

ANALISIS BIOINFORMATICO DE DATOS DE EXPRESION GENICA GLOBAL DURANTE EL DESARROLLO DE LA SEMILLA DE *Vitis vinifera*

**MINDY STEPHANIA DE LOS ANGELES MUÑOZ MIRANDA
INGENIERO EN BIOINFORMÁTICA**

RESUMEN

Entender los cambios que ocurren durante el desarrollo de la semilla en *Vitis vinifera* es importante para la industria agrícola en Chile. La *Vitis* es el principal cultivo para la industria de exportación de fruta. En este trabajo se analizó la expresión génica global durante el desarrollo de la semilla de *Vitis vinifera* en descendientes de la cruce de Moscatel Rosado y Ruby seedless, en dos estadios de desarrollo; flor abierta y fruto cuajado, con el objetivo de identificar genes que son importantes para el proceso de desarrollo y que pudieran utilizarse en la generación de nuevas variedades apirénicas. La progenie de los cruces desarrollaron bayas con semillas y con rudimentos seminales, las cuales fueron estudiadas durante 4 años para identificarlas según su fenotipo. Se aisló ARN total de flores de plantas crecidas en condiciones de campo al momento de la polinización y posteriormente de frutos a los 20 días post-polinización. El ARN total se utilizó para hibridar las micromatrices de Affymetrix *Vitis vinifera* GeneChip®. Los datos crudos fueron normalizados utilizando el método de RMA implementado en la función `rma()` del paquete “`affy`” de Bioconductor. Para identificar genes expresados diferencialmente se utilizó estadístico-t empírico bayesiano, implementado en el paquete “`limma`” de Bioconductor. Para controlar por error tipo I, se ajustó la tasa de descubrimiento de falsos positivos a un valor de 5%. Este análisis identificó 891 genes inducidos y 702 genes reprimidos de *Vitis vinifera*, frente a dos condiciones, “rudimento seminal” contra “semilladas”, en dos períodos de tiempo de acuerdo al diseño experimental de este trabajo. El análisis de la función de los genes con expresión diferencial mediante la clasificación de Ontologías Génicas agrupó al 70% de los genes en las categorías funcionales de proceso celular y metabólico, parte de la célula, actividad catalítica y unión. El análisis de grupo jerárquico aglomerativo con enlace promedio de los genes con expresión diferencial identificó 14 conjunto de genes, que permitieron reunir a los genes que presentaban similar patrón de expresión destacando los agrupamientos 3 y 5 que contienen 3 genes importantes para el proceso de desarrollo de la semilla, *VvPI*, *VvAGL11* y *VvPMEI2*. Estos resultados contribuyen a comprender los procesos moleculares que ocurren durante el desarrollo de la semilla en *Vitis vinifera*.

ABSTRACT

Understanding the changes that occur during seed development in *Vitis vinifera* is important for the agricultural industry in Chile. *Vitis* is the main crop for the fruit export industry. In this work, we analyzed global gene expression during seed development in the progeny of Moscatel Rosado and Ruby seedless, in two stages of development: open flower and fruit set. The aim of this project is to identify candidate genes that are important for seed development and could be used for the generation of new seedless varieties. The Progeny of the crossing developed berries with seeds and seminal rudiments, which were studied for four years to classify them according to their phenotype. Total RNA was isolated from the flowers collected from plants grown under a countryside environment at the time of the pollination and then from fruits 20 days after the pollination. The total RNA was used to hybridize the *Vitis vinifera* Affymetrix GeneChip® micromatrix. The raw data was normalized using the RMA method implemented in the `rma()` function from the “affy” package of Bioconductor. The empirical bayes t-statistic implemented in the “limma” package from Bioconductor was used to identify differentially expressed genes. To control type I error, the false discovery rate was adjusted at a 5% cutoff. This analysis identified 891 induced genes and 702 repressed, against two conditions, seminal rudiments to normal seeds in two time periods according to experimental design. Functional analysis of the genes with differential expression using the Gene Ontology classification scheme grouped 70% of the genes in the functional categories cellular process, metabolic process, cell part, catalytic activity and binding. Agglomerative hierarchical clustering analysis with average linkage of the differentially expressed genes identified 14 groups of genes that shared similar expression profiles. Clusters 3 and 5 were interesting as they contained 3 important genes for seed development processes, VvPI, VvAGL11 and VvPMEI2. These results will contribute to understand the molecular processes happening during seed development in *Vitis vinifera*.