Recent Advances in Phytochemistry, volume 35
Regulation of Phytochemicals by Molecular Techniques
Edited by John T. Romeo, James A. Saunders, and Benjamin F. Matthews

Chapter Five

TRANSCRIPTION FACTORS AND METABOLIC ENGINEERING: NOVEL APPLICATIONS FOR ANCIENT TOOLS

Edward L. Braun, Anusha P. Dias, Todd J. Matulnik and Erich Grotewold

Department of Plant Biology & Plant Biotechnology Center
The Ohio State University
Columbus, OH 43210

Introduction..80
Transcriptional Control of Plant Metabolic Pathway Genes.......................80
  Phenylpropanoid Accumulation...81
  Flavonoid Accumulation...84
  Indole Alkaloid Accumulation...85
  Summary – Phytochemical Accumulation.......................................87
Metabolic Engineering by Using Plant Transcription Factors......................88
Identification of Novel Transcriptional Regulators................................91
  Bioinformatics and Predicting Transcription Factor Function....................92
Strategies for Using Plant Transcription Factors..................................96
  The Value of Gain-of-Function Approaches......................................96
  Limitations of Using Transcription Factors.....................................99
Conclusions...102

INTRODUCTION

Metabolic engineering involves the deliberate modification of biochemical networks with the objective of altering the accumulation of specific metabolites. A general prerequisite to engineering a specific metabolic pathway is a good biochemical knowledge of the processes that govern the production and accumulation of the desired product(s). Over the past few decades a wealth of information has been obtained regarding the biosynthesis of a large number of metabolites in plants, animals, fungi, and bacteria. Many of these compounds have economic importance, and substantial effort has gone into increasing their accumulation by genetic selection and biotechnology.

A favorite strategy to increase or alter product formation has been to overexpress enzymes assumed to correspond to rate-limiting steps of particular pathways utilizing transgenic approaches. This strategy has been employed to alter lignin composition in *Arabidopsis thaliana* by overexpression of ferulate-5-hydroxylase (F5H)[1] and increase phenylpropanoid accumulation in tobacco by overexpression of phenylalanine ammonia lyase (PAL).[2] These approaches are limited by flux considerations,[3] the intrinsic plasticity of plant metabolism,[4] the possible participation of enzymes in complexes or metabolic channels,[5] and homeostatic control of metabolic pathways.[6] Moreover, pathways entirely absent from particular cell types can not be manipulated in this fashion because the expression of a single gene is not sufficient for product formation in the absence of the other enzymes of the pathway.

The simultaneous introduction of multiple genes from a specific pathway can overcome these limitations, at least in principle. Nevertheless, this strategy creates additional technical challenges and may result in unpredictable levels of product accumulation or the silencing of the introduced genes.[7] However, in some cases, this strategy has been successful in engineering new metabolic pathways in crop plants. The expression of multiple genes from a single promoter to create polyproteins that are then cleaved by proteases provides an attractive alternative.[9]

TRANSCRIPTIONAL CONTROL OF PLANT METABOLIC PATHWAY GENES

The limitations of the approaches to engineering plant secondary metabolism described above suggest that new strategies are necessary. The simultaneous activation of all genes of a particular pathway by expression of one or a few transcriptional activators provides an emerging technology that may largely overcome the limitations of most currently available methods (also reviewed by C. Martin[10]). However, just as detailed knowledge of the relevant biochemical pathways is necessary for conventional approaches to metabolic engineering, precise

knowledge of the biosynthetic genes subject to transcriptional regulation is necessary for metabolic engineering with transcription factors. Here, we review known roles for transcriptional regulators in controlling the accumulation of selected phytochemicals, focusing upon specific examples with excellent prospects for this novel approach to metabolic engineering.

Phenylpropanoid Accumulation

Plants are specialized to channel carbon on demand, or in response to specific stimuli, from primary metabolism through phenylalanine to the phenylpropanoid pathway, for the biosynthesis of a large number of compounds including lignin and flavonoids. The biosynthesis of phenylpropanoids requires the efficient flow of carbon into the shikimate pathway for the biosynthesis of phenylalanine. This pathway, responsible for the biosynthesis of aromatic amino acids,[11] plays a major role in the production of precursors for aromatic compounds in plants. Chorismate, the end product of the shikimate pathway, is converted to quinones, indoles, and aromatic amino acids, which are the precursors of many secondary metabolites (Fig. 5.1).

The first committed step in the phenylpropanoid pathway is catalyzed by PAL, which converts phenylalanine into cinnamic acid.[12,13] PAL is probably the best-studied biosynthetic enzyme in plant phenolic metabolism, and it is typically encoded by small multigene families in the higher plants (*e.g.*, the three characterized *Arabidopsis PAL* genes[13]). Accumulation of *PAL* mRNA increases in response to many stimuli, such as UV light and pathogens.[12] Genes encoding PAL were among the first isolated and were shown to respond to elicitors from plant pathogens by changes in mRNA accumulation.

Transcriptional activation of genes encoding enzymes involved in phenylpropanoid metabolism (Fig. 5.1), such as PAL, 4-coumarate CoA ligase (4CL), and cinnamyl alcohol dehydrogenase (CAD), probably represents a key step in the regulation of these pathways. However, the transcription factors that specify temporal and spatial expression of these genes remain poorly characterized. The snapdragon (*Antirrhinum majus*) *AmMYB308* and *AmMYB330* genes constitute excellent candidates for transcription factors involved in the regulation of genes specific to the phenylpropanoid pathway.[14,15] In addition, there are a number of other transcription factors that have been identified based upon their ability to bind elements in the promoters of phenylpropanoid biosynthetic enzymes (see below, this section), although their contribution to the regulation of this pathway remains unclear.

The coordinate regulation of the *PAL*, *4CL*, and *CAD* genes in many plant species suggests that specific transcription factors activate the expression of these genes. The promoters of several genes in the phenylpropanoid pathway have been

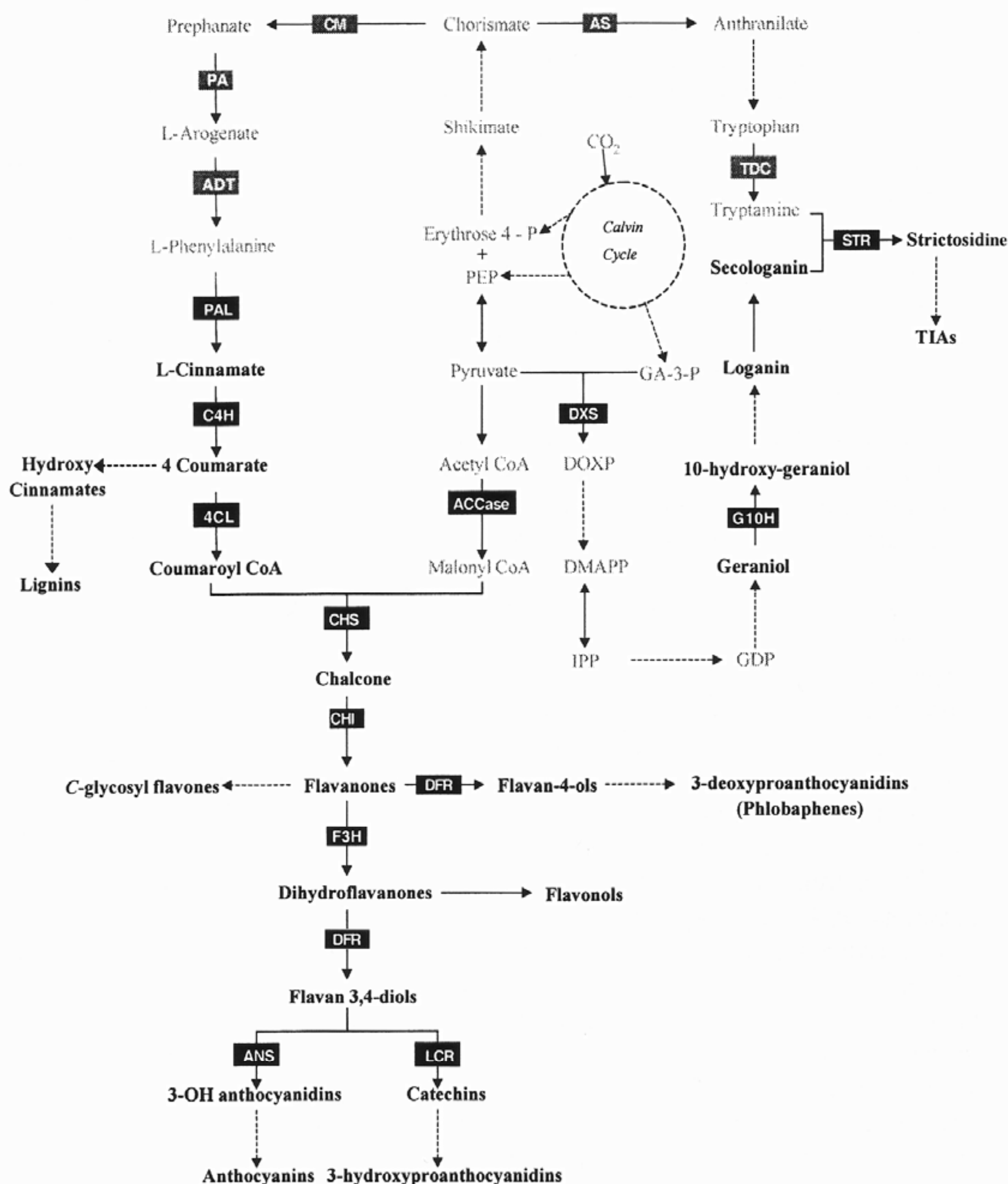

Figure 5.1: Schematic representation of the phenylpropanoid and terpenoid indole alkaloid biosynthetic pathways. Only core enzymes of each pathway are shown and their abbreviations are given in the text. The names of classes of intermediates or final compounds are indicated. Compounds and enzymes from primary metabolism are shown in gray.

dissected, and *cis*-acting regulatory elements important for expression have been identified. For example, members of the *PAL* gene family in parsley (*Petroselinum crispum*) contain the characteristic P-, A-, and L-boxes, also present in other phenylpropanoid biosynthetic genes.[16] Additional other *cis*-regulatory elements have been identified in genes of the pathway, including the G- and H-boxes. G-boxes (with the core consensus sequence CACGTG) mediate the response to various stimuli, including light and abscisic acid, and are recognized by bZIP proteins designated GBFs (G-box factors)[17] and potentially by other factors.[17] H-boxes, which are similar to the P-box identified by Sablowski and colleagues,[19] have the core consensus sequence CCTACC and are found in many phenylpropanoid biosynthetic genes. The G- and H-box regulatory elements in soybean (*Glycine max*) phenylpropanoid biosynthetic genes participate in the early response to pathogen attack.[20]

There are a number of candidate transcription factors that regulate the genes encoding enzymes involved in phenylpropanoid biosynthesis by recognizing H-boxes present in the promoters of these genes. The core consensus sequence of the H-box resembles the consensus binding site of plant R2R3 Myb transcription factors,[21-23] and the product of the snapdragon *AmMyb305* gene recognizes the H-box in the promoter of the tobacco (*Nicotiana tabacum*) *PAL2* gene.[19] This snapdragon Myb protein and a closely related paralog (AmMYB340) provide one of the first examples of transcriptional regulators that control structural genes from the phenylpropanoid pathway as well as flavonoid biosynthetic pathway, since they also activate genes encoding chalcone isomerase (CHI) and flavanone 3-hydroxylase (F3H) in tobacco.[24]

Despite the clear effects of AmMyb305 and AmMyb340 upon *PAL* mRNA accumulation,[24] the regulatory role of H-boxes is complex. Activation of transcription from an AmMYB305 binding site in tobacco protoplasts requires the presence of a G-box,[19] suggesting the need for additional factors *in vivo*. At least two different factors (KAP-1 and KAP-2) have been identified as H-box binding factors in the French bean (*Phaseolus vulgaris*),[25] suggesting that the R2R3 Myb proteins closely related to AmMYB305 and AmMYB340 are unlikely to be the only transcription factors that bind to H-boxes. An H-box binding factor designated BPF-1 (identified as a Box P-binding factor) has been purified from parsley, and transcription of this factor increases in response to elicitor treatment,[26] which also increases *PAL* mRNA accumulation. However, the correlation between *BPF*-1 and *PAL* mRNA accumulation is imperfect, and overexpression studies to establish that increased BPF-1 activity increases *PAL* transcription have not been conducted. Surprisingly, BPF-1 is closely related to the maize (*Zea mays*) IBP proteins[27] and the rice (*Oryza sativa*) RTBP1 protein,[28] which are single-repeat Myb homologs initially identified based upon their ability to interact with the initiator of the maize *Shrunken* gene and with plant telomeres, respectively. Thus, the biological role of these

divergent Myb homologs remains unclear (also see below, section on *Indole Alkaloid Accumulation*).

In addition to Myb homologs, the bZIP protein G/HBF-1 has been identified in soybean,[29] based upon its ability to bind both G- and H-boxes from the *chs15* chalcone synthase promoter. G/HBF-1 is phosphorylated in response to elicitor treatment, and this enhances the binding to the soybean *chs15* promoter. It is unclear whether the apparent complexity of transcriptional regulation by H-boxes (or P-boxes) reflects the existence of multiple factors capable of binding H-boxes *in vivo*, heterogeneity among different plants studied, or differences among H-boxes examined in each study. However, the differences among studies make it difficult to use this information about transcriptional regulators to manipulate phenylpropanoid biosynthesis. From the standpoint of metabolic engineering, overexpression studies, such as those using the snapdragon R2R3 Myb proteins,[19,24] may provide the most useful information.

Flavonoid Accumulation

Flavonoids have multiple functions in plants, including roles as floral pigments for the attraction of pollinators, signalling molecules for beneficial microorganisms in the rhizosphere, and antimicrobial defense compounds.[30,31] In addition, flavonoids are emerging as important nutraceuticals because of their strong antioxidant properties, and several flavonoids have shown anti-tumor activities.[32] The regulation of flavonoid biosynthetic genes (which encode enzymes listed in Fig. 5.1) provides the best-described example of how plants control the expression of genes involved in secondary metabolite biosynthesis.[31] The conspicuous pigmentation provided by flavonoid-derived compounds has allowed the identification of mutants in a large number of structural and regulatory genes in this pathway in various model plants, such as maize (*Zea mays*), petunia (*Petunia hybrida*), and *Arabidopsis*.[33,34]

In maize, anthocyanin accumulation is controlled by two classes of regulatory proteins: a Myb-domain containing class (encoded by the *C1* and *Pl* genes[35,36]) and a basic helix-loop-helix (bHLH)-domain containing class (members of the *R/B* gene families[37]). Anthocyanin production requires interaction between a member of the Myb-domain C1/Pl family and a member of the bHLH-domain R/B family,[31] and the pattern of anthocyanin pigmentation in any particular plant part is controlled by the combinatorial, tissue-specific expression of these regulatory genes. Orthologs of the maize C1 and R regulators—homologs of C1 and R related to the maize proteins by the divergence of species rather than gene duplication[35]—have been identified in other plants, such as petunia and snapdragon.[31] These transcriptional regulators are exchangeable between monocots and higher dicots,[39] suggesting that it may be possible to use transcription factors from a single species for metabolic engineering in a variety of plants.

In addition to 3-hydroxy flavonoids and anthocyanins, maize and its close relatives like sorghum accumulate 3-deoxy flavonoids and derived pigments, which include the phlobaphenes. A single known genetic factor (*P*) controls 3-deoxy flavonoid and phlobaphene biosynthesis in maize.[40-42] *P* encodes a protein with a Myb domain more than 80% identical to the Myb domain of C1,[43] although evolutionary analyses suggest that the divergence between the genes encoding these transcription factors was fairly ancient.[44] In maize floral organs, *P* regulates the accumulation of a subset of mRNAs encoding flavonoid biosynthetic genes, namely *C2* (a chalcone synthase, CHS) and *A1* (dihydroflavonol 4-reductase, DFR). In sharp contrast, C1 is able to regulate the accumulation of mRNAs for *F3H*, *A2* (anthocyanidin synthase, ANS; an enzyme that has also been designated leucoanthocyanidin dioxygenase), *Bz1* (UDP-glucose flavonoid 3-O-glucosyltransferase), and *Bz2* (a glutathione *S*-transferase involved in the vacuolar localization of anthocyanins), as well as *C2* and *A1*,[23] when R or B are present. The regulatory activity of P is independent of the presence of the bHLH proteins R or B,[21,23,45] despite the fact that P and C1 recognize identical elements in the promoter of the *A1* gene.[21]

INDOLE ALKALOID ACCUMULATION

Alkaloids are classified in distinct groups based upon their core nitrogen-containing skeletons. This classification has a logical biochemical basis, since alkaloids with distinct skeletons are often derived from different primary metabolites, such as orinithine/arginine (nicotine and tropane alkaloids), tyrosine (isoquinoline alkaloids), or tryptophan (indole alkaloids).[46] We focus on the indole alkaloids in this review, because substantial progress has been made in elucidating the transcriptional regulation for genes encoding various enzymes involved in their biosynthesis.[47-49]

The regulation of tryptophan biosynthesis has been predicted to be more complex in plants than in other groups of organisms, such as the fungi and bacteria, since a large number of important compounds, including the plant hormone auxin, are derived from tryptophan or its precursors.[50] Further stressing the participation of *Myb* genes in plant metabolic control, an R2R3 Myb transcription factor designated ATR1[47] (altered tryptophan regulation) was identified that positively regulates accumulation of the *Arabidopsis ASA1* mRNA, which encodes the α subunit of anthranilate synthase (AS).[51] Since evolutionary analyses indicate that the *R2R3 Myb* gene family underwent a dramatic expansion early in the evolution of the land plants,[44] the existence of an R2R3 Myb protein that activates transcription of tryptophan biosynthetic enzymes is consistent with the prediction that plants regulate this pathway in unique ways (also see below, section on *Bioinformatics and Predicting Transcription Factor Function*).

Accumulation of the *ASA1* transcript increases substantially in response to wounding or infiltration with *Pseudomonas syringae*,[51] further stressing the unique

patterns of regulation that genes encoding tryptophan biosynthetic enzymes exhibit in plants. However, any potential role of the ATR1 Myb protein in the induction of *ASA1* in response to specific environmental stimuli remains unclear. Indeed, any possible role of this Myb homolog or related paralogs in the regulation of other indole-derived secondary metabolites from *Arabidopsis*, such as the indolic phytoalexin camalexin,[52,53] has not been examined.

The greatest progress in elucidating the role of transcriptional regulation in the control of indole alkaloid accumulation has been made for compounds related to the *Vinca* alkaloids (vincristine and vinblastine). These are terpenoid indole alkaloids (TIAs) with important roles as chemotherapeutics in the treatment of cancer, which has generated substantial interest in their biosynthesis. TIAs have been identified from the Apocynaceae, Loganiaceae, and Rubiaceae, which are members of the Gentianales.[54] Extensive biosynthetic studies for TIAs have been conducted in the Madagascar periwinkle (*Catharanthus roseus*) to establish the identity of enzymes involved in their biosynthesis (Fig. 5.1).

The indole moiety of TIAs is derived from tryptophan after conversion to tryptamine by the enzyme tryptophan decarboxylase (TDC),[55] while the terpene moiety is derived from the iridoid glucoside secologanin.[55] Secologanin is derived from geraniol, synthesized by the chloroplast localized deoxyxylulose phosphate pathway.[56] Strictosidine synthase (STR) catalyzes the stereospecific condensation of tryptamine and secologanin, resulting in the formation of strictosidine, from which all other TIAs are derived.[55] Accumulation of these phytochemicals is induced by exposure to elicitors or jasmonic acid,[57] consistent with a role for TIAs in defense against pathogens and predators. However, the biosynthesis of TIAs appears to be complex, since the enzymes necessary for their biosynthesis are localized in distinct cell types, suggesting that intercellular translocation of intermediates is necessary.[58] Thus, the enzymes involved in TIA biosynthesis are subject to developmental control and respond to environmental stimuli, suggesting complex transcriptional control for the genes encoding them.

Members of the AP2/EREBP family of transcription factors, a diverse group of DNA-binding proteins thus far identified only in the plants[59] (and chlorophyte algae, based upon a *Chlamydomonas reinhardtii* EST encoding an AP2 homolog, accession number AW720749), have a role in the regulation of genes encoding TIA biosynthetic enzymes.[49,57] Two distinct periwinkle AP2-domain transcription factors, ORCA2 and ORCA3 (octadecanoid-derivative responsive *Catharanthus* AP2 domain proteins), have been identified as regulators of TIA biosynthetic genes. Both regulators show increased accumulation in response to jasmonic acid,[49,57] suggesting that they play a role in the induction of TIA accumulation by this plant hormone.

ORCA2 was identified in a yeast one-hybrid screen by using the jasmonic acid- and elicitor-response element from the periwinkle *STR* gene,[57] and ORCA2 is capable of trans-activating a fusion of the *STR* promoter to a GUS reporter gene. ORCA3 was identified by screening periwinkle cell lines containing T-DNA

activation tags for lines with elevated TDC activity.[49] ORCA3 overexpression increased mRNA accumulation for TIA biosynthetic genes, including *TDC*, *STR*, and *D4H* (desacetoxyvindoline 4-hydroxylase), and genes necessary for biosynthesis of TIA precursors (the anthranilate synthase α subunit [AS] and the deoxyxylulose phosphate pathway enzyme D-1-deoxyxylulose 5-phosphate synthase [DXS]). ORCA3 binds to elements in the promoters of *TDC* and *SPR* and is capable of activating transcription from these promoters. However, ORCA3 overexpression does not result in TIA accumulation, because the gene encoding geraniol 10-hydroxylase (G10H), necessary for synthesis of loganin, was not induced in cells overexpressing ORCA3.[49] A third AP2 protein (ORCA1) was identified by the yeast one-hybrid screen by using the periwinkle *STR* promoter,[57] although this protein could not transactivate in plant cells from this promoter element.

The induction of genes encoding enzymes involved in both early (TDC and STR) and late (D4H) steps of TIA biosynthesis (see Fig. 5.1) by a single transcription factor is surprising, since these enzymes appear to be localized to different cell types in periwinkle plants.[58] This suggests that additional regulatory factors exist and that these factors show differential regulation *in planta* and in suspension cultures (or after exposure to elicitors). Indeed, a number of additional candidate transcription factors that may play a role in the regulation of TIA biosynthesis have been identified. For example, the promoter of the periwinkle *STR* gene contains several regions that are recognized by the tobacco GT-1 protein,[60] a member of a small family of plant transcription factors characterized by the trihelix DNA-binding domain.[61] One of the promoter regions recognized by GT-1 is capable of driving transcription of a reporter gene in response to elicitor treatment in transgenic tobacco,[60] suggesting that GT-factors may play a role in the regulation of TIA biosynthetic genes. A periwinkle BPF-1 homolog induced by elicitor in a jasmonic acid-independent manner has been shown to bind a novel element in the *STR* promoter, raising the possibility that this divergent class of Myb homologs play a role in the regulation of TIA biosynthetic genes.[62] Given the variety of biological processes that BPF-1 homologs have been implicated in (see above, section on *Phenylpropanoid Accumulation*), it seems critical to establish the roles that these divergent Myb proteins play in plants.

SUMMARY - PHYTOCHEMICAL ACCUMULATION

The regulation of phytochemical accumulation by transcription factors is clearly complex, with distinct types of transcription factors contributing to the regulation of various enzymes involved in both primary and secondary metabolism. This complexity may pose substantial problems for metabolic engineering. Indeed, differences between suspension cultures and regulation *in planta* (*e.g.*, for TIA

biosynthesis) suggest the existence of differentially expressed and uncharacterized transcriptional regulators. Despite problems, there are good prospects for metabolic engineering of many pathways (see below, next section). Better understanding of the roles played by these regulators will allow the use of transcription factors to change the accumulation of secondary metabolites.

METABOLIC ENGINEERING BY USING PLANT TRANSCRIPTION FACTORS

Clearly, transcriptional regulators provide unique tools to modify plant traits.[10] Plant transcription factors often control several genes of a particular pathway in a coordinate fashion, providing an opportunity to activate entire biochemical pathways by expressing single transgenes. Despite the apparent power of this strategy, few examples of the successful use of transcription factors to manipulate plant metabolism have been described (Table 5.1). One reason for this is our limited knowledge of the transcriptional control of plant metabolic pathways. Nevertheless, the few reported examples suggest that regulatory factors will become a favorite strategy to manipulate plant metabolism.

Since the cloning of the first plant transcription factor more than a decade ago,[35] the regulators of anthocyanin biosynthesis have been extensively studied. One of the most intriguing findings is the exchangeability of these factors among plant species. Ectopic expression of the bHLH factors Lc (a member of the maize R gene family) or DELILA (the snapdragon ortholog of R) in transgenic tobacco or *Arabidopsis* plants resulted in a dramatic increase in anthocyanin production.[63,64] These and other studies[39] suggest that for pathways conserved in different plants, such as the anthocyanin pathway, orthologous regulators are likely to be exchangeable. Indeed, this may allow exceptional flexibility in the selection of transcription factors and plant species used for metabolic engineering.

Recently, the two main branches of maize flavonoid biosynthesis have been activated independently by ectopic expression of R2R3 Myb transcription factors in maize Black Mexican Sweet (BMS) cultured cells.[65] Expression of P resulted in the accumulation of 3-deoxy flavonoids and *C*-glycosyl flavones, similar to the compounds controlled by the *P* gene in maize floral organs.[42] In addition, the expression of P in BMS cells also resulted in an unexpected increase in the accumulation of the phenylpropanoids ferulic and chlorogenic acids,[65] suggesting an additional role of *P* in the control of phenylpropanoid biosynthesis. Consistent with this observation, a *PAL* transcript showed increased accumulation in response to the ectopic expression of P in a subsequent study.[66] Simultaneous expression of C1 and R in BMS cells resulted in a high level of anthocyanin accumulation,[65] and the induction of many mRNAs encoding flavonoid biosynthetic enzymes was readily

Table 5.1: Examples of successful metabolic engineering of phytochemical accumulation using transcription factors.

Regulator	Genes Controlled	Pathway Controlled[a]	Evidence Derived from:[b]
Zm-C1 (R2R3 Myb) & Zm-R (bHLH)	*C2* (CHS), *F3H*, *A1* (DFR), *A2* (ANS), *BZ1*, *BZ2*	Anthocyanins (+)	Loss-of-function mutants[126] and gain-of-function[77] analyses
Zm-P (R2R3 Myb)	*C2* (CHS), *A1* (DFR)	Phlobaphenes (+) *C*-Glycosyl flavones (+)	Loss-of-function mutants[53] and gain-of-function[77] analyses
Am-MYB308 (R2R3 Myb)	*C4H*, *4CL*, *CAD*	Phenylpropanoids (-) Lignin (-)	Gain-of-function analyses[c] [24]
Am-MYB330 (R2R3 Myb)	*4CL*	Phenylpropanoids (-) Lignin (-)	Gain-of-function analyses[c] [24]
Cr-ORCA3 (AP2)	*ASα*, *TDC*, *DXS*, *CPR*, *STR*, *D4H*, (not *G10H*)	Terpenoid indole alkaloids[d] (+)	Gain-of-function analyses[61]

[a] The specific phytochemicals that show altered accumulation and the nature of that change (activation [+] or repression [-]).
[b] Evidence that a particular transcriptional regulator controls a given pathway is provided by the mutant phenotypes (loss-of-function mutants) and increased accumulation of particular compounds when the regulator is overexpressed in plants or plant cells (gain-of-function evidence).
[c] Gain-of-function analyses were carried out in a different species of plant from the plant encoding the regulator.
[d] Terpenoid indole alkaloid biosynthesis was observed only when loganin was added to the cell cultures, probably reflecting the absence of increased *G10H* expression.

detectable.[66] Although P and C1+R expression in BMS cells resulted in the accumulation of 3-deoxy- and 3-hydroxy-flavonoids, respectively, neither *CHI*[67] (chalcone flavanone isomerase), mRNA accumulation, nor CHI enzyme activity was detectable in BMS cells expressing P or C1+R,[65] despite the participation of CHI in

the biosynthesis of both types of flavonoids (Fig. 5.1). These findings indicate that it may not be necessary to have all the genes that play a role in specific pathways coordinately increased to dramatically increase flux by the corresponding regulators.

Since the induction of anthocyanin accumulation requires the simultaneous expression of both C1 and R, a chimeric factor (CRC) with the activities of both proteins was created.[66] Surprisingly, expression of P or CRC in BMS cells resulted in the induction of many more genes than those previously expected to be regulated by P or C1+R.[66] These genes encode glutathione *S*-transferases, MRP transporters, at least one cytochrome P450, and several other proteins predicted to participate in the modification and subcellular localization of the phytochemicals controlled by P and C1+R.[66] Thus, ectopic expression of transcription factors also provides a novel way to investigate currently unknown functions played by regulatory proteins (also see below, section on *Strategies For Using Plant Transcription Factors*).

Novel functions for transcription factors also can be identified by expression in a heterologous system, demonstrated by studies of the snapdragon AmMYB308 and AmMYB340 R2R3 Myb proteins in tobacco. These paralogs exhibit 94% identity within their DNA binding domains,[68] suggesting that they have similar DNA-binding activities. However, they are expressed in different regions of the snapdragon plant, with *AmMYB308* normally expressed throughout the plant and *AmMYB340* expressed only in flowers. When expressed from the *CaMV* 35S promoter in tobacco, the transgenic plants show reduced growth and patches of white, dead cells on mature leaves, in sharp contrast to the normal brown phenotype of senescent tobacco.[15] HPLC analysis of methanol-soluble extracts from these transgenic plants showed a reduced accumulation for a number of phenolics, with esters of caffeic and ferulic acid esters reduced 70-98%. Flavonoids were also affected, accumulating to levels about half of those in untransformed plants. Significant reduction in the accumulation of mRNAs encoding phenylpropanoid biosynthetic enzymes C4H, 4CL, and CAD was also observed. Lignin formation, which requires phenylpropanoid precursors, was also affected and showed an estimated total reduction of more than 50% in mature tissues and almost 20% in younger tissues. This finding is especially interesting for paper production, as lignin removal from wood pulp is currently achieved by environmentally-challenging chemical treatments.

A particularly elegant application of metabolic engineering to the identification of transcription factor function is provided by the identification of the TIA biosynthesis regulator encoded by *ORCA3* in periwinkle.[49] The strategy used corresponds essentially to metabolic engineering accomplished by overexpressing random genes, and can be applied to any metabolic pathway for which a convenient screen can be devised. To examine the regulation of TIA biosynthesis, the ability of TDC to detoxify tryptophan analogs such as 4-methyltryptophan was exploited, allowing the identification of TDC overproducing cell lines that were generated by the transformation of periwinkle cells with a T-DNA activation tag. This allowed the

identification of a single transcription factor (ORCA3) that increased the expression of genes involved in both secondary metabolism and the biosynthetic pathways for precursors (see above, section on *Indole Alkaloid Accumulation*). However, ORCA3 overexpression is not sufficient for accumulation of TIAs in periwinkle cells, since at least one gene encoding an enzyme necessary for biosynthesis of the terpene moiety (G10H) is not induced by expression of this transcription factor (see below, section on *Limitations Of Using Transcription Factors*). A similar activation tag strategy has resulted in the identification of an *Arabidopsis* Myb homolog (*pap1-D*) capable of inducing anthocyanin accumulation throughout the plant when overexpressed (results of J. Borewitz, *et al.*, described by R. A. Dixon and C. L. Steel[69]).

The use of transcription factors to manipulate plant traits need not be limited to altering the accumulation of specific phytochemicals. For example, the expression of transcription factors could be used to alter stress resistance. One such experiment involved the *Arabidopsis* DREB1A AP2-domain transcription factor, which recognizes the DRE (dehydration response element) found in the promoters of several dehydration- and cold-stress inducible genes. When DREB1A was expressed in *Arabidopsis* plants from the promoter of the *rd29A* gene, which contains DRE elements itself and is induced by various stress conditions, plants highly tolerant to drought, salt, and freezing were obtained.[70] Expression of DREB1A from the *rd29A* promoter produced minimal negative effects on the plant in non-stress conditions, while expression of the same gene from the constitutive *CaMV* 35S promoter produced many undesirable effects. However, it is unclear how an activation loop in the *rd29*A::DREB1A plants was prevented, since *rd29A* is regulated by DREB1A.[70] Regardless of the specific features of the promoter used, this study indicates that other important plant traits can be manipulated by the ectopic expression of transcription factors.

IDENTIFICATION OF NOVEL TRANSCRIPTIONAL REGULATORS

Although the utilization of transcription factors as tools to engineer plant metabolism provides an attractive alternative to more classical approaches, there are formidable obstacles to the application of this approach outside of those examples that we list in Table 5.1, since there is almost no information about the role of transcriptional regulation in other plant metabolic pathways. Biochemical methods aimed at characterizing factors that interact with specific *cis*-acting DNA-regulatory elements in the promoters of genes encoding enzymes of particular pathways have been extensively used. However, these approaches are restricted by the low level of expression for many regulatory proteins and the limitations imposed by the specific cell types where particular phytochemicals accumulate.

The screening for mutants in regulatory genes has been helpful when the pathways controlled yielded visible intermediates or final products, as is the case for flavonoids[33] or the tryptophan pathway.[47] However, secondary metabolites are defined as compounds that are not essential for viability[71] so mutations in most secondary metabolic pathways will have relatively modest phenotypic effects. This, together with the functional redundancy characteristic of higher plants, makes the identification of mutants in regulators of secondary metabolism a low-payoff approach. This assertion is dramatically exemplified by the enormous effort invested in identifying DNA insertions (T-DNA or transposons) in 36 *Arabidopsis R2R3 Myb* genes that yielded few detectable phenotypes in the large variety of conditions tested.[72] However, the large number of distinct transcriptional regulators present in plants (see below, next section) suggests that the identification of specific candidate transcription factors for use in metabolic engineering represents a necessary initial step for this strategy. The growing interest in plant genomics is likely to provide substantial information that can be used to identify regulators that represent candidates for further study.

Bioinformatics and Predicting Transcription Factor Function

Transcriptional regulators may represent as much as 10% of the genes present in plant genomes (based upon analyses of the complete sequence of chromosome 4 from *Arabidopsis*[73]). Transcription factors have been classified into families based on the presence of conserved DNA-binding domains, with many of the families of transcriptional regulators that have been identified in the animals having plant homologs.[74] However, the cellular processes associated with particular classes of regulatory proteins in plants and animals often differ. For example, plant MADS-box proteins control a number of homeotic functions,[75] largely performed by homeodomain proteins in animals, while plant homeodomain proteins participate in functions associated with meristem identity and boundary establishment.[76] In addition to these broadly distributed groups of transcription factors, groups apparently specific to plants, such as the AP2 proteins[59] and trihelix factors,[61] also have been identified.

Plants also express members of transcription factor families with broad phylogenetic distributions that exhibit substantial qualitative differences from their homologs in animals and fungi. For example, animals express a handful of Myb-domain transcription factors, with the best characterized proteins—those related to the vertebrate c-Myb proto-oncoprotein—having three Myb-homologous direct repeats.[77] Although similar factors are present in plants,[78-80] the majority of the plant *Myb* gene family encodes proteins derived from the three repeat proteins by loss of the first repeat.[78] Thus, the majority of plant Myb genes encode proteins characterized by the presence of only two Myb repeats directly related to the second and third repeats in the vertebrate c-Myb protein (designated R2 and R3).[44,79,81] This

gene family underwent a striking expansion 250-550 million years ago,[44] early in evolution of land plants, resulting in the presence of more than 150 paralogous *Arabidopsis* genes encoding R2R3 Myb proteins (Wenning, Braun and Grotewold, unpublished observation).

It seems likely that the expansion of gene families encoding transcription factors, such as the *R2R3 Myb* gene family, was directly associated with the evolution of novel functions that these factors regulate. Additional transcription factor families, such as the AP2 and the MADS-box families, underwent similar expansions. This suggests that one may be able to make functional inferences from their patterns of evolution. Analyses of genes encoding MADS-box proteins from the ferns *Ceratopteris pteroides* and *Ophioglossum pedunculosum* revealed that many of the MADS-box proteins involved in floral homeotic functions arose by gene duplication after the divergence of ferns and seed plants.[82] Likewise, the majority of characterized R2R3 Myb proteins control plant-specific functions thought to have arisen during the colonization of the land or the development of the vascular system.[44,79,81] One might infer that transcription factor genes that diverged prior to these periods might encode proteins with distinctive functions.

For example, phylogenetic analyses indicate the existence of deep-branching R2R3 Myb homologs related to the *Arabidopsis* AtMyb1 protein[83] and are characterized by the absence of a leucine insertion in the R2 Myb repeat[78] that is present in typical R2R3 Myb proteins from plants. This atypical class arose before the origin of land plants, since a member of this group is present in the chlorophyte alga *C. reinhardtii* (an EST with accession number AV396320). Unlike typical R2R3 Myb proteins, these atypical Myb proteins represent poor candidates for regulators of secondary metabolite accumulation. These differences emphasize the importance of carefully differentiating among distinct groups of homologous transcription factors, such as the typical R2R3 Myb proteins, the atypical (AtMyb1-like) R2R3 Myb proteins, the three repeat Myb homologs,[78-80] and the distantly related BPF-1/IBP-like Myb homologs.[26-28,62]

Evolutionary analyses can provide information suggesting that a regulator is unlikely to be involved in certain processes by providing evidence that a specific subfamily arose prior to the evolution of the process. Indeed, focusing upon gene duplications correlated with the evolution of specific phenotypes (such as the accumulation of a distinctive phytochemical) as candidates for regulators of that phenotype seems reasonable. Recent progress in angiosperm phylogeny[84] suggests that it may soon be possible to infer the divergence times for many different groups of plants fairly accurately, allowing this strategy to be applied to plants that accumulate interesting metabolites. However, the identification of recent duplications in genes encoding transcription factors must consider the patterns of duplication for other genes in the organism under consideration. Thus, the identification of more than two closely related genes would be necessary for plants that have undergone a genome duplication (such as maize[85]). Preliminary analyses of

this type have been conducted,[86] revealing six groups of maize *R2R3 Myb* genes that underwent more than one gene duplication in the past 40 million years.

One of these groups of closely related maize *R2R3 Myb* genes contains at least nine closely-related paralogs. This group includes the maize *P* gene and has been designated the proline-to-alanine clade,[44] because of a specific amino acid substitution at a slowly evolving position that characterizes the group. Since the regulatory activity of *P* (activation of a subset of flavonoid biosynthetic genes resulting in phlobaphene pigment accumulation[23]) has been demonstrated only in maize and sorghum thus far,[87] the recent expansion of this group is consistent with an association between gene duplications and the evolution of novel regulatory functions. Some members of the proline-to-alanine clade exhibit a surprising degree of sequence divergence, suggesting a higher rate of evolution.[44] Establishing the functions of these divergent proteins will be informative, since many paralogs with different functions exhibit evolutionary rate differences.[88] Changes in the ratio of synonymous to nonsynonymous differences may prove useful for this, since functional changes should alter the accumulation of nonsynonymous rather than synonymous changes. However, correlations between the rate of evolution at synonymous and nonsynonymous sites[89] and the saturation of synonymous divergence[44] may complicate this type of analysis.

To determine whether a similar correlation was apparent for AP2-domain regulators of TIA biosynthesis, we used the ORCA proteins as queries for database searches. ORCA2 and ORCA3 are each other's top hit in BLAST searches, and both proteins recognize the same *Arabidopsis* protein (an uncharacterized ORF designated At2g44840) as their top hit in another organism (Braun and Grotewold, unpublished observations), suggesting a gene duplication after the divergence of the Gentianales and Brassicales. Searches indicated that ORCA1 is closely related to the *Arabidopsis* DREB2A protein,[57] consistent with the functional differences between ORCA1 and ORCA2/3. However, this relationship may be more complex, since ORCA2, ORCA3, and At2g44840 exhibit similar synonymous and nonsynonymous distances from each other (Braun and Grotewold, unpublished observations), consistent with a gene duplication prior to the divergence of the Gentianales and Brassicales. Resolving the timing of these gene duplications and establishing whether these genes exhibit differences in evolutionary rate will prove interesting.

Although the identification of genes encoding transcription factors with novel functions may be facilitated by identifying recently expanded gene clades or genes exhibiting evolutionary rate differences, functional inferences will require additional data. If transcription factors that exhibit a close phylogenetic relationship also exhibit similar regulatory functions, the use of evolutionary information will be powerful. Functional analyses of orthologous genes encoding MADS-box and R2R3 Myb proteins have indicated that functions are often conserved in distantly related plant taxa.[31, 90] Closely related paralogous transcription factors may also have similar regulatory functions. For example, the conserved domains of both MADS-box

proteins[90] and R2R3 Myb proteins (Wolfe, Braun and Grotewold, unpublished observations) with related functions are often closely related phylogeneticaly.

As predicted by a correlation between phylogeny and function, the maize C1 and petunia AN2 regulators of anthocyanin biosynthesis cluster together in a larger group that also includes P, which may constitute a "phenylpropanoid group." Consistently, other studies positioned the snapdragon AmMYB330 and AmMYB308 in a clade sister to that containing C1.[91] Surprisingly, the snapdragon AmMYB305 and AmMYB340 proteins appear to be distantly related to this phenylpropanoid group,[91] suggesting some independent recruitment of genes to phenylpropanoid regulation. A major advantage of rigorous phylogenetic analyses relative to simple database searches (*e.g.*, BLAST) is the existence of well-accepted methods to evaluate support for specific clades, such as bootstrap or jackknife resampling[92,93] or the Bremer support index.[94] However, these rigorous analyses reveal that many phylogenetic associations among paralogous transcription factor groups (including the phenylpropanoid group) are poorly supported,[44] and researchers should be aware of artifacts driven by factors such as differences in evolutionary rates or amino acid composition.[95] The correlation between evolutionary relationships and regulatory functions of transcription factor genes requires further evaluation.

Genomic and cDNA sequence data may provide rich sources of information regarding the functions of specific transcriptional regulators in addition to the phylogenetic relationships that can be inferred. For example, differences in the composition of synonymous positions of plant genes (such as third codon position GC content) and intron size may have important functional correlates.[96] The identification of transcription factors with distinctive domain organization or evidence for recombination between relatively distantly related transcription factor genes might also represent excellent methods to identify candidate transcriptional regulators. The continued accumulation of sequence data and functional information from plant genomics should improve the selection of criteria for the identification of candidate transcription factors and suggest additional criteria that can be employed to this end.

Overall, we are optimistic regarding the use of evolutionary genomic analyses to identify candidates for transcription factors that regulate the accumulation of specific phytochemicals. For example, preliminary analyses suggest that another member of the proline-to-alanine clade has regulatory activities that overlap with those of P (Dias and Grotewold, unpublished observations), consistent with our assertion that the association between phylogeny and function is sufficiently strong to be useful for functional prediction. However, the use of sequence analyses should be viewed as a means to identify candidates for specific regulatory functions, to be followed with experiments (such as gain-of-function experiments) to evaluate the

regulatory function of the candidate transcription factors in a direct fashion (see below, next section).

STRATEGIES FOR USING PLANT TRANSCRIPTION FACTORS

If a transcription factor controlling a particular metabolic pathway is known, the regulator can be expressed in plants or cultured plant cells from constitutive or inducible promoters, and the accumulation of the metabolite can be evaluated (Fig. 5.2). Some of the basic steps to follow after the identification of candidate transcription factors controlling particular metabolic pathways, to the characterization of gene products or phytochemicals that accumulate in cultured plant cells expressing the regulator are detailed below. This or similar strategies have been successfully applied in the examples we list in Table 5.1.

The Value of Gain-of-Function Approaches

The gain-of-function approach can be used to establish functions of particular regulators. The gene for the transcription factor is expressed in the plant or cultured plant cells under a constitutive or inducible promoter. After selection and propagation of the plant or plant cells expressing the regulator, several strategies are possible to determine which particular pathway the factor controls in the absence of an obvious phenotype. Differentially accumulated chemicals can be quantified with HPLC or GC-MS. This type of approach allowed the description of novel regulatory roles played by the maize *P* gene.[65] Alternatively, other methods to evaluate differential gene expression of possible target genes are available (Fig. 5.2), such as the Curagen GeneCalling technology. GeneCalling was efficiently applied to demonstrate that the ectopic expression of P or C1+R (the CRC chimera) induces or represses a large number of genes in maize BMS cells.[66] While this technology is expensive, the rapid development of microarrays for many model plants[97] provides a good alternative. Indeed, smaller arrays focusing on known metabolic pathways may suffice in many cases.

Comparison of mRNAs from cells expressing the transcriptional activator to control cells will allow the identification of differentially expressed genes, allowing the metabolic pathway(s) that are controlled to be established. This would permit subsequent biochemical confirmation by looking for the accumulation of the pathway products. These data, together with information on the temporal and tissue-specific expression of the regulator, can facilitate the prediction of loss-of-function phenotypes for the regulatory gene, closing the circle in this analysis (Fig. 5.2). Such data will provide substantial information regarding the criteria for the computational identification of candidate transcription factor genes.

This approach does not establish whether differentially expressed genes are directly or indirectly activated or repressed by the corresponding regulators. Fusing transcriptional regulators to the glucocorticoid receptor (GR) and comparing the activation of the putative regulated genes in the presence and absence of dexamethasone (DEX) and a protein synthesis inhibitor, such as cycloheximide, can identify genes that represent direct targets of specific transcriptional regulators. Similar studies have already been used to identify direct targets of the floral homeotic gene AP3.[98] Alternatively, one could examine the binding of the transcription factor to relevant promoters *in vitro* and examine the activation of reporter gene fusions *in planta*, similar to the studies performed with many of the proteins listed in Table 5.1. However, experiments of that type are relatively difficult to apply on a large scale, unlike the use of GR fusions where the mRNAs that accumulate in the presence of both DEX and cycloheximide can be assessed using microarrays or the Curagen GeneCalling technology.

While today is the time of plant genomics, tomorrow's goal will be relating the expression of particular transcription factors to the accumulation of specific proteins or enzymatic activities. Altering enzyme activities represents the ultimate goal of metabolic engineering with transcription factors. Detailed comparisons of mRNA accumulation and protein synthesis in the budding yeast *Saccharomyces cerevisiae* have shown profound differences in the translatability of specific mRNAs,[99] suggesting that differences exist in other groups of organisms as well. In fact, the point mutation that originally defined the *Arabidopsis ATR1* gene (allele *atr1D*) generated a stop codon in an upstream open reading frame, suggesting that ATR1 is subject to translational control.[47] Although few proteomic studies have been conducted in plants,[100] the comparison of isogenic cell lines distinguished only by ectopic expression of transcriptional regulators provided an opportunity to conduct a pilot proteomic study. The patterns of ^{35}S-methionine-labeled proteins obtained from BMS cells expressing P or C1+R were compared by two-dimensional polyacrylamide electrophoresis, and proteins induced by these regulators were revealed (Fig. 5.3). The number of proteins that P or C1+R induce in these analyses[101] is closer to the number of genes expected to be regulated by these factors[65] and significantly lower than the number estimated by using the Curagen GeneCalling technology.[66] These apparent differences may reflect translational regulation or differences in the sensitivity of the methods used.

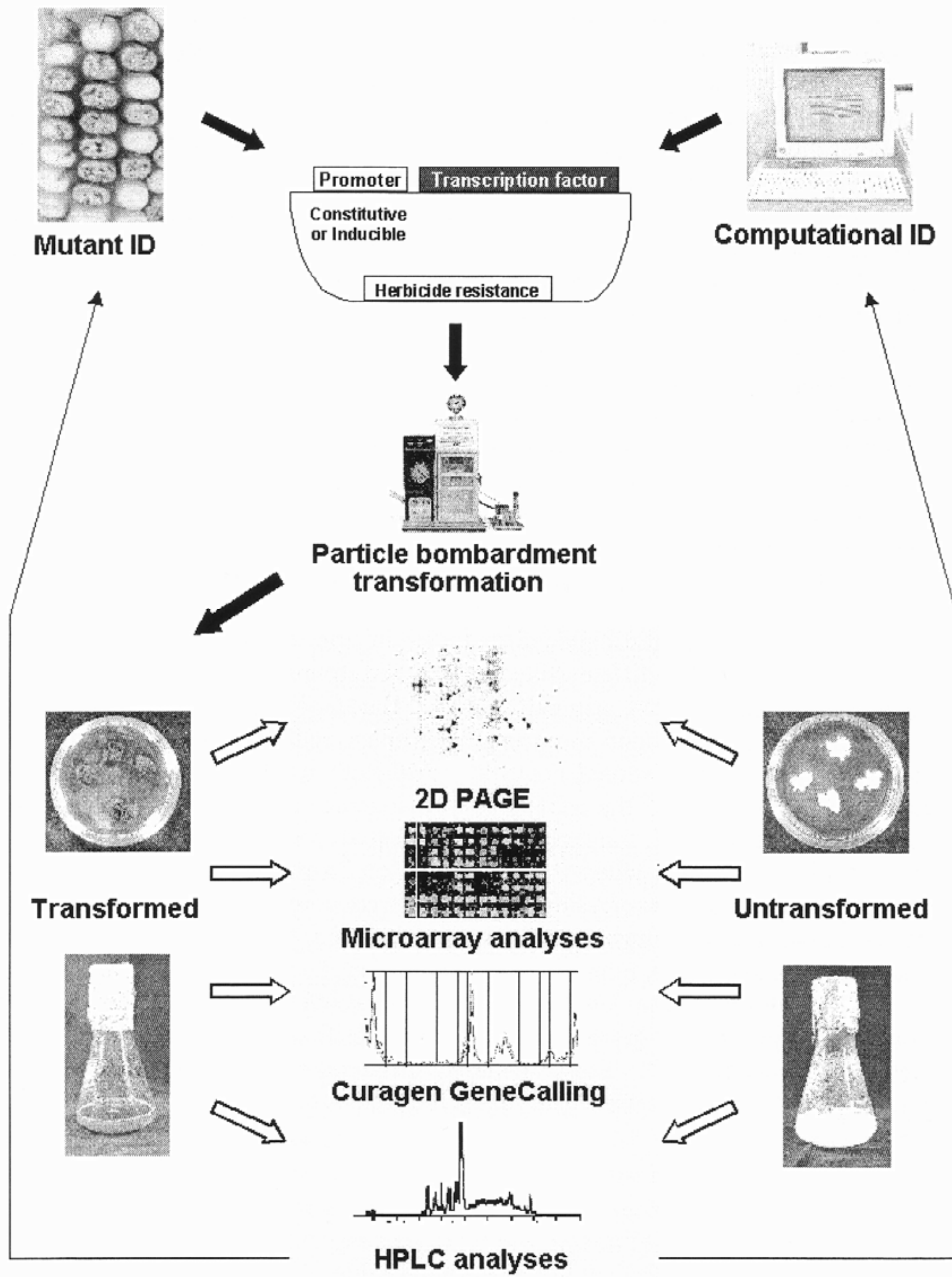
Mutant ID
Promoter
Transcription factor
Constitutive
or Inducible
Herbicide resistance
Computational ID
Particle bombardment
transformation
2D PAGE
Transformed
Untransformed
Microarray analyses
Curagen GeneCalling
HPLC analyses

Figure 5.2: General scheme suggested for the engineering of plant metabolic pathways using transcription factors. Transcription factors either characterized by mutant analysis (mutant ID) or deduced from sequence information (computational ID) are cloned under constitutive or inducible plant promoters. The maize ear shows the phenotype of the *Ac/Ds* transposable element insertions in the *R* (*r-m3* allele, spotted aleurone), or *P* (*P-vv* allele, variegated pericarp) genes. The transcription factors are introduced into plant cells (BMS cells are shown in this case) using the general method of particle bombardment. Transgenic plant cell lines (calli expressing C1+R from the *CaMV* 35S promoters are shown) are selected for herbicide resistance and HPLC or other methodologies are used to investigate the accumulation of different phytochemicals. Alternatively, or in addition, specific genes or proteins induced by the regulators can be identified using a number of available technologies. This information can be used for transcription factors of unknown function to predict function and help identify loss-of-function mutations or to improve the criteria used for computational identification (represented by the thin feedback arrows).

Limitations of Using Transcription Factors

The main limitation of the use of transcription factors for plant metabolic engineering is our knowledge of regulators that control plant metabolic pathways. We propose here several new approaches to determine the function of plant regulators that should complement other more classical methods. These studies will reveal the specificity of transcription factors under various conditions. Maintenance of this specificity will determine the future usefulness of transcription factors for metabolic engineering.

A second limitation is that a single transcription factor might not be sufficient to induce all genes of a given metabolic pathway. For example, the maize regulatory factor C1 is incapable of activating transcription in the absence of the co-activator R.[31] This is overcome either by expressing both regulators together[65] or by generating the CRC chimera.[66] The dogma has been that all genes in a particular pathway are coordinately activated. This, however, may turn out not to be the general case. The maize P or C1+R regulators activate all known flavonoid biosynthetic genes, with the exception of *CHI*.[65,66] Preliminary experiments indicate that maize *CHI1* is regulated by mechanisms different from the other genes in the pathway (Grotewold, unpublished observations). In higher dicots (often called eudicots[84]), C1 orthologs only activate "late genes" in the pathway, and not the early

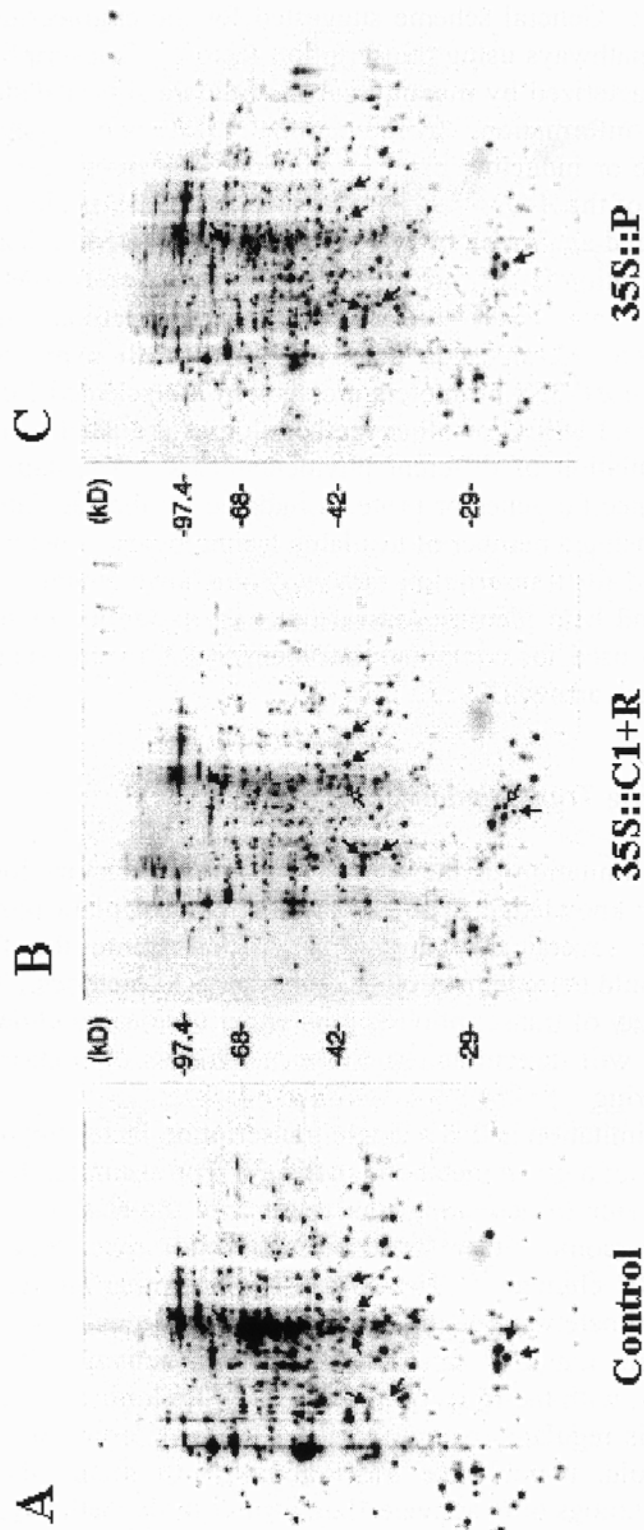
A
B
C
(kD)
-97.4-
-68-
-42-
-29-
(kD)
-97.4-
-68-
-42-
-29-
Control
35S::C1+R
35S::P

Figure 5.3: Proteomic analysis of maize BMS callus cells expressing P and C1+R. The results of 2-D PAGE analyses for maize BMS callus cells (**A**); cells transformed with 35S::C1+R (**B**); or cells transformed with 35S::P (**C**). Arrows in (**B**) and (**C**) indicate protein spots absent in the control (**A**) line. Arrows indicate the positions of the relevant spots in the untransformed lines. Empty arrows in (**B**) correspond to protein spots seen only in the 35S::C1+R lines.

ones,[31] such as *chs* or *chi* (Fig. 5.1). Variation in the specific genes activated in different eudicots suggests that functionally-relevant changes in the promoters of flavonoid biosynthetic genes occur on short evolutionary time scales. The evolutionary flexibility of promoters for other genes involved in secondary metabolism is unclear, but we expect substantial variation to exist. Despite this variation, these regulators are capable of inducing anthocyanin accumulation in a number of eudicots examined (see above, section on *Flavonoid Accumulation*).

Just as the expression of C1+R resulted in the activation of a subset of flavonoid biosynthetic genes in maize BMS cells, overexpression of ORCA3 in periwinkle cells induced only a subset of TIA biosynthetic enzymes[49] (see above, section on *Indole Alkaloid Accumulation*). Cells overexpressing ORCA3 failed to accumulate TIAs unless externally fed with loganin,[49] demonstrating that failure to induce all of the genes participating in a process can be problematic. However, engineered cells expressing part of a biosynthetic pathway may allow the conversion of relatively inexpensive precursors to a more valuable end product.

The use of transcription factors to manipulate plant metabolism typically depends upon the use of constitutive or inducible promoters that drive high levels of expression of the regulator. It is conceivable that the over-expression of these transcription factors could result in non-specific effects associated with an increased binding of the regulator to DNA elements in the promoters of other genes, "squelching" other cellular factors or components of the transcriptional machinery.[102] However, studies that use maize cells stably or transiently expressing P or C1+R argue that these factors retain their biological specificity, even when over-expressed.[31,65,66] Thus, over-expression of transcription factors does not necessarily lead to non-specific effects. Nevertheless, ectopic expression of *P* in *Arabidopsis* plants resulted in phenotypes (greatly reduced plant size and pigment accumulation) that we have been unable to explain.[103] However, neither the function of the *Arabidopsis P* ortholog (encoded by *AtMyb12* and its close relatives, based upon our phylogenetic reconstructions[44,78]) nor the specific ways in which the regulatory function of the maize *P* gene differs from that of its orthologs in other species have been established. Thus, it is premature to conclude that the phenotype conferred by

ectopic expression of P in *Arabidopsis* reflects non-specific transcriptional activation. Expression of C1 in *Arabidopsis* has no obvious phenotypic effect unless expressed together with the co-activator R, which results in anthocyanin accumulation,[63] suggesting that the biological specificity of C1 is maintained.

CONCLUSIONS

The plant kingdom provides a vast source of compounds with important biological activities, and ambitious projects to exploit this diversity have been undertaken recently. The time when plants and cultured plant cells will be routinely used as factories to produce compounds with importance to medicine and agriculture has yet to come, but important progress in the area of plant metabolic engineering has recently been made. Transcription factors are emerging as important tools for these processes, as they allow the activation of entire pathways with just one or a small number of transgenes. Clearly, one of the main challenges in the future will be establishing the regulation of important plant metabolic pathways.

ACKNOWLEDGEMENTS

We are grateful to Vince de Luca for critical reading of this manuscript and to Johan Memelink and his colleagues for sharing unpublished information. We thank our colleagues at the Ohio State Plant Biotechnology Center and Dan Crawford for many helpful interactions and we thank the editors for comments that improved this manuscript. We are extremely grateful to Ben Bowen for helpful interactions, suggestions, and inspiration while conducting much of the work described in this review. Our work is supported in part by grants from the United States Department of Agriculture and the National Science Foundation.

REFERENCES

1. MEIER, K., SHIRLEY, A.M., CUSUMANO, J.C., BELL-LELONG, D.A., CHAPPLE, C., Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in *Arabidopsis*, *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 6619-6623.
2. HOWLES, P.A., SEWALT, V.J.H., PAIVA, N.L., ELKIND, Y., BATE, N.J., LAMB, C., DIXON, R. A., Overexpression of L-phenylalanine ammonia-lyase in transgenic tobacco plants reveals control points for flux into phenylpropanoid biosynthesis, *Plant Physiol.*, 1996, **112**, 1617-1624.
3. NIEDERBERGER, P., PRASAD, R., MIOZZARI, G., KACSER, H., A strategy for increasing an *in vivo* flux by genetic manipulations, *Biochem. J.*, 1992, **287**, 473-479.

4. AP REES, T., Prospects of manipulating plant metabolism, *Trends Biotech.*, 1995, **13**, 375-378.
5. WINKEL-SHIRLEY, B., Evidence of enzyme complexes in the phenylpropanoid and flavoniod pathways, *Physiol. Plant.*, 1999, **107**, 142-149.
6. FRIDLYAND, L.E., BACKHAUSEN, J.E., SCHEIBE, R., Homeostatic regulation upon changes of enzyme activities in the Calvin cycle as an example for general mechanisms of flux control, *Photosyn. Res.*, 1999, **61**, 227-239.
7. LEECH, M.J., MAY, K., HALLARD, D., VERPOORTE, R., DE LUCA, V., CHRISTOU, P., Expression of two consecutive genes of a secondary metabolic pathway in transgenic tobacco: Molecular diversity influences levels of expression and product accumulation, *Plant Mol. Biol.*, 1998, **38**, 765-774.
8. YE, X., AL-BABILI, S., KLOTI, A., ZHANG, J., LUCCA, P., BEYER, P., POTRYKUS, I., Engineering the provitamin A (β-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm, *Science*, 2000, **287**, 303-305.
9. BECK VON BODMAN, S., DOMIER, L.L., FARRAND, S.K., Expression of multiple eukaryotic genes from a single promoter in *Nicotiana*, *Biotechnology*, 1995, **13**, 587-591.
10. MARTIN, C., Transcription factors and the manipulation of plant traits, *Curr. Opin. Biotech.*, 1996, **7**, 130-138.
11. WEAVER, L.M., HERRMANN, K. M., Dynamics of the shikimate pathway in plants, *Trends Plant Sci.*, 1997, **2**, 346-351.
12. KUHN, D.N., CHAPPELL, J., BOUDET, A., HAHLBROCK, K., Induction of phenylalanine ammonia-lyase and 4-coumarate:CoA ligase mRNAs in cultured plant cells by UV or fungal elicitor, *Proc. Natl. Acad. Sci. USA*, 1984, **81**, 1102-1106.
13. WANNER, L.A., LI, G., WARE, D., SOMSSICH, I.E., DAVIS, K.R., The phenylalanine ammonia-lyase gene family in *Arabidopsis thaliana*, *Plant Mol. Biol.*, 1995, **27**, 327-338.
14. TAMAGNONE, L., MERIDA, A., PARR, A., MACKAY, S., CULIANEZ-MACIA, F.A., The AmMYB308 and AmMYB330 transcription factors from *Antirrhinum* regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco, *Plant Cell*, 1998, **10**, 135-154.
15. TAMAGNONE, L., MERIDA, A., STACEY, N., PLASKITT, K., PARR, A., CHANG, C.-F., LYNN, D., MAXWELL DOW, J., ROBERTS, K., MARTIN, C., Inhibition of phenolic acid metabolism results in precocious cell death and altered cell morphology in leaves of transgenic tobacco plants, *Plant Cell*, 1998, **10**, 1801-1816.
16. LOGEMANN, E., PARNISKE, M., HAHLBROCK, K., Modes of expression and common structural features of the complete phenylalanine ammonia-lyase gene family in parsley, *Proc. Natl. Acad. Sci. USA*, 1995, **92**, 5905-5909.
17. MENKENS, A.E., SCHINDLER, U., CASHMORE, A.R., The G-box: A ubiquitous regulatory element in plants bound by the GBF family of bZIP proteins, *TIBS*, 1995, **20**, 506-516.

18. DE PATER, S., PHAM, K., MEMELINK, J., KIJNE, J., RAP-1 is an *Arabidopsis* MYC-like R protein homologue, that binds to G-box sequence motifs, *Plant Mol. Biol.*, 1997, **34**, 169-174.
19. SABLOWSKI, R.W.M., MOYANO, E., CULIANEZ-MACIA, F.A., SCHUCH, W., MARTIN, C., BEVAN, M., A flower-specific Myb protein activates transcription of phenylpropanoid biosynthetic genes, *EMBO J.*, 1994, **13**, 128-137.
20. WEISSHAAR, B., JENKINS, G. I., Phenylpropanoid biosynthesis and its regulation, *Curr. Opin. Plant Biol.*, 1998, **1**, 251-257.
21. SAINZ, M.B., GROTEWOLD, E., CHANDLER, V.L., Evidence for direct activation of an anthocyanin promoter by the maize C1 protein and comparison of DNA binding by related Myb domain proteins, *Plant Cell*, 1997, **9**, 611-625.
22. WILLIAMS, C.E., GROTEWOLD, E., Differences between plant and animal Myb domains are fundamental for DNA-binding and chimeric Myb domains have novel DNA-binding specificities, *J. Biol. Chem.*, 1997, **272**, 563-571.
23. GROTEWOLD, E., DRUMMOND, B., BOWEN, B., PETERSON, T., The Myb-homologous *P* gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset, *Cell*, 1994, **76**, 543-553.
24. MOYANO, E., MARTINEZ-GARCIA, F., MARTIN, C., Apparent redundancy in *myb* gene function provides gearing for the control of flavonoid biosynthesis in *Antirrhinum* flowers, *Plant Cell*, 1996, **8**, 1519-1532.
25. YU, L.M., LAMB, C.J., DIXON, R. A., Purification and biochemical characterization of proteins which bind to the H-box *cis*-element implicated in transcriptional activation of plant defense genes, *Plant J.*, 1993, **3**, 805-816.
26. DA COSTA E SILVA, O., KLEIN, L., SCHMELZER, E., TREZZINI, G.F., HAHLBROCK, K., BPF-1, a pathogen-induced DNA-binding protein involved in the plant defense response, *Plant J.*, 1993, **4**, 125-135.
27. LUGERT, T., WERR, W., A novel DNA-binding domain in the *Shrunken* initiator-binding protein (IBP1), *Plant Mol. Biol.*, 1994, **25**, 493-506.
28. YU, E.Y., KIM, S.E., KIM, J.H., KO, J.H., CHO, M.H., CHUNG, I.K., Sequence-specific DNA recognition by the Myb-like domain of plant telomeric protein RTBP1, *J. Biol. Chem.*, 2000, **275**, 24206-24214.
29. DRÖGE-LASER, W., KAISER, A., LINDSAY, W.P., HALKIER, B.A., LOAKE, G.J., DOERNER, P., DIXON, R.A., LAMB, C., Rapid stimulation of a soybean protein-serine kinase that phosphorylates a novel bZIP DNA-binding protein, G/HBF, during the induction of early transcription-dependent defenses, *EMBO J.*, 1997, **16**, 726-738.
30. KOES, R.E., QUATTROCCHIO, F., MOL, J.N.M., The flavonoid biosynthetic pathway in plants: Function and evolution, *BioEssays*, 1994, **16**, 123-132.
31. MOL, J., GROTEWOLD, E., KOES, R., How genes paint flowers and seeds, *Trends Plant Sci.*, 1998, **3**, 212-217.
32. DI CARLO, G., MASCOLO, N., IZZO, A.A., CAPASSO, F., Flavonoids: Old and new aspects of a class of natural therapeutic drugs, *Life Sci.*, 1999, **65**, 337-353.
33. DOONER, H.K., ROBBINS, T.P., JORGENSEN, R.A., Genetic and developmental control of anthocyanin biosynthesis, *Annu. Rev. Genet.*, 1991, **25**, 173-199.

34. SHIRLEY, B.W., Flavonoid biosynthesis: 'New' functions for an 'old' pathway, *Trends Plant Sci.*, 1996, **1**, 377-382.
35. PAZ-ARES, J., GHOSAL D., WEINLAND, U., PETERSON, P.A., SAEDLER, H., The regulatory *c1* locus of *Zea mays* encodes a protein with homology to *myb* proto-oncogene products and with structural similarities to transcriptional activators, *EMBO J.*, 1987, **6**, 3553-3558.
36. CONE, K.C., COCCIOLONE, S.M., BURR, F.A., BURR, B., Maize anthocyanin regulatory gene *pl1* is a duplicate of *c1* that functions in the plant, *Plant Cell*, 1993, **5**, 1795-1805.
37. LUDWIG, R., HABERA, L.F., DELLAPORTA, S.L., WESSLER, S.R., *Lc*, a member of the maize *R* gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcriptional activators and contains the *myc*-homology region, *Proc. Natl. Acad. Sci. USA*, 1989, **86**, 7092-7096.
38. FITCH, W.M., Distinguishing homologous from analogous proteins, *Systematic Zool.*, 1970, **19**, 99-113.
39. QUATTROCCHIO, F., WING, J.F., LEPPEN, H.T.C., MOL, J.N.M., KOES, R.E., Regulatory genes controlling anthocyanin pigmentation are functionally conserved among plant species and have distinct sets of target genes, *Plant Cell*, 1993, **5**, 1497-1512.
40. STYLES, E.D., CESKA, O., Genetic control of 3-hydroxy- and 3-deoxy-flavonoids in *Zea mays*, *Phytochemistry*, 1975, **14**, 413-415.
41. STYLES, E.D., CESKA, O., The genetic control of flavonoid synthesis in maize, *Can. J. Genet. Cytol.*, 1977, **19**, 289-302.
42. STYLES, E.D., CESKA, O., Pericarp flavonoids in genetic strains of *Zea mays*, *Maydica*, 1989, **34**, 227-237.
43. GROTEWOLD, E., ATHMA, P., PETERSON, T., Alternatively spliced products of the maize *P* gene encode proteins with homology to the DNA-binding domain of Myb-like transcription factors, *Proc. Natl. Acad. Sci. USA*, 1991, **88**, 4587-4591.
44. RABINOWICZ, P.D., BRAUN, E.L., WOLFE, A.D., BOWEN, B., GROTEWOLD, E., Maize *R2R3 Myb* genes: Sequence analysis reveals amplification in higher plants, *Genetics*, 1999, **153**, 427-444.
45. GROTEWOLD, E., Does P protein require a partner, as C1 protein does?, *Maize Genet. Coop. News*, 1995, **69**, 32.
46. HASHIMOTO, T., YAMADA, Y., Alkaloid biogenesis: Molecular aspects, *Annu. Rev. Plant Physiol.*, 1994, **45**, 257-285.
47. BENDER, J., FINK, G.R., A Myb homologue, ATR1, activates tryptophan gene expression in *Arabidopsis*, *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 5655-5660.
48. MENKE, F.L.H., PARCHMANN, S., MUELLER, M. J., KIJNE, J. W., MEMELINK, J., Involvement of the octadecanoid pathway and protein phosphorylation in fungal elicitor-induced expression of terpenoid indole alkaloid biosynthetic genes in *Catharanthus roseus*, *Plant Physiol.*, 1999, **199**, 1289-1296.
49. VAN DER FITS, L., MEMELINK, J., ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism, *Science*, 2000, **289**, 295-297.

50. RADWANSKI, E.R., LAST, R.L., Tryptophan biosynthesis and metabolism: Biochemical and molecular genetics, *Plant Cell*, 1995, **7**, 921-934.
51. NIYOGI, K.K., FINK, G., Two anthranilate synthase genes in *Arabidopsis*: Defense-related regulation of the tryptophan pathway, *Plant Cell*, 1992, **4**, 721-733.
52. TSUJI, J., JACKSON, E.P., GAGE, D.A., HAMMERSCHMIDT, R., SOMERVILLE, S., Phytoalexin accumulation in *Arabidopsis thaliana* during the hypersensitive reaction to *Psuedomonas syringae* pv *syringae*, *Plant Physiol.*, 1992, **98**, 1304-1309.
53. ZHAO, J., LAST, R.L., Coordinate regulation of the tryptophan biosynthetic pathway and indolic phytoalexin accumulation in *Arabidopsis*, *Plant Cell*, 1996, **12**, 2235-2244.
54. ROBERTS, M.F., STRACK, D., Biochemistry and physiology of alkaloids and betalains, *in* Biochemistry of Plant Secondary Metabolism (M. Wink, ed.) CRC Press LLC, Baton Rouge, 1999, pp. 358.
55. MEIJER, A.H., VERPOORTE, R., HOGE, J.H.C., Regulation of enzymes and genes involved in terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*, *J. Plant Res.*, 1993, **3**, 145-164.
56. CONTIN, A., VAN DER HEIJDEN, R., LEFEBER, A.W.M., VERPOORTE, R., The iridoid glucoside secologanin is derived from the novel triose phosphate/pyruvate pathway in a *Catharanthus roseus* cell culture, *FEBS Lett.*, 1998, **434**, 413-416.
57. MENKE, F.L.H., CHAMPION, A., KIJNE, J.W., MEMELINK, J., A novel jasmonate- and elicitor-responsive element in the periwinkle secondary metabolite biosynthetic gene *Str* interacts with a jasmonate- and elicitor-inducible AP2-domain transcription factor, ORCA2, *EMBO J.*, 1999, **18**, 4455-4463.
58. ST-PIERRE, B., VAZQUEZ-FLOTA, F., DE LUCA, V., Multicellular compartmentation of *Catharanthus roseus* alkaloid biosynthesis predicts intercellular translocation of a pathway intermediate, *Plant Cell*, 1999, **11**, 887-900.
59. REICHMANN, J.L., MEYEROWITZ, E.M., The AP2/EREBP family of plant transcription factors, *Biol. Chem.*, 1998, **379**, 633-646.
60. PASQUALI, G., ERVEN, A.S.W., OUWERKERK, P.B.F., MENKE, F.L.H., MEMELINK, J., The promoter of the strictosidine synthase gene from periwinkle confers elicitor-inducible expression in transgenic tobacco and binds nuclear factors GT-1 and GBF, *Plant Mol. Biol.*, 1999, **39**, 1299-1310.
61. ZHOU, D.-X, Regulatory mechanism of plant gene transcription by GT-elements and GT-factors, *Trends Plant Sci.*, 1999, **4**, 210-214.
62. VAN DER FITS, L., ZHANG, H., MENKE, F.L.H., DENEKA, M., MEMELINK, J., A *Catharanthus roseus* BPF-1 homologue interacts with an elicitor-responsive region of the secondary metabolic biosynthetic gene *Str* and is induced by elicitor via a JA-independent signal transduction pathway, *Plant Mol. Biol.*, 2000, **44**, 675-685.
63. LLOYD, A.M., WALBOT, V., DAVIS, R.W., *Arabidopsis* and *Nicotiana* anthocyanin production activated by maize regulators *R* and *C1*, *Science*, 1992, **258**, 1773-1775.
64. GOODRICH, J., CARPENTER, R., COEN, E.S., A common gene regulates pigmentation pattern in diverse plant species, *Cell*, 1992, **68**, 955-964.

65. GROTEWOLD, E., CHAMBERLAIN, M. , ST. CLAIRE, G., SWENSON, J., SIAME, B.A., BUTLER, L.G., SNOOK, M., BOWEN, B., Engineering secondary metabolism in maize cells by ectopic expression of transcription factors, *Plant Cell*, 1998, **10**, 721-740.
66. BRUCE, W., FOLKERTS, O., GARNAAT, C., CRASTA, O., ROTH, B., BOWEN, B., Expression profiling of the maize flavonoid pathway genes controlled by estradiol-inducible transcription factors CRC and P, *Plant Cell*, 2000, **12**, 65-79.
67. GROTEWOLD, E., PETERSON T., Isolation and characterization of a maize gene encoding chalcone flavanone isomerase, *Mol. Gen. Genet.*, 1994, **242**, 1-8.
68. JACKSON, D., CULIANEZ-MACIA, F., PRESCOTT, A.G., ROBERTS, K., MARTIN C., Expression patterns of *myb* genes from *Antirrhinum* flowers, *Plant Cell*, 1991, **3**, 115-125.
69. DIXON, R.A., STEELE, C.L., Flavonoids and isoflavonoids - a gold mine for metabolic engineering, *Trends Plant Sci.*, 1999, **4**, 394-400.
70. KASUGA, M., LIU, Q., MIURA, S., YAMAGUCHI-SHINOZAKI, K., SHINOZAKI, K., Improving plant drought, salt and freezing tolerance by gene transfer of a single stress-inducible transcription factor, *Nature Biotechnol.*, 1999, **17**, 287-291.
71. KOSSEL., Über die chemische zusammensetzung der zelle. Verhandlungen der Berliner , 1891, 181-186.
72. MEISSNER, R.C., HAILING, J., COMINELLI, E., DENEKAMP, M., FUERTES, A., GRECO, R., KRANZ, H., PENFIELD, S., PETRONI, K., URZAINQUI, A., MARTIN, C., PAZ-ARES, J., SMEEKENS, S., TONELLI, C., WEISSHAAR, B., BAUMANN, E., KLIMYUK, V., JONES, J.J.D., PEREIRA, A., WISMAN, E., BEVAN, M., Function search in a large transcription factor gene family in *Arabidopsis*: Assessing the potential of reverse genetics to identify insertional mutations in R2R3 *MYB* genes, *Plant Cell*, 1999, **11**, 1827-1840.
73. THE EUROPEAN UNION ARABIDOPSIS GENOME SEQUENCING CONSORTIUM & THE COLD SPRING HARBOR, WASHINGTON UNIVERSITY IN ST. LOUIS AND PE BIOSYSTEMS ARABIDOPSIS SEQUENCING CONSORTIUM., Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*, *Nature*, 1999, **402**, 769-777.
74. LIU, L., WHITE, M.J., MACRAE, T.H., Transcription factors and their genes in higher plants, *Eur. J. Biochem.*, 1999, **262**, 247-257.
75. THEISSEN, G., BECKER, A., DI ROSA, A., KANNO, A., KIM, J.T., MÜNSTER, T., WINTER, K.-U., SAEDLER, H., A short history of MADS-box genes in plants, *Plant Mol. Biol.*, 2000, **42**, 115-149.
76. REISER, L., SANCHEZ-BARACALDO, P., HAKE, S., Knots in the family tree: Evolutionary relationships and functions of knox homebox genes, *Plant Mol. Biol.*, 2000, **42**, 151-166.
77. LIPSICK, J. S., One billion years of Myb, *Oncogene*, 1996, **13**, 223-235.
78. BRAUN, E. L., GROTEWOLD, E., Newly discovered plant *c-myb*-like genes rewrite the evolution of the plant *myb* gene family, *Plant Physiol.*, 1999, **121**, 21-24.

79. JIN, H., MARTIN, C., Multifunctionality and diversity within the plant *MYB*-gene family, *Plant Mol. Biol.*, 1999, **41**, 577-585.
80. KRANZ, H., SCHOLTZ, K., WEISSHAAR, B., c-MYB oncogene-like genes encoding three MYB repeats occur in all major plant lineages, *Plant J.*, 2000, **21**, 231-235.
81. MARTIN, C., PAZ-ARES, J., MYB transcription factors in plants, *Trends Genet.*, 1997, **13**, 67-73.
82. MÜNSTER, T., PAHNKE, J., DI ROSA, A., KIM, J.T., MARTIN, W., SAEDLER, H., THEISSEN, G., Floral homeotic genes were recruited from homologous MADS-box genes preexisting in the common ancestor of ferns and seed plants, *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 2415-2420.
83. SHINOZAKI, K., YAMAGUCHI-SHINOZAKI, K., URAO, T., KOIZUMI, M., Nucleotide sequence of a gene from *Arabidopsis thaliana* encoding a *myb* homologue, *Plant Mol. Biol.*, 1992, **19**, 493-499.
84. SOLTIS, P.S., SOLTIS, D.E., CHASE, M.W., Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology, *Nature*, 1999, **402**, 402-404.
85. GAUT, B.S., DOEBLEY, J.F., DNA sequence evidence for the segmental allotetraploid origin of maize, *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 6809-6814.
86. BRAUN, E.L., GROTEWOLD, E., Diversification of the *R2R3 Myb* gene family and the segmental allotetraploid origin of the maize genome, *Maize Genet. Coop. Newsl.*, 1999, **73**, 26-27.
87. CHOPRA, S., BRENDEL, V., ZHANG, J., AXTELL, J., PETERSON T., Molecular characterization of a mutable pigmentation phenotype and isolation of the first active transposable element from *Sorghum bicolor, Proc. Natl. Acad. Sci. USA*, 1999, **96**, 15330-15335.
88. GU, X., Statistical methods for testing functional divergence after gene duplication, *Mol. Biol. Evol.*, 1999, **16**, 1664-1674.
89. MOUCHIROUD, D., GAUTIER, C., BERNARDI, G., Frequencies of synonymous substitutions in mammals are gene-specific and correlated with frequencies of non-synonymous substitutions, *J. Mol. Evol.*, 1995, **40**, 107-113.
90. SMYTH, D., A reverse trend - MADS functions revealed, *Trends Plant Sci.*, 2000, **5**, 315-317.
91. ROMERO, I., FUERTES, A., BENITO, M.J., MALPICA, J.M., LEYVA, A., PAZ-ARES, J., More than 80 *R2R3-MYB* regulatory genes in the genome of *Arabidopsis thaliana*, *Plant J.*, 1998, **14**, 273-284.
92. FELSENSTEIN, J., Confidence limits on phylogenies - an approach using the bootstrap, *Evolution*, 1985, **39**, 783-791.
93. FARRIS, J.S., ALBERT, V.A., KALLERSJO, M., LIPSCOMB, D., KLUGE, A.G., Parsimony jackknifing outperforms neighbor joining, *Cladistics*, 1996, **12**, 99-124.
94. BREMER, K., The limits of amino acid sequence data in angiosperm phylogenetic reconstruction, *Evolution*, 1988, **42**, 795-803.
95. PHILIPPE, H., LAURENT, J., How good are deep phylogenetic trees?, *Curr. Opin. Genet. Dev.*, 1998, **8**, 616-623.

96. CARELS, N., BERNARDI, G., Two classes of genes in plants, *Genetics*, 2000, **154**, 1819-1825.
97. SOMERVILLE, C., SOMERVILLE, S., Plant functional genomics, *Science*, 1999, **285**, 380-383.
98. SABLOWSKI, R.W.M., MEYEROWITZ, E.M.,A Homolog of *NO APICAL MERISTEM* is an immediate target of the floral homeotic genes *APETALA3/PISTILLATA*, *Cell*, 1998, **92**, 93-103.
99. FUGE, E.K., BRAUN, E.L., WERNER-WASHBURNE, M., Protein synthesis in long-term stationary-phase cultures of *Saccharomyces cerevisiae*, *J. Bacteriol.*, 1994, **176**, 5802-5813.
100. THIELLEMENT, H., BAHRMAN, N., DAMERVAL, C., PLOMION, C., ROSSIGNOL, M., SANTONI, V., DE VIENNE, D., ZIVY, M., Proteomics for genetic and physiological studies in plants, *Electrophoresis*, 1999, **20**, 2013-2026.
101. GROTEWOLD, E., Ectopic expression of P and R+C1 induce few new proteins, *Maize Genet. Coop. Newsl.*, 1999, **73**, 23-24.
102. PTASHNE, M., How eukaryotic transcriptional activators work, *Nature*, 1988, **335**, 683-689.
103. RABINOWICZ, P. D., MA, H., GROTEWOLD, E., Consequences of the ectopic expression of the Myb-domain protein P, *Maize Genet. Coop. Newsl.*, 1997, **71**, 21-22.