

## II - Tabla de contenido

I - Agradecimientos .....	2
II - Tabla de contenido.....	3
III - Índice de Figuras.....	6
IV - Índice de tablas.....	7
1 Resumen .....	8
2 Abstract.....	10
3 Introducción. ....	11
3.1 Regulación genética a través de Factores de Transcripción.....	11
3.1.1 La familia de factores de transcripción MADS-BOX. ....	11
3.2 Funciones de MADS-Box en organismos vegetales .....	15
3.2.1 Desarrollo floral.....	15
3.2.2 MADS-BOX durante la maduración de frutos: modelos climatéricos y no-climatéricos.....	16
3.3 Métodos experimentales para el estudio de interacción de DNA-Proteínas.....	18
3.3.1 Inmuno precipitación de cromatina (ChIP).....	18
3.3.2 CHIP-Sequencing.....	19
3.4 Herramientas y métodos para búsqueda de elementos en CIS de plantas. ....	19
3.4.1 Modelos HMM:.....	20
3.4.2 Suite MEME.....	20
3.4.3 BLAST: .....	21
3.5 Índices de evaluación estadística.....	21
3.5.1 E-value: .....	21
3.5.2 P-value: .....	22
3.5.3 Q-Value: .....	22
3.6 Lenguajes de Programación .....	23

3.6.1	Perl & BioPerl: .....	23
3.6.2	AWK: .....	23
3.7	Categorización funcional .....	24
3.7.1	Gene Ontology:.....	24
4	Planteamiento del problema .....	26
5	Hipótesis de trabajo .....	27
6	Objetivo general.....	27
7	Objetivos específicos .....	27
8	Materiales y Métodos.....	28
8.1	Base de datos de regiones promotoras.....	28
8.2	Definición de los sitios de unión a MADS-Box (CArG-Box).....	29
8.3	Búsqueda sobre la base de datos de promotores de <i>Fragaria vesca</i> .	30
8.4	Metodología para el análisis de resultados .....	30
8.4.1	Análisis de secuencias palíndromas.....	31
8.4.2	Anotación por homología contra la base de datos NR.....	31
8.4.3	Validación de predicción por referencia biológica.....	32
9	Resultados.....	34
9.1	Base de datos biológica .....	34
9.2	Determinación de los motivos CArG-Box.....	35
9.2.1	Búsqueda de secuencias.....	35
9.2.2	Alineamiento de secuencias y determinación de CArG-Box.....	41
9.2.3	Validación <i>in silico</i> de los motivos CArG-Box encontrados.....	43
9.3	Búsqueda de los motivos en los promotores de <i>Fragaria vesca</i> .	43
9.4	Análisis y cuantificación de los resultados .....	45
9.4.1	Promotores con presencia de sitios de unión .....	45
9.4.2	Sitios de unión CArG-Box encontrados. ....	45
9.4.3	Secuencias palíndromas .....	45

9.4.4	Análisis de secuencias promiscuas .....	46
9.4.5	Secuencias de unión más frecuentes .....	47
9.4.6	Ubicación de los sitios de unión dentro de la región promotora ...	48
9.5	Anotación por homología de secuencias.....	49
9.6	Categorización funcional con Gene Ontology .....	50
9.6.1	Proceso biológico: .....	50
9.6.2	Función Molecular: .....	51
9.6.3	Componente Celular:.....	53
9.7	Reconstrucción de rutas metabólicas.....	54
9.8	Validación por referencia biológica.....	57
10	Discusión .....	58
11	Conclusión.....	68
12	Referencias:.....	70
13	Anexos.....	76
13.1	Presentación a congresos.....	76
13.2	Material suplementario.....	77
13.2.1	Rutas metabólicas.....	77
13.2.2	Frecuencia de secuencias de unión para cada CArG-Box .....	82
13.2.3	Gráficos de clasificación funcional de primer nivel .....	85
13.2.4	Script eliminación de secuencias palíndromas.....	88
13.2.5	Script de aislamiento de promotores. ....	89
13.2.6	Script parseo de resultados BLAST .....	90

### III - Índice de Figuras

Figura 1 Estructura y función de los dominios MADS-BOX tipo II: .....	13
Figura 2 Logo de la secuencia consenso CArG-Box: .....	14
Figura 3 Efectos fenotípicos de los factores de transcripción MADS-Box: .....	17
Figura 4 Mapeo y anotación funcional del genoma de <i>F. vesca</i> : .....	25
Figura 5 Método. ....	28
Figura 6 Gráfica de región promotora: .....	29
Figura 7 Formato de la base de datos de promotores de <i>F. vesca</i> . ....	34
Figura 8 Sitios de unión y promotores con presencia de CArG-Box: .....	45
Figura 9 CArG-box promiscuos entre AP3, FLC, PI y SVP. ....	47
Figura 10 Sitios de unión exclusivos de SEP4, SRF y FLC: .....	47
Figura 11 Sitios de unión más frecuentes: .....	48
Figura 12 Distribución de CArG-Box en promotores: .....	49
Figura 13 Clasificación ontológica de Proceso biológico Nivel 2: .....	50
Figura 14 Clasificación ontológica de Proceso biológico Nivel 3: .....	51
Figura 15 Clasificación ontológica de Función molecular Nivel 3: .....	52
Figura 16 Clasificación ontológica de Función Molecular Nivel 2: .....	52
Figura 17 Clasificación ontológica de Componente celular Nivel2: .....	53
Figura 18 Clasificación ontológica de Componente celular Nivel 3: .....	53
Figura 19 Clasificación ontológica de Componente celular Nivel 8: .....	54
Figura 20 Ruta metabólica KEEG de Biosíntesis de Antocianinas: .....	56
Figura 21 Poster Fruit Biotechnology: .....	76
Figura 22 Ruta KEEG de interconversión entre pentosas y glucoronato: .....	77
Figura 23 Ruta KEEG de Biosíntesis de ácido grasos: .....	78
Figura 24 Ruta KEEG de metabolismo de purinas: .....	79
Figura 25 Ruta KEEG de Degradación de valina, leucina e isoleucina: .....	80
Figura 26 Ruta KEEG de biosíntesis de fenilpropanoides: .....	81
Figura 27 Frecuencia de sitios de unión de SVP: .....	82
Figura 28 Frecuencia de sitios de unión de SRF: .....	82
Figura 29 Frecuencia de sitios de unión de SEP3: .....	82
Figura 30 Frecuencia de sitios de unión de SEP4: .....	83
Figura 31 Frecuencia de sitios de unión de AGAMOUS: .....	83
Figura 32 Frecuencia de sitios de unión de AGL15: .....	83
Figura 33 Frecuencia de sitios de unión de AP3: .....	84
Figura 34 Frecuencia de sitios de unión de PI: .....	84
Figura 35 Frecuencia de sitios de unión de FLC: .....	84
Figura 36 Clasificación ontológica Función molecular: .....	85
Figura 37 Clasificación ontológica Proceso biológico: .....	86
Figura 38 Clasificación ontológica Componente celular: .....	87

#### IV - Índice de tablas

Tabla 1: Descripción de genes del modelo de desarrollo floral ABCDE para <i>Arabidopsis thaliana</i> .	16
Tabla 2 Sitios CArG-Box, Búsqueda del motivo y Logo generado con MEME:	42
Tabla 3 Resultado general de búsqueda sobre base de datos de promotores:	44
Tabla 4 Análisis de promotor Antocianidina:	55
Tabla 5 Resultados validación por referencia biológica:	57