

---

**MAPA GLOBAL DE LA REGULACIÓN DE FACTORES DE TRANSCRIPCIÓN  
TIPO MADS-BOX EN FRUTILLA SILVESTRE (*Fragaria vesca*).**

**GONZALO ARTURO SEPÚLVEDA HERMOSILLA  
INGENIERO EN BIOINFORMÁTICA**

**RESUMEN**

Las interacciones entre proteínas y DNA son particularmente conservadas a lo largo de la evolución. El estudio de estas redes de interacción pueden dar pie al entendimiento de cómo se regulan los principales procesos celulares. Los factores de transcripción tipo MADS-BOX han sido descritos como reguladores maestros del desarrollo de flores y su arquitectura en angiospermas. Algunos de sus miembros como: APETALLA, AGAMOUS, FLOWERING LOCUS y, los pertenecientes a la subfamilia SEPALLATA entre otros, presentes en tomate, manzana y frutilla comercial; juegan un rol clave en los procesos de maduración del fruto. *Fragaria vesca*, cuyo genoma ha sido recientemente secuenciado y a diferencia de otras especies de este género, como la mencionada frutilla comercial (*Fragaria x ananassa*) y la frutilla chilena (*Fragaria chiloensis*), las cuales son altamente complejas a nivel genético, es una especie diploide lo que facilita su uso como modelo de estudio a nivel genético y fisiológico. El objetivo de esta investigación, es determinar los posibles genes regulados transcripcionalmente por los factores tipo MADS-BOX en *Fragaria vesca*, estableciendo la presencia y densidad de su correspondiente sitio de unión (CArG-BOX) dentro de las secuencias de 2Kb río arriba de cada uno de los 24.000 genes anotados en este genoma. Esto se realizó mediante la utilización de la suite para el descubrimiento y búsqueda de patrones MEME. Los resultados basados en los nueve miembros de la familia MADS-Box analizados, muestran un total de 5.828 regiones promotoras con presencia de sitios de respuesta afines, que se distribuyen de manera uniforme a lo largo de la región promotora a través de 64 secuencias de unión no siempre específicas para cada factor de transcripción. Dicha inespecificidad concuerda con el mecanismo de unión al DNA a través de multímeros, propio de esta familia. Los genes putativamente regulados por MADS-Box determinados en la presente investigación participan principalmente en procesos metabólicos con actividad catalítica o de unión, aunque se destacan 90 genes clasificados funcionalmente relacionados a desarrollo y 22 con rol de

factores de transcripción. se encontraron tres enzimas que regulan el inicio de las cascadas metabólicas en la síntesis de antocianinas y fenilpropanoides. Así como otras que participan en la degradación de aminoácidos esenciales como valina, leucina e isoleucina. Dado los presentes resultados, se revela la importancia de la regulación de los factores de transcripción en dichos procesos metabólicos, los cuales tienen un rol directo en el cómo los frutos ejecutan los procesos que determinan su color (o la falta de este) durante su periodo de maduración. Se proyecta, la validación experimental de los elementos de unión predichos a través de modelos estructurales e interacción DNA-Proteína *in silico*. Esta validación sería el punto de partida para comprobar que dichos elementos en regulatorios, son funcionales y responsables de que factores de transcripción de la familia MADS-box regulen efectivamente dichos genes blancos.

## ABSTRACT

Interactions between proteins and DNA are particularly conserved throughout evolution. The study of these interaction networks can unravel the key cellular processes. The MADS-box transcription factors have been described as master regulators of flower development in angiosperms and in its architecture, in which some of its members as APETALLA, AGAMOUS, FLOWERING LOCUS and, according to the results found in tomato, apple and commercial strawberry, those belonging to the subfamily SEPALLATA, among others, play a key role in the process of fruit ripening. *Fragaria vesca*, whose genome has recently been sequenced and, unlike other species of this genus as the mentioned commercial strawberry (*Fragaria ananassa*) and the Chilean strawberry (*Fragaria chiloensis*), it is a diploid species which facilitates its use as a model organism in genetics and at a physiological level.

The objective of this research is to identify potential genes transcriptionally regulated by MADS-box transcription factors in *Fragaria vesca*, establishing the presence and density of its corresponding binding site, called CArG-box within 2Kb upstream sequences of each one of the 24,000 annotated genes in this genome. This was done using the suite for search and discovery of patterns MEME.

The results based on the nine types of MADS-Box analyzed, which were defined based on information present in the current literature, shows a total of 5828 promoter regions with presence of at least one related response site, which are distributed uniformly along promoter regions through 64 specific sequences that are not always specific (shared) for each transcription factor, which matches with the DNA binding mechanism through self-multimers, common on this family.

The genes putatively regulated by MADS-Box identified in this research are mainly involved in metabolic processes as binding or catalytic activity, although we also found 90 genes functionally related to development and 22 as transcription factors. Interestingly three enzymes which regulate the start of the metabolic cascade in anthocyanin and phenylpropanoid synthesis has been also found, as well as part of the breakdown of essential amino acids such as valine, leucine and isoleucine. Given these results, it's clear the importance of regulation of transcription factors in these metabolic processes, which have a direct role in how fruits acquire its color during the ripening period.

---

As further work, it will be necessary to structurally model the binding sites candidates in order to measure the affinity of DNA-protein interaction to determine their viability and establish a starting point for the experimental validation of the best candidates here described