Comparative study of Chalcone synthase promoters across plant families

Estudio comparativo de promotores de la Chalcón Sintasa en diferentes familias de plantas

Buitrago, F.[1], Sierra, R.[2], Grajales, A.[3], Rodríguez-R, L.M.[4], Jiménez, P.[5], Bernal, A.[6], Restrepo, S.[7]

Abstract

In the post – genomic era the understanding of gene regulation has become a challenge and a research priority. In this research, we performed a comparative study of the regulator sequences of the chalcone synthase gene across plant families. Twenty-two sequences of chalcone synthase promoters were compared considering three regulator Cis elements: G-Box, H-Box and TATA Box. Our results show that these Cis elements are conserved among species and even at the family level. However, in some species all of the Cis elements were not found, showing that the expression and regulation of these promoters via the Cis elements can be variable. Additionally, a comparison between promoters from a species with a chalcone synthase multigene family showed that the duplicate genes are variable in the composition of the Cis elements, suggesting that these genes could be expressing in different ways.

Key Words: Promoter, Chalcone synthase, Cis elements, Floral expression

1 Francisco Buitrago, B. Sc. Cra 1 No. 18A - 10 Bloque J205 Laboratorio de Micología y Fitopatología, Universidad de los Andes, Bogotá, Colombia. fra-buit@uniandes.edu.co

2 Roberto Sierra, M.Sc. Cra 1 No. 18A - 10 Bloque J205 Laboratorio de Micología y Fitopatología, Universidad de los Andes, Bogotá, Colombia. ro-sierr@uniandes.edu.co

3 Alejandro Grajales, M.Sc. Cra 1 No. 18A - 10 Bloque J205 Laboratorio de Micología y Fitopatología, Universidad de los Andes, Bogotá, Colombia. alejogr@gmail.com

4 Luis Miguel Rodríguez-R, B. Sc. Cra 1 No. 18A - 10 Bloque J205 Laboratorio de Micología y Fitopatología, Universidad de los Andes, Bogotá, Colombia. luisrodr@uniandes.edu.co

5 Pedro Jiménez, Ph.D. Cra. 11 No. 101-80 Laboratorio de Fitopatología, Universidad Militar Nueva Granada, Bogotá, Colombia. 2005pjm@gmail.com

6 Adriana Bernal, Ph.D. Cra 1 No. 18A - 10 Bloque J205 Laboratorio de Micología y Fitopatología, Universidad de los Andes, Bogotá, Colombia. abernal@uniandes.edu.co

7 Silvia Restrepo, Ph.D. Cra 1 No. 18A - 10 Bloque J205 Laboratorio de Micología y Fitopatología, Universidad de los Andes, Bogotá, Colombia. srestrep@uniandes.edu.co

Resumen

En la era post-genómica, el entendimiento de la regulación génica se ha convertido en un reto y una prioridad de investigación. En este trabajo realizamos un estudio comparativo de las secuencias reguladoras del gen de la chalcón sintetasa de varias familias botánicas. Veintidós secuencias de promotores de *Chalcone Synthase* fueron comparados teniendo en cuenta tres elementos Cis reguladores: Caja-G, Caja-H y Caja-TATA, que podrían estar actuando como una sola unidad cooperativa. Nuestra comparación muestra que estos elementos puede que se conserven en algunas especies e inclusive que se conserven a nivel de familia. Sin embargo, en algunas especies no todos los elementos Cis fueron encontrados, mostrando que no todas las especies se regulan bajo los mismos parámetros. Adicionalmente, una comparación entre promotores de una misma especie con una familia de multigenes Chs, mostró que los genes duplicados son variables en la composición del elemento Cis, sugiriendo que estos genes pueden estarse expresando de maneras diferentes.

Palabras claves: Promotores, Chalcón sintetasa, Elementos Cis, Expresión floral

Recibido: agosto 21 de 2009 Aprobado: noviembre 23 de 2009

Introduction

Flavonoids play important roles in many biological processes such as pigmentation of flowers, fruits and vegetables, plant-pathogen interactions, fertility and protection against UV light (Brouillard & Cheminat, 1988; Gronquist *et al.*, 2001; Bovy *et al.*, 2007). Anthocyanins belong to the flavonoid class of molecules and contribute with the formation of color in flowers (Buchanan *et al.*, 2000). Color in flowers has been an important subject for study in plant biotechnology and changes in flower color have been carried out by regulation or co-suppression of the genes involved in the expression of the anthocyanins (Nakatsuka *et al.*, 2007; Nakatsuka *et al.*, 2008; Tanaka & Ohmiya, 2008).

Chalcone synthase is a key enzyme in the production of anthocyanins and other flavonoids. Anthocyanin biosynthesis starts with the condensation of 4-coumaroil-CoA and manolil-CoA mediated by Chs (Hanumappa *et al.*, 2007). Chalcone synthase genes form a multigene family in many species; nevertheless, some plants as those belonging to the genera *Antirrhinium* and *Arabidopsis* have a unique copy of the gene (Harborne, 1994).

Although several studies have focused on the study of the gene and particularly on its evolution (Durbin *et al.*, 2000), the promoter sequences have been less studied and no comparative studies across plant families has been done. The chalcone synthase gene promoter has a complex series of regulator Cis elements involved in the expression control (Meer *et al.*, 1990). Faktor *et al.* (1997) described two adjacent motifs, the G-Box (CACGTG) and the H-Box (CCTACC) near to a TATA Box that are essential for specific expression in flowers and roots of the chalcone synthase and suggested that the three motifs act as a unit. Additionally, Koch *et al.* (2001) found additional regulator Cis elements in several species of the Brassicaceae family (Koch *et al.*, 2001).

This study started as an initiative to take advantage of the available data on Cis elements on the promoter sequence of the Chs and expand this knowledge aiming for furture biotechnological applications. Comprehending Cis elements can lead to improve the regulation of protection against UV light in plants, the generation of resistant plants and the possibility of modifiying the flower colors. All of these correspond to interesting topics for future studies in our laboratory. In this research we performed a comparative study of the chal-

cone synthase promoters across different plant families. We describe the presence and conservation of the regulator Cis elements mentioned previously (Faktor *et al.*, 1997).

Materials and Methods

Alignments

A set of twenty-two sequences of chalcone synthase promoters were retrieved from GenBank, details of species and accession numbers are given in table 1. Sequences were aligned in the program Geneious v.4.7.5 (Biomatters Ltd., Auckland) using MUSCLE (8 iterations and the rest of parameters by default) (Edgar, 2004). Then, the alignment was constrained using the Cis elements (G-Box, H-Box and TATA Box) as guides. A different set; consisting of five chalcone synthase promoters from *Pisum sativum* (Ch1, Ch2, Ch3, Ch5, Ch7) was also aligned in the Geneious software using MUSCLE.

Table 1. Summary table of the sequences and taxa in this study.

Taxa	GenBank Accesion No.	Length (bp)	Gene[1]
Aethionema grandiflora	AF249000	424	Chs
Arabidopsis griffitihana	AF248989	1184	Chs
Arabidopsis halleri	AF248986	575	Chs
Arabidopsis lyrata	AF248987	605	Chs
Arabidopsis thaliana	AF248988	581	Chs
Arabis alpina	AF248995	632	Chs
Arabis jacquinii	AF248994	616	Chs
Arabis pauciflora	AF248988	485	Chs
Arabis turrita	AF248996	547	Chs
Barbarea vulgaris	AF249991	483	Chs
Cardamine amara	AF248993	476	Chs
Cochlearia excelsa	AF248999	468	Chs
Gentiana triflora	AB005484	1162	Chs
Ginkgo biloba	EF091691	871	Chs
Lepidium campestre	AF248990	513	Chs
Matthiola incana	AF248997	479	Chs
Nicotiana tabacum	FJ655994	529	Chs
Petunia hybrida	EF199747	550	ChsA
Phaseolus vulgaris	AY268022	1453	Chs8
Pinus radiata	AF337656	1055	Chs1
Pisum sativum	AF060238.1	314	Chs1
Rorippa amphibia	AF248992	469	Chs

[1] Chs: chalcone synthase

Phylogenetic Analysis

A genealogy was built with the first alignment, using a bayesian approach with the software BEAST (Drummond & Rambaut, 2007). The program was run with default parameters and 10,000,000 replicates. The first 1000 iterations were discarded as burn-in. The consensus tree was visualized in FigTree (http://tree.bio.ed.ac.uk/software/figtree/) (Fig. 1).

Finding a chalcone synthase promoter in tomato

We retrieved all sequenced BAC clones of tomato (*Solanum lycopersicum*) from the SOL Genomics Network (http://sgn.cornell edu/) database, and performed a tBlastx search using the mRNA of Chs-A from *Petunia hybrida* as a query. Then, we retrieved a fragment of 2Kb upstream from the best hit and searched for any CIS motif.

Results and Discussion

Phylogenetic Analysis

All the sequences corresponding to the Brassicaceae grouped in a monophyletic clade with high branch support (Figure 1, blue clade). All the Cis elements described by Faktor *et al.* (1997) were found in this family. The G-Box (a typical Cis element of the family) was detected at the beginning of these sequences. In this plant family, besides having the three Cis elements, an extra G-Box was found after the first G-Box identified by Faktor *et al.* This Cis element may be acting as an extra site for binding nuclear proteins for a successful expression of Chs (Harter *et al.*, 1994).

All of the sequences from the genus *Arabidopsis* were grouped in a monophyletic clade. *Arabis pauciflora* grouped with *Choclearia excelesa* showing that this promoter is highly variable

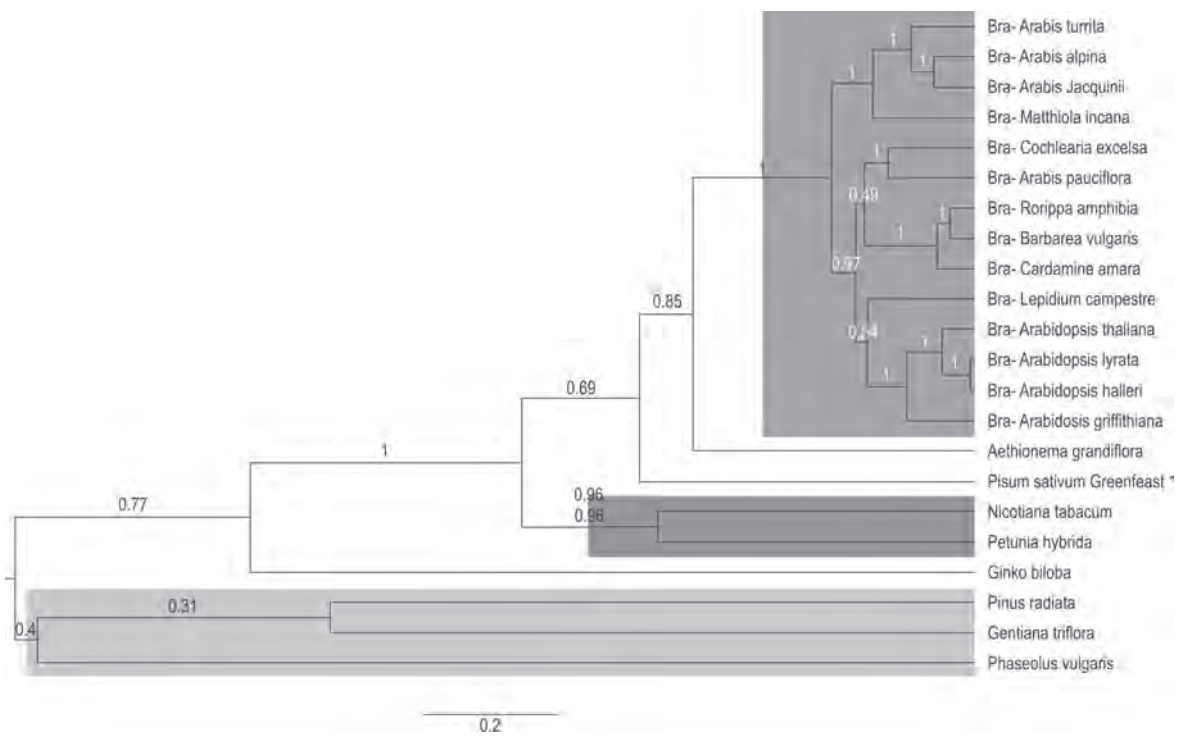

Figure 1. Schematic phylogenetic relationships of taxa under study derived from chalcone synthase promoters, posterior values are given along the branches. In dark gray all Brassicaceae species were grouped, in gray the two species of Solanaceae family were grouped and in light gray three species with very low posterior values are showed.

in the genus *Arabis* an insertion consisting of one adenine at the position 6 of the 3' end of the G-Box was found in *Arabis pauciflora* and six *Arabidopsis* spp. This insertion may be modifying the function of the Cis element since it has been shown that the G-box and the H-box make major contributions to the transcription of some promoters in vivo (Hartmann *et al.*, 1998). Further studies should be done in order to know if the insertion is modifying the function of this Cis element. The species *Aethionema grandiflora* (Brassicaceae) and *Pisum sativum* (Fabaceae) presented the three Cis elements described by Faktor *et al.* (1997), but did not show an extra G-Box, characteristic of the Brassicaceae family. The position of *A. grandiflora* at the base of the Brassicaceae family may suggest that the additional G-Box might be a derived character that characterizes some genus of this family. Additionally, we found one transition in the H-Box from cytosine to thymine in *Pisum sativum*.

Two members of the Solanaceae family used in this study, *Petunia hybrida* and *Nicotiana tabacum*, formed a single group with a posterior probability of 0.96 (figure 1, orange group). This group showed the three Cis elements and an additional G-Box before the G-Box and H-Box, as described by Faktor *et al.* (1997). Thus, we found four Cis elements: G-G-H-TATA Boxes. As mentioned before, the extra G-Box could act like an extra element for specific expression or like an extra site for binding nuclear proteins for a successful expression of chalcone synthase (Harter *et al.*, 1994). This might be due to the fact that the G-Box sometimes carries out its regulatory function combined with other Cis elements (Donald & Cashmore, 1990). Additionally, both species have one transition in the H-Box from a cytosine to a thymine, a feature also found in *Pisum sativum*.

Four species, *Ginko biloba, Gentiana triflora, Pinus radiata,* and *Phaseolus vulgaris,* did not present an H-Box. This absence could have implications in UV light resistance; the LRU (Light Regulator Unit) that contains the G-Box and the H-Box is sufficient for UV/blue light-regulated expression of Chs (Hartmann *et al.*, 1998). These two Cis elements are accounted as the most common in some Chs promoters (Meer *et al.*, 1990; Hartmann *et al.*, 1998; Koch *et al.*, 2001). One possibility is that the regulation of the promoter is mediated by undetected Cis elements. Our results showed that the three Cis elements described by Faktor *et al.* (1997) are not present in all species. This may be indicative that in some plant families these three Cis elements play an important role in the expression of the chalcone synthase gene, while others may present different and currently un-described Cis elements.

Only some gene copies from the Chalcone Synthase gene in *Pisum sativum* presented all the Cis elements identified by Faktor *et al.* (1997) (figure 2). In the promoter Chs1

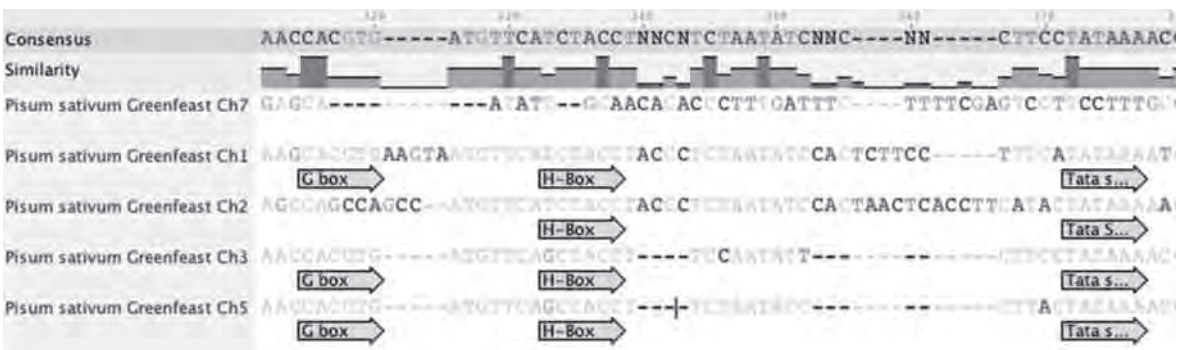

Figure 2. Comparison between chalcone synthase promoters of *Pisum sativum*. The Cis elements described by Faktor *et al.* (1997) are shown in the arrows.

we found the three Cis elements with the previously mentioned transition in the H-Box. In the Chs2 the transition in the H-Box was also found but no G-Box was found. In the Chs3 and Chs4, the three Cis elements were found, but the H-box presented a transversion of one cytosine for one guanine in the first position. These changes or absences of the boxes in these promoters could change the functionality or regulation of the promoter (Hartmann *et al.*, 1998).

Probably these Chs promoters (Chs2, Chs3, Chs5) are only expressed in the plant under stress conditions or are non-functional promoters, flavonioids (mediated in the first step of biochemical reaction by chalcone synthase) can attenuate some adverse effects like heat stress on fertilization or early seed maturation. (Coberly & Rausher, 2003) in the Chs7 promoter no Cis elements were found, in this case, the Chs7 could be a duplicate non-functional copy.

Finding a Chs promoter in tomato

In the sequence of 2Kb of the BAC C12HBa0183M06.1.v1 corresponding to the best tBlastx hit (87% of identity), no Cis elements were found. This sequence may correspond to a non-functional promoter, as in the case of Chs7 of *Pisum sativum.*

Conclusions

There is an apparent evolution of Cis elements that grouped by families as it was seen for the Brassicaceae and Solanaceae family. This must be confirmed with a larger set of sequences from related species and other botanical families. The three Cis elements described by Faktor *et al.* (1997) may be present in promoters of some species or families with different modifications or duplications but were not present in all of the promoters under study. Thus indicating that there may be other Cis elements involved in the expression. Additionally, in *Pisum sativum,* a species with several copies of the Chs gene, the regulation of the expressed promoters of the Chs are carried out by the recognized Cis elements (G-Box, H-Box and Tata-Box) and promoters without this Cis elements may be expressed under stress conditions.

References

Bovy AG, Schijlen EGWM, Hall RD 2007. Metabolic engineering of flavonoids in tomato (*Solanum lycopersicum*): The potential for metabolomics.In.

Brouillard R, Cheminat A. 1988. Flavonoids and plant color. *Prog Clin Biol Res* 280: 93-106.

Buchanan BB, Gruissem W, Jones RL. 2000. *Biochemistry & molecular biology of plants*. Rockville: American Society of Plant Physiologists.

Coberly LC, Rausher MD. 2003. Analysis of a chalcone synthase mutant in *ipomoea purpurea* reveals a novel function for flavonoids: Amelioration of heat stress. *Molecular Ecology* 12(5): 1113-1124.

Donald RG, Cashmore AR. 1990. Mutation of either g box or i box sequences profoundly affects expression from the arabidopsis rbcs-1a promoter. *EMBO J* 9(6): 1717-1726.

Drummond A, Rambaut A. 2007. Beast: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7(1): 214.

Durbin ML, McCaig B, Clegg MT. 2000. Molecular evolution of the chalcone synthase multigene family in the morning glory genome. *Plant Mol Biol* 42(1): 79-92.

Edgar R. 2004. Muscle: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5(1): 113.

Faktor O, Loake G, Dixon RA, Lamb CJ. 1997. The g-box and h-box in a 39 bp region of a french bean chalcone synthase promoter constitute a tissue-specific regulatory element. *The Plant Journal* 11(5): 1105-1113.

Gronquist M, Bezzerides A, Attygalle A, Meinwald J, Eisner M, Eisner T. 2001. Attractive and defensive functions of the ultraviolet pigments of a flower (hypericum calycinum). *Proc Natl Acad Sci U S A* 98(24): 13745-13750.

Hanumappa M, Choi G, Ryu S, Choi G. 2007. Modulation of flower colour by rationally designed dominant-negative chalcone synthase. *Journal of Experimental Botany* 50(10): 2471-2478.

Harborne JB. 1994. *The flavonoid advances in research since 1986*: Chapman & Hall/CRC.

Harter K, Kircher S, Frohnmeyer H, Krenz M, Nagy F, Schafer E. 1994. Light-regulated modification and nuclear translocation of cytosolic g-box binding factors in parsley. *Plant Cell* 6(4): 545-559.

Hartmann U, Valentine WJ, Christie JM, Hays J, Jenkins GI, Weisshaar B. 1998. Identification of uv/blue light-response elements in the arabidopsis thaliana chalcone synthase promoter using a homologous protoplast transient expression system. *Plant Molecular Biology* 36(5): 741-754.

Koch MA, Weisshaar B, Kroymann J, Haubold B, Mitchell-Olds T. 2001. Comparative genomics and regulatory evolution: Conservation and function of the chs and apetala3 promoters. *Mol Biol Evol* 18(10): 1882-1891.

Meer IM, Spelt CE, Mol JNM, Stuitje AR. 1990. Promoter analysis of the chalcone synthase (chsa) gene of petunia hybrida: A 67 bp promoter region directs flower-specific expression. *Plant Molecular Biology* 15(1): 95-109.

Nakatsuka T, Mishiba K-i, Abe Y, Kubota A, Kakizaki Y, Yamamura S, Nishihara M. 2008. Flower color modification of gentian plants by rnai-mediated gene silencing. *Plant Biotechnology* 25(1): 61-68.

Nakatsuka T, Pitaksutheepong C, Yamamura S, Nishihara M. 2007. Induction of differential flower pigmentation patterns by rnai using promoters with distinct tissue-specific activity. *Plant Biotechnology Reports* 1(4): 251-257.

Tanaka Y, Ohmiya A. 2008. Seeing is believing: Engineering anthocyanin and carotenoid biosynthetic pathways. *Current Opinion in Biotechnology* 19(2): 190-197.