come il Merlot e il Cabernet Sauvignon; i risultati hanno mostrato che l'Aglianico è un modello per astringenza. Un'altra varietà di elevato pregio enologico è la Falanghina. Essa è rinomata sia per l'equilibrio della composizione che per la perfetta armonia degli aromi e discenderebbe dall'antico "Falerno bianco". Questo vitigno è autoctono della regione Campania e viene coltivato esclusivamente sul 5% dell'intera area regionale dedicata ai vigneti, esso racchiude in sé tutta la storia centenaria di questa regione. La valorizzazione del patrimonio genetico della vite in Campania viene effettuata da anni da innumerevoli ignoti viticoltori ai quali siamo debitori per il lavoro che hanno svolto nella selezione dei vitigni e nella conservazione di tante risorse genetiche. Con l'esponenziale progresso nell'ambito del miglioramento genetico e della biologia molecolare applicata, la comunità scientifica ha ad oggi a disposizione una serie di nuove tecnologie che possono enormemente contribuire all'aumento delle conoscenze della viticoltura locale e alla sua tutela e conservazione. Una delle tecnologie maggiormente informative è la decodifica del genoma, del trascrittoma e del metaboloma.

Il genoma della vite è stato sequenziato per la prima volta nel 2007 da due consorzi indipendenti (Jaillon *et al.*, 2007; Velasco *et al.*, 2007). Le due iniziative, una italo-francese e l'altra italo-americana, hanno decodificato il genoma di due cloni di Pinot Nero, rispettivamente il PN40024, clone sperimentale non coltivato, e l'ENTAV 115, clone coltivato. I dati prodotti hanno svelato numerose caratteristiche del genoma tra cui la sua lunghezza (oltre 500 milioni di nucleotidi) e il numero di geni. Quest'ultimi, circa 30.000 nei vitigni

sequenziati, costituiscono le unità di base responsabili della trasmissione dei singoli caratteri. I genomi di Pinot Noir decodificati rappresentano i riferimenti per il sequenziamento (in questo caso si parla di *re-sequencing*) di Aglianico e Falanghina. La conoscenza delle sequenze genomiche di questi vitigni pregiati per la regione Campania aumenterebbe la comprensione del controllo molecolare e genetico che sta alla base di eventi morfologici, fisiologici e biochimici che avvengono durante lo sviluppo della pianta e dei suoi frutti. Un altro aspetto largamente approfondito è la messa a punto di studi basati sulle analisi di espressione genica. Questi studi hanno mosso i primi passi negli anni '90 con analisi limitate a singoli o pochi geni e si sono evolute rapidamente fino alle prime applicazioni dell'analisi di espressione genica su larga scala (trascrittomica) che prevede la rilevazione contemporanea dei livelli di centinaia o migliaia di trascritti di un determinato campione. Il trascrittoma consiste nell'intera collezione di molecole di RNA presenti all'interno di una cellula, di un tessuto, di un organo o di un intero organismo. Questa collezione include molecole di mRNA, che codificano proteine e quindi posseggono il maggior potenziale informativo e pertanto le applicazioni trascrittomiche hanno riguardato quasi esclusivamente l'analisi dei cambiamenti qualitativi e quantitativi proprio di mRNA, i cosiddetti profili di espressione. La caratterizzazione dei cambiamenti del trascrittoma risulta essere molto informativa poiché fornisce una visione globale riguardo ai processi molecolari associati allo sviluppo e/o alle interazioni di un organismo con l'ambiente esterno. Le analisi trascrittomiche in vite sono state principalmente rivolte allo studio del processo di sviluppo della bacca, dall'allegagione fino alla maturazione e post-maturazione, e all'influenza dell'ambiente (fattori biotici e abiotici) sullo sviluppo della bacca stessa e sulle sue caratteristiche qualitative alla raccolta. La costruzione di profili d'espressione è ad oggi un approccio sperimentale molto efficace utilizzato per studiare variazioni in termini di abbondanza di geni coinvolti nei processi biologici d'interesse. Correlazioni positive o negative possono essere utilizzate per sviluppare ipotesi riguardanti la funzione genica e per costruire modelli di network metabolici. Infatti, al fine di valutare in modo più accurato i cambiamenti che sono alla base dei processi biologici di interesse, occorre associare agli studi di espressione genica quelli di proteomica (studio di struttura, funzione, attività, quantità e interazioni delle proteine) e/o metabolomica (studio delle reazioni metaboliche che avvengono negli organismi viventi). Al pari della trascrittomica, la metabolomica è essenziale per determinare la varietà di metaboliti primari e secondari caratteristici di un organismo. I primi sono prodotti in processi biochimici essenziali per la vita dell'organismo, come ad esempio la fotosintesi, la sintesi proteica e la respirazione. I metaboliti secondari invece, non hanno un ruolo diretto nei processi essenziali della cellula,

ma hanno una funzione di difesa e di competizione interspecifica. L'obiettivo degli studi metabolomici è quello di identificare e quantificare l'intero metaboloma, definito come tutti quei composti chimici presenti in un campione biologico.

2. Theoretical background

Sono ormai passati oltre 8 anni dal primo sequenziamento genomico di *Vitis vinifera* durante i quali le tecnologie di sequenziamento sono molto progredite. In particolare ora sono disponibili strumentazioni molto veloci e ad alta processività. E' attualmente possibile "risequenziare" il genoma di un qualunque genotipo di *Vitis vinifera* con poche decina di migliaia di Euro nell'arco di una settimana. Sebbene le informazioni molecolari generate dai progetti di sequenziamento, seppure interessanti, non trovano un'immediata applicazione, esse consentiranno lo sviluppo futuro di strumenti genomici innovativi. Ad esempio, il risequenziamento di genotipi di *Vitis vinifera* può essere applicato con successo allo studio della struttura genetica del germoplasma e impiegato attivamente nei programmi di miglioramento genetico. Un altro aspetto con implicazioni applicative potenziali di grande interesse è l'uso dell'analisi del trascrittoma (espressione genica) per studiare processi fondamentali, quali la maturazione e l'appassimento della bacca, la risposta della pianta a condizioni di crescita diverse, l'interazione tra porta-innesto e innesto. Ciò contribuirà a fornire dei parametri funzionali per la caratterizzazione del binomio «vitigno-terroir». La caratterizzazione del trascrittoma di un campione può essere effettuata, ad esempio, attraverso metodologie basate sul sequenziamento (es. cDNA-AFLP, sequenziamento di librerie a cDNA, RNA-seq) che non richiedono informazioni di sequenza a priori e generano esse stesse tale informazione. Questi approcci, ad esclusione dell'RNA-seq che si basa su tecniche di sequenziamento ad elevata processività sviluppate in anni recenti sono stati molto utilizzati nella vite quando l'informazione relativa a sequenze geniche era ancora limitata, e sono stati quasi esclusivamente rivolti allo studio delle variazioni dell'espressione genica che si verificano durante la formazione (fase di crescita erbacea) e la maturazione della bacca (Ablett et al., 2000; Burger e Botha, 2004; Davies e Robinson, 2000; Moser et al., 2001; Peng et al., 2007; Terrier et al., 2001; Venter et al., 2001; Wang et al., 2012b; Zamboni et al., 2008). Ad oggi, l'RNA-seq rappresenta la metodologia analitica più accurata ed efficace per le misure di espressione genica. L'RNA-seq utilizza le recenti tecniche di sequenziamento ad elevata processività che consentono in pochi giorni di ottenere milioni di sequenze relativamente corte (poche centinaia di bp). Queste sequenze possono essere utilizzate per l'analisi di espressione determinando se e quante volte ogni sequenza viene ritrovata nella sequenza del

genoma di riferimento. Ciò che si ottiene è una precisa mappa di trascrizione su scala genomica che fornisce informazioni sulla complessità e sulle dinamiche del trascrittoma. L'RNA-seq ha anche il vantaggio di generare misure di espressione assolute non limitate dalla presenza di segnali di background, che forniscono un range di espressione molto ampio e consentono un'analisi più accurata delle differenze di espressione tra i campioni in analisi (Zenoni et al., 2012). Un'ottima correlazione di dati può essere effettuata grazie all'accoppiamento di analisi genomiche e trascrittomiche con quelle di tipo metabolomico, in quanto la vite produce migliaia di metaboliti secondari responsabili della qualità dei vini. Diversi studi proteomici e metabolomici sono stati riportati con successo negli ultimi anni per quanto riguarda la caratterizzazione di alcuni aspetti riguardanti la maturazione degli acini di vite e la risposta della pianta a stress biotici e abiotici (Genetics, genomics and breeding of grapes 2011). Ad esempio, l'utilizzo della 2DGE (two-dimensional gel electrophoresis) e delle tecniche iTRAQ (isobaric Tagging for Relative and Absolute Quantitation) abbinate alla spettrometria di massa (MS), hanno consentito di determinare la natura delle proteine accumulate nel pericarpo dell'acino durante la maturazione (Giribaldi *et al.*, 2007), nonché le variazioni qualitative e quantitative in seguito a stress abiotici (Jellouli *et al.*, 2008) e biotici (Basha *et al.*, 2010, Marsh *et al.*, 2010). Al pari della trascrittomica (studio dell'espressione dei trascritti di un organismo), la metabolomica è essenziale per determinare la varietà di metaboliti primari e secondari caratteristici di un organismo. I primi sono prodotti in processi biochimici essenziali per la vita dell'organismo, come ad esempio la fotosintesi, la sintesi proteica e la respirazione. I metaboliti secondari invece non hanno un ruolo diretto nei processi essenziali della cellula, ma hanno una funzione di difesa e di competizione interspecifica. Possiamo dividere quest'ultima classe di metaboliti in terpeni, fenoli, alcaloidi, acidi organici e altri. In vite, essi sono circa 1.600, praticamente due terzi dei 2.314 presenti nel Plant Metabolic Network Database (<http://www.plantcyc.org/>). L'obiettivo degli studi metabolomici è quindi quello di identificare e quantificare l'intero metaboloma, definito come tutti quei composti chimici presenti in un campione biologico. Diversi risultati si sono ottenuti in vite, tra cui l'identificazione di più di 150 metaboliti in un singolo campione attraverso l'utilizzo della GC/MS (Gas Chromatography Mass Spectrometry) (Fenoll *et al.*, 2009), oppure lo studio del profilo di metaboliti primari e secondari usando la risonanza magnetica nucleare (NMR) (Son *et al.*, 2009). Comunque, le tecniche applicate alla proteomica ed alla metabolomica in pianta sono in continua evoluzione, offrendo protocolli d'estrazione sempre più efficienti e metodi di rilevamento sempre più sensibili.

3. Obiettivi della ricerca

Tra i vitigni maggiormente coltivati e utilizzati per la produzione di vini a denominazione DOCG, spiccano l'Aglianico del Taburno e la Falanghina del Beneventano. Al fine di poter tutelare e valorizzare questi vitigni, gli obiettivi del presente lavoro sono stati:

- 1) Il sequenziamento dei loro genomi e l'analisi delle variazioni genetiche strutturali;
- 2) L'analisi del trascrittoma delle varietà Aglianico del Taburno e Falanghina Beneventana in tre differenti tessuti dell'acino (buccia, polpa e vinaccioli) e in cinque stadi di sviluppo (allegagione, pre-invaiatura, invaiatura, maturazione e raccolta) e lo studio delle vie metaboliche di particolare interesse dal punto di vista enologico, riproduttivo, e di risposta a stress biotici ed abiotici.
- 3) lo studio del metaboloma al fine di rilevare nuovi composti (volatili e non-volatili) in gradi di accumularsi nella bacca e dotati di proprietà importanti un ruolo importante dal punto di vista enologico.

4. Metodologia

Analisi del genoma di Aglianico e Falanghina

Il materiale vegetale utilizzato per l'estrazione del genoma è costituito da foglie raccolte da piante micropropagate *in vitro* per evitare contaminazioni da agenti esterni. Al fine di aumentare la resa del DNA estratto sono state scelte foglie giovani con minor quantitativo di strutture cellulari complesse. Il materiale raccolto è stato macinato in azoto liquido con l'ausilio di mortaio e pestello ed è stato poi utilizzato per testare diversi protocolli di estrazione, tra i quali quello riportato in Japelaghi e collaboratori (2011) è risultato il migliore in resa e qualità di DNA estratto. Dai campioni estratti delle varietà di *Vitis vinifera* Falanghina e Aglianico sono state prodotte le sequenze di DNA utilizzando la tecnologia Illumina. Quest'ultima fa parte delle tecnologie di sequenziamento di ultima generazione che prendono il nome di Next Generation Sequencing (NGS). Le fasi principali della tecnologia utilizzata prevedono la frammentazione del DNA genomico, l'aggiunta degli adattatori necessari per la formazione dei gruppi di sequenze (detti cluster) e il finale sequenziamento degli stessi. In base al tipo di adattatori utilizzati viene effettuato un diverso tipo di generazione e sequenziamento di cluster di sequenze.

Nel presente studio, il genoma di Aglianico e Falanghina è stato sequenziato a partire da librerie di Illumina dette pair-end (come riportato nella Figura 1). L'insieme di cluster di sequenze paired-end è stato diviso in due gruppi dette librerie contenenti rispettivamente 137'000'000 di paired-end con lunghezza media del frammento di 400bp e 77'000'000 di

paired-end con lunghezza media del frammento di 700bp. A questo punto le librerie create sono state assemblate al fine ultimo di ottenere un intero genoma ossia effettuare una ricostruzione della sequenza del DNA in una serie continua e univoca a partire da sequenze relativamente brevi. Questa fase del procedimento di determinazione della sequenza di un genoma complesso come quello della vite prende il nome di assemblaggio della sequenza. Ciò è reso possibile dal fatto che sequenziando un numero così elevato di frammenti casuali capita di sequenziare frammenti che contengano porzioni in comune e il riconoscimento di queste porzioni di sequenza comune, in termini tecnici sovrapposizioni, consente di ricostruire finalmente la sequenza completa. L'assemblaggio è reso possibile dall'applicazione di programmi bioinformatici sofisticati, in grado di gestire questa enorme quantità di dati. Nel presente studio, le sequenze sono state assemblate con il software SOAPdenovo127kmer1. Una volta disponibile una sequenza assemblata di qualità, incomincia il lungo lavoro di interpretare i dati di sequenza in termini di geni, di sequenze spaziatrici, di elementi di regolazione, ovvero si procede alla cosiddetta annotazione.

Analisi del trascrittoma di Aglianico e Falanghina

Per le analisi del trascrittoma, il materiale è stato raccolto in tre repliche biologiche dai vigneti facenti parte della proprietà della Cantina "La Guardiense", situato a Guardia Sanframondi (provincia di Benevento). Esso è costituito da bacche di Aglianico del Taburno e Falanghina del Beneventano campionate in cinque fasi di sviluppo e seguendo uno schema di campionamento random. Gli stadi di maturazione sono stati definiti secondo il metodo DAF (*Days After Flowering*), che consiste nel conteggio dei giorni dopo la fioritura. La prima fase (15 giorni dopo la fioritura [DAF]; Eichhorn-Lorenz Sistema [EL] 29) corrisponde all'allegagione, quando le giovani bacche iniziano ad aumentare di volume (> 3 mm di diametro); la seconda fase (35 DAF, EL 32) è la pre-invaiatura, quando le bacche cominciano a toccarsi tra loro (> 7 mm); la terza fase (70 DAF, EL 35) è l'invaiatura, quando le bacche cominciano a cambiare colore; la quarta tappa (84 DAF, EL 36) corrisponde alla maturazione; e la fase finale (115 DAF, EL 38) è rappresentata dalla raccolta. Dopo la raccolta, tutti i frutti a partire dall'allegagione sono stati sezionati manualmente in buccia, polpa e semi e macinati sino a ottenere una polvere fine. Questo passaggio è fondamentale per ottenere risultati di buona qualità. Il materiale è stato utilizzato per testare diversi protocolli di estrazione, tra i quali quello riportato in Japelaghi e collaboratori (2011) è risultato il migliore in resa e qualità di RNA estratto. La lista dei campioni utilizzati per l'estrazione dell'RNA è riportata in Tabella 2.

Per l'analisi del trascrittoma è stata utilizzata la tecnica di RNA-Seq (RNA sequencing) tramite la tecnologia Illumina, descritta in precedenza. I dati ottenuti dal sequenziamento sono stati elaborati in sinergia con l'azienda SEQUENTIA in quanto hanno richiesto un'ingente capacità di calcolo. Di seguito sono descritte le analisi effettuate per ognuno dei campioni:

- **Rimozione degli adattatori e controllo qualitativo/composizionale:** partendo dal dato grezzo del sequenziamento del trascrittoma sono state rimosse le sequenze degli adattatori impiegate per la costruzione delle librerie di sequenziamento. In concomitanza a questa analisi, prima e dopo la rimozione degli adattatori, è stato effettuato un controllo qualitativo del dato e una successiva selezione dei frammenti lunghi almeno 35 paia di basi, e con un alto indice di qualità:
- **Mappaggio contro il genoma di riferimento:** i frammenti selezionati sono stati allineati contro il genoma di riferimento di *Vitis vinifera* (versione 2.1) ed è stato determinato il numero (count) di frammenti (read) allineati ad ogni locus specifico del genoma di riferimento;
- **Analisi statistiche e individuazione dei geni differenzialmente espressi:** mediante l'ambiente software di biostatistica *R* (versione 3.1.2) è stato calcolato il grado di similarità tra tutti i campioni, è stato normalizzato il dato di espressione genica tra le repliche biologiche di ogni tessuto specifico in un determinato stadio e per i campioni provenienti da buccia e polpa è stato effettuato un confronto di due campioni per volta, con tutte le combinazioni possibili. In questo modo sono stati individuati i geni che mostrano una variazione, positiva o negativa, nel livello d'espressione.
- **Annotazione funzionale e analisi di arricchimento del termine GO:** ai geni differenzialmente espressi sono state aggiunte le informazioni biologiche note e sono stati identificati i termini GO più significativi, ovvero maggiormente ricorrenti tra i geni differenzialmente espressi individuati nei confronti precedenti.
- **Identificazione del numero di trascritti unici totali in buccia e polpa durante lo sviluppo:** utilizzando un dataset con l'elenco dei loci unici identificati ed il valore d'espressione normalizzato in RPKM (Reads Per Kilobase of gene per Milion of mapped reads), sono stati calcolati i valori medi di espressione tra i replicati biologici. Perdi più, sono stati individuati i loci per geni espressi unicamente in una specifica fase della maturazione di un determinato tessuto..
- **Analisi quantitativa e heatmap dei geni differenzialmente espressi:** i dati di tutti i geni differenzialmente espressi durante le diverse fasi fenologiche per la buccia e per la polpa sono stati integrati e visualizzati sotto forma di heatmap.

Analisi del metaboloma di Aglianico e Falanghina

Per una completa analisi del metaboloma di bacche di vite di Aglianico e Falanghina sono stati utilizzati i campioni riportati in Tabella 1. Il suddetto materiale è stato liofilizzato e sottoposto

ad un'analisi a livello biochimico-metabolico tramite l'uso di tecnologia LC-MS a elevata risoluzione al fine di individuare i metaboliti primari (amminoacidi, zuccheri, acidi organici, acidi grassi, etc) e secondari (flavonoidi, carotenoidi, etc) caratterizzanti le diverse varietà. Simultaneamente, sono stati effettuati studi di metabolomica con la tecnologia GC-MS ad elevata risoluzione al fine di definire i complessi di metaboliti volatili (C13-nor-isoprenoidi, terpeni, tioli etc) prodotti nel corso della maturazione dei differenti vitigni.

5. Risultati e discussione

Sono state sequenziate librerie *paired-end* con lunghezza media del frammento di 400-700bp.

Sono stati effettuati diversi assemblaggi di prova e il migliore è risultato essere quello con:

- numero di sequenze assemblate: 1'209'171
- lunghezza media delle sequenze assemblate: 745.9bp
- lunghezza minima/massima dei contig: 100/648'344bp

La lunghezza totale dell'assemblaggio è risultata pari al 93% delle dimensioni del genoma di riferimento sequenziato di *Vitis vinifera* (Pinot Noir clone PN40024). Ciò suggerisce che gli assemblaggi ottenuti si avvicinano con buona approssimazione alle dimensioni totali attese. Successivamente sono state realizzate le analisi bionformatiche volte a valutare la composizione di sequenze ripetute dei genomi, classificare gli elementi trasponibili dispersi, valutare l'impatto di quest'ultimi sulla variazione della dimensione del genoma, stimare l'eterozigotità e predire le varianti geniche e i loro effetti.

Mediante il sequenziamento del trascrittoma (RNA-seq) si possono determinare le differenze di espressione genica nei vari tessuti e nei diversi stadi di sviluppo riportati nella Tabella 2. Come già detto, le procedure di acquisizione dati ed analisi risultano piuttosto complesse. I dati ottenuti dal sequenziamento sono detti *count* e rappresentano il numero di volte che ciascun frammento è letto: va da sé che il valore dei *count* rappresenta una misura del livello di espressione del gene. Prima di poter calcolare il livello di espressione è necessario effettuare una selezione dei frammenti in base alla loro qualità e successivamente allineare i frammenti migliori contro il genoma di riferimento di *Vitis vinifera*; questa procedura prende il nome di mappaggio e ne è riportato un esempio nella figura 2. Nella Figura 1 è riportato un grafico a torta rappresentante la qualità degli allineamenti del campione AAL (bacca intera di Aglianico del Taburno in allegazione) contro il genoma di riferimento. Nel suddetto grafico è possibile notare che il 91% degli allineamenti ha un valore dell'indice della qualità del mappaggio (MAPQ) superiore o uguale a 30, mentre lo 0,1% delle read non è stato mappato.

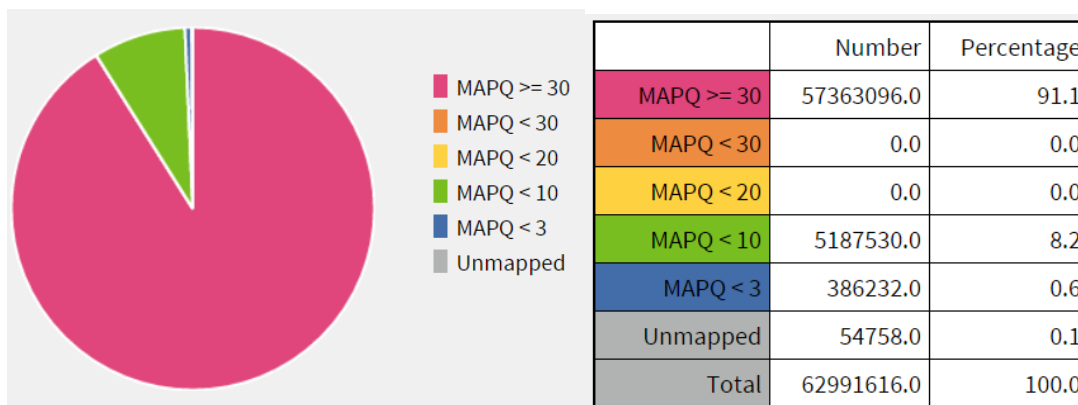


Figura 1. Output grafico di SamStat per il campione AAL_1.

Per ogni confronto effettuato, il livello d'espressione dei geni differenzialmente espressi è stato ulteriormente normalizzato come Log FC (Fold Change su scala logaritmica). L'annotazione funzionale ha permesso di aggiungere specifiche informazioni biologiche ai geni differenzialmente espressi, e di sottoporle alle successive analisi di arricchimento del termine GO. Quest'ultime hanno permesso di categorizzare e interpretare il ruolo funzionale dei trascritti con un ruolo comune. A fronte di 26826 loci unici totali individuati in buccia, e di 26878 individuati in polpa, è stato rappresentato graficamente il loro andamento in buccia durante le diverse fasi fenologiche e lo stesso dato è stato ulteriormente scomposto per permettere una correlazione visiva tra il livello d'espressione e il numero di loci unici identificati (Figura 2).

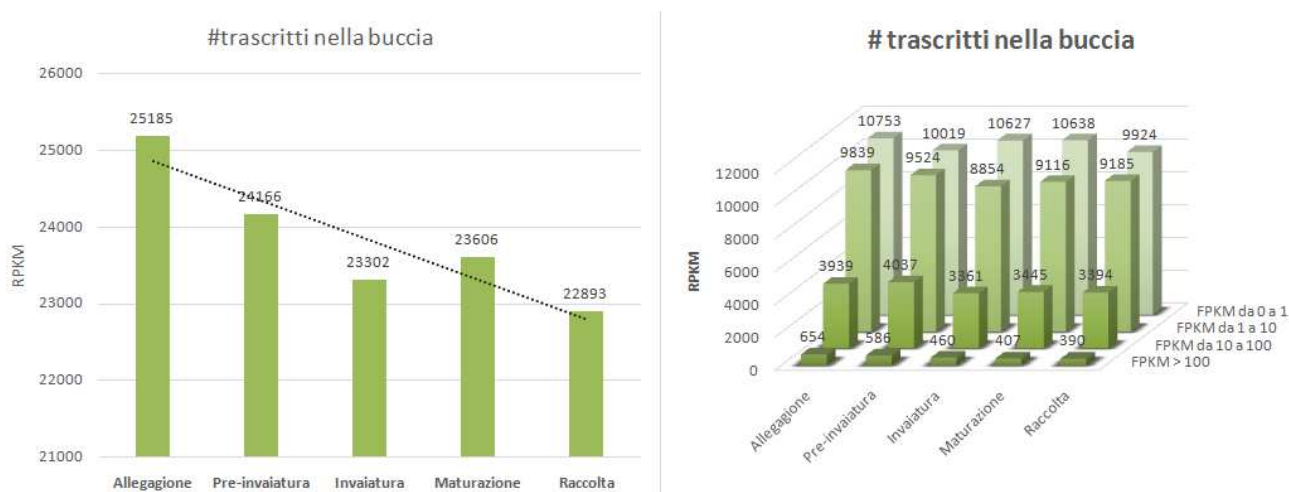


Figura 2. Numero di trascritti unici presenti nei vari stadi durante lo sviluppo delle buccia. Nel grafico di destra è riportato lo stesso dato ma in relazione al livello d'espressione normalizzato.

L'analisi Venn per buccia e polpa ha permesso d'individuare i geni comuni e i fase-specifici (Fig. 3). In particolare: nella la buccia sono presenti 21109 geni costitutivi, 284 geni espressi solo in pre-invaiatura, 141 geni espressi solo in invaiatura, 232 geni espressi solo in maturazione, e 137 geni espressi solo in raccolta; nella la buccia sono presenti 21025 geni costitutivi, 733 geni espressi solo in pre-invaiatura, 115 geni espressi solo in invaiatura, e 209 geni espressi solo in raccolta.

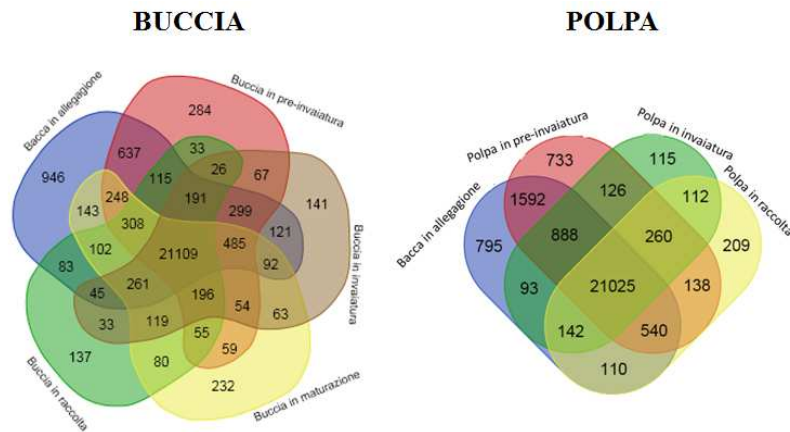


Figura 3. Diagramma di Venn rappresentante il numero di trascritti in buccia e polpa in correlazione con i diversi stadi di sviluppo.

Il numero complessivo dei geni differenzialmente espressi durante lo sviluppo di buccia e polpa è stato riportato nella Tabella 1, dove sono anche riportati i DEGs con $-1,5 \leq \log FC \leq 1,5$.

Tabella 1. Numero complessivo di geni differenzialmente espressi in buccia e polpa e per un Log FC maggiore di 1,5 e minore di -1,5.

Geni Differenzialmente Espressi	Complessivi		Con un Log FC > 1,5 e < -1,5	
	UP	DOWN	UP	DOWN
BUCCIA				
Allegagione - Pre-invaiatura	6451	6454	1609	1888
Pre-invaiatura - Invaiatura	6440	6411	1337	2433
Invaiatura - Maturazione	5017	5091	543	486
Maturazione - Raccolta	1649	1107	7	60
POLPA				
Allegagione - Pre-invaiatura	6451	6454	1289	1642
Pre-invaiatura - Invaiatura	2433	6411	1232	2691
Invaiatura - Raccolta	5017	5091	76	742

Analisi del metaboloma di Aglianico e Falanghina

Per quanto concerne il metaboloma, le analisi LC-MS e GC-MS hanno permesso di definire i profili di accumulo di circa 400 metaboliti delle frazioni polari/apolari e volatili. I dati hanno evidenziato differenze significative nella composizione in composti volatili tra i due vitigni e, nell'ambito di ciascuno di essi, tra i cinque stadi di sviluppo della bacca. In generale le prime fasi di maturazione sono state caratterizzate da alti livelli di alcuni carotenoidi e apocarotenoidi, come ad esempio il β -ciclocitrato e il β -ionone, terpenoidi, (E)- e (Z) linalolo ossido e diversi furani, mentre le fasi finali della maturazione sono caratterizzati da un elevato quantitativo di etanolo, benzenoidi, fenilacetaldeide, 2-feniletanolo, il 3-metilbutanolo e il 2-metilbutanolo, oltre ad un elevato numero di derivati dei lipidi. In aggiunta, sono stati osservati numerosi metaboliti vitigno-specifici, che potrebbero contribuire in maniera significativa a contraddistinguere i profili aromatici distintivi dei due vitigni.

Conclusioni

Grazie alle attività di ricerca condotte nell'ambito del progetto VITIS è stato possibile decodificare per la prima volta il genoma, il trascrittoma e il metaboloma dell'Aglianico del

Taburno e della Falanghina Beneventana. Grazie a questo studio è stato possibile accedere ai segreti più profondi dei due vitigni e comprendere il programma genetico che concorre alla loro tipicità. La sfida per il prossimo futuro sta nell'integrazione dei dati prodotti, di fondamentale importanza per definire gli elementi alla base delle proprietà organolettiche dei vini. Inoltre, attraverso la conoscenza del genoma, del trascrittoma e del metaboloma, sarà possibile studiare le relazioni tra geni e metaboliti di una bacca e studiarne le variazioni legate alla stagionalità, all'ambiente e all'opera dell'uomo.

Bibliografia

1. FAOSTAT (2013) FAO statistical databases, production statistics.
2. Mipaaf: <http://www.politicheagricole.it>
3. Zenoni S., Zamboni A., Dal Santo S., Fasoli M., Pezzotti M., Tornielli G.B. (2012) Lo sviluppo delle conoscenze genomiche in vite e il loro potenziale utilizzo nella viticoltura attuale e future. Review n. 17 Italus Hortus 19 (2):29-40;
4. Rinaldi, A., Jourdes, M., Teissedre, P.L. and Moio, L. (2014) A preliminary characterization of Aglianico (*Vitis vinifera* L. cv.) grape proanthocyanidins and evaluation of their reactivity towards salivary proteins. Food Chem 164, 142-149;
5. Ablett E., Seaton G., Scott K., Shelton D., Graham M.W., Baverstock P., Lee L.S., Henry R., 2000. Analysis of grape ESTs: global gene expression patterns in leaf and berry. Plant Sci. 159: 87-95.
6. Burger A.L., Botha F.C., 2004. Ripening-related gene expression during fruit ripening in *Vitis vinifera* L. cv. Cabernet Sauvignon and Clairette blanche. Vitis 43: 59-63.
7. Davies C., Robinson S.P., 2000. Differential screening indicates a dramatic change in mRNA profiles during grape berry ripening. Cloning and characterization of cDNAs encoding putative cell wall and stress response proteins. Plant Physiol. 122: 803-812.
8. Peng F.Y., Reid K.E., Liao N., Schlosser J., Lijavetzky D., Holt R., Martinez Zapater J.M., Jones S., Marra M., Bohlmann J., Lund S.T., 2007. Generation of ESTs in *Vitis vinifera* wine grape (Cabernet Sauvignon) and table grape (Muscat Hamburg) and discovery of new candidate genes with potential roles in berry development. Gene 402: 40-50.
9. Terrier N., Ageorges A., Abbal P., Romieu C., 2001. Generation of ESTs from grape berry at various developmental stages. J Plant Physiol 158: 1575-1583.
10. Terrier N., Glissant D., Grimplet J., Barrieu F., Abbal P., Couture C., Ageorges A., Atanassova R., Leon C., Renaudin J.P., Dedaldechamp F., Romieu C., Delrot S., Hamdi S., 2005. Isogene specific oligo arrays reveal multifaceted changes in gene expression during grape berry (*Vitis vinifera* L.) development. Planta 222: 832-847.
11. Zamboni A., Minoia L., Ferrarini A., Tornielli G.B., Zago E., Delledonne M., Pezzotti M., 2008. Molecular analysis of post-harvest withering in grape by AFLP transcriptional profiling. J. Exp. Bot. 59: 4145-4159.
12. Venter M., Burger A.L., Botha F.C., 2001. Molecular analysis of fruit ripening: The identification of differentially expressed sequences in *Vitis vinifera* using cDNA-AFLP technology. Vitis 40: 191-196.
13. Wang X.C., Guo L., Shangguan L.F., Wang C., Yang G., Qu S.C., Fang J.G., 2012b. Analysis of expressed sequence tags from grapevine flower and fruit and development of simple sequence repeat markers. Mol Biol Rep.

14. De Nisco M, Manfra M, Bolognese A, Sofò A, Scopa A, Tenore GC, Pagano F, Milite C, Russo MT (2013) Nutraceutical properties and polyphenolic profile of berry skin and wine of *Vitis vinifera* L. (cv. Aglianico) Food Chem. 140 (4):623-9.
15. Ergul A, Perez-Rivera G, Soybelezoglu G, Kazan K and Arroyo-Garcia R. (2011) Genetic diversity in Anatolian wild grapes (*Vitis vinifera* subsp. *sylvestris*) estimated by SSR markers. Plant Genetic Resources. 9, 375-383.
16. Wang Z, Gerstein M, Snyder M: RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 2009, 10:57-63.
17. Venturini L, Ferrarini A, Zenoni S, Tornielli GB, Fasoli M, Dal Santo S, Minio A, Buson G, Tononi P, Zago ED, Zamperin G, Bellin D, Pezzotti M, Delledonne M: *De novo* transcriptome characterization of *Vitis vinifera* cv. Corvina unveils varietal diversity. BMC Genomics 2013, 14:41.
18. Zenoni S, Ferrarini A, Giacomelli E, Xumerle L, Fasoli M, Malerba G, Bellin D, Pezzotti M, Delledonne M: Characterization of transcriptional complexity during berry development in *Vitis vinifera* Using RNA-Seq. Plant Physiol 2010, 152:1787-1795.
19. Fasoli M, Dal Santo S, Zenoni S, Tornielli GB, Farina L, Zamboni A, Porceddu A, Venturini L, Bicego M, Murino V, Ferrarini A, Delledonne M, Pezzotti M: The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program. Plant Cell 2012, 24:3489-3505.
20. Sweetman C, Wong DC, Ford CM, Drew DP: Transcriptome analysis at four developmental stages of grape berry (*Vitis vinifera* cv. Shiraz) provides insights into regulated and coordinated gene expression. BMC Genomics 2012, 13:691.
21. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013).
22. Lassmann, T., Hayashizaki, Y. & Daub, C. O. SAMStat: monitoring biases in next generation sequencing data. Bioinformatics 27, 130–131 (2011).
23. Liao, Y., Smyth, G. K. & Shi, W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Research 41, e108–e108 (2013).
24. Li, R., et al. (2010) De novo assembly of human genomes with massively parallel short read sequencing, Genome Res, 20, 265- 272.