Plant secondary metabolism comprises an enormous diversity in compounds and enzymes, and wide spectra of mechanisms of gene regulation and of transport of metabolites and enzymes. Genetic approaches using the model plant Arabidopsis thaliana have contributed importantly to recent progress in understanding glucosinolate biosynthesis and its intricate linkage with auxin homeostasis. Arabidopsis genetics have also caused revolutionary changes in the existing views on the metabolic intermediates and enzyme activities that are involved in phenylpropanoid biosynthesis. Some progress has been achieved in understanding the transcriptional regulation of the flavonoid pathway. Transcriptional regulators have also been identified for glucosinolate and terpenoid indole alkaloid biosynthesis.

Classical genetics is a phenotype-driven approach to the dissection of biological processes, which does not require prior knowledge of the molecular mechanisms or information about genes, proteins, enzyme activities or metabolic intermediates. Using appropriate mutagens, such as transposons or T-DNA, it is possible to quickly connect the mutant phenotype to a certain gene. Reverse genetics comprises the engineering of changes in the genome using existing information about the targeted gene. Common approaches are ectopic overexpression of the gene and, as a complementary approach, post-transcriptional silencing by antisense RNA or RNA interference (RNAi). The introduction of targeted genetic changes by homologous recombination is still impossible in plants and, as an alternative, TILLING (Targeting Induced Local Lesions IN Genomes) has been developed. In a classical genetic approach, the selection of mutants relies on a screenable phenotype. In secondary metabolism, changes in the levels of metabolites that absorb specific light wavelengths are easily screenable phenotypes. The classic example is the screening with visible light for changes in the levels of anthocyanin pigments, which has allowed the isolation of mutants that are affected in many structural genes, in regulatory genes and in genes that encode transporters. Successful mutant screens have also been conducted for metabolites that absorb UV light. In addition, more elaborate screens using high-throughput metabolite detection methods are possible.

In petunia, maize and snapdragon, endogenous transposons have been used to generate mutations, allowing the rapid recovery of mutated genes. In Arabidopsis, which lacks naturally active transposons, foreign transposons have been introduced. Gene cloning after ethylmethanesulfonate (EMS) mutagenesis in Arabidopsis is facilitated by rapid advances in map-based cloning techniques. In addition to mutant selection, the natural genetic variation that occurs among ecotypes or cultivars can be exploited to discover quantitative trait loci (QTL) [7]. Isolating the genes that correspond to QTLs is a major challenge, however, which in Arabidopsis is facilitated by map-based cloning approaches. Knocking out known genes in Arabidopsis using reverse genetics approaches is now facilitated enormously by the development of saturated collections of transposon and T-DNA insertion lines. Genetic tools have not been established for the large majority of plant species, and this causes difficulties in the
mapping and subsequent cloning of mutated genes. Reverse genetics, on the other hand, is applicable to all plant species that can be genetically transformed. In addition, reverse genetic techniques that introduce dominant traits are applicable to plant cell cultures, which cannot be taken through meiosis.

**Glucosinolate biosynthesis in Arabidopsis**

In the past couple of years, remarkable advances in the understanding of glucosinolate biosynthesis in *Arabidopsis* have been achieved [8]. The research has deployed phenotype-driven classical genetics as well as gene-based reverse genetic studies that use mutants from transposon- and T-DNA-tagged *Arabidopsis* collections. As such, recent work on glucosinolate biosynthesis forms a showcase for the power of both *Arabidopsis* genetics and the available genetic tools for studying secondary metabolism. For example, piling up insertionional gene knockouts in a single plant line allows the analysis of functionally redundant genes, which escape classical genetic screening, for involvement in a process of interest.

The individual functions of two cytochromes P450 that are involved in aliphatic glucosinolate biosynthesis have been unraveled using transposon knockouts [9]. CYP79F1 turns out to be involved in the biosynthesis of short-chain and long-chain aliphatic glucosinolates from (n)homomethionine, whereas CYP79F2 is involved only in long-chain aliphatic glucosinolate biosynthesis. In the double-knockout mutant, aliphatic glucosinolate biosynthesis is completely abolished [10]. The *ref2* mutant, which had reduced levels of phenylpropanoid end products, was isolated using a screen for reduced epidermal fluorescence under UV light [11]. Surprisingly, this mutant turned out to be affected in the gene that encodes CYP83A1, which is involved in aliphatic glucosinolate biosynthesis one enzymatic step downstream of the CYP79-catalyzed step. Indeed, the *ref2* mutant had reduced levels of both short-chain and long-chain aliphatic glucosinolates. It is not well understood how CYP83A1 affects phenylpropanoid product levels. A detailed characterization of the *superroot1* (*sur1*) mutant, which is affected in a gene with hitherto unknown function, showed that SQR1 is the C-S lyase that produces the thiobhydroximic acid in glucosinolate biosynthesis [12]. The *sur1* mutant is completely devoid of aliphatic and indole glucosinolates. The super-root phenotype is attributed to increased auxin levels resulting from the accumulation of indole-3-acetaldoxime, the common upstream precursor for indole glucosinolates and auxin. This finding indicates an intricate linkage between glucosinolate biosynthesis and auxin homeostasis. A similar linkage was revealed by an inserional loss-of-function mutant that is affected in the gene encoding the glucosyltransferase UGT74B1 [13]. This knockout mutant had reduced glucosinolate levels, showing that UGT74B1 catalyzes the glucosylation reaction that occurs one enzymatic step downstream from the SQR1-catalyzed step in glucosinolate biosynthesis. The mutant had increased auxin levels and a corresponding phenotype.

Further analysis of a double-knockout mutant for the genes encoding CYP79B2 and CYP79B3, which catalyze the first step in indole glucosinolate biosynthesis from tryptophan (Trp), showed a total lack of the indole phytoalexin camalexin [14]. This finding demonstrates that the product of the CYP79-catalyzed reaction, indole-3-acetaldoxime, is a key branching point between indole glucosinolate, camalexin and auxin biosynthesis. Analysis of a knockout mutant of the gene *ATR1* (*ALTERED TYRPTOPHAN REGULATION1*, also known as *AtMYB34*) combined with the study of a transformant that overexpressed ATR1 showed that this MYB-type transcription factor controls the expression of genes that are involved in Trp biosynthesis, as well as of the *CYP79B2, CYP79B3* and *CYP83B1* genes, which are involved in Trp-derived indole glucosinolate biosynthesis [15*]. Interestingly, in a loss-of-function mutant affected in the *CYP83B1* gene, which encodes the cytochrome P450 that channels indole-3-acetaldoxime into indole glucosinolate biosynthesis, expression of the *ATR1* gene is upregulated. Probably as a consequence of this upregulation, the Trp biosynthesis gene *ANTHRANILATE SYNTHASE1* (*ASA1*) and the *CYP79B2* gene are also upregulated [16]. This indicates that the super-root phenotype of the *cyp83b1* mutant, which is also known as *sur2* [17], could be caused by increased auxin levels that occur both as a result of the accumulation of indole-3-acetaldoxime (because of lack of further processing) and because of increased production of this auxin precursor (because of the upregulation of the biosynthesis genes). The molecular mechanism behind the effect of the *cyp83b1* mutation on gene expression is not clear because *ATR1* gene expression was not induced, but rather was repressed, by exogenous auxin application [16].

**Phenylpropanoid biosynthesis in Arabidopsis**

*Arabidopsis* genetics has also facilitated progress in the understanding of phenylpropanoid biosynthesis, which has brought about some revolutionary changes in the existing views. Using bioinformatics-guided gene discovery and biochemistry, the postulated enzyme activity *p*-coumarate 3-hydroxylase (*C3H*), which introduces the hydroxyl group at the 3-position of the phenylpropanoid ring, was identified as a cytochrome P450 that uses the shikimate and quinate esters of *p*-coumarate as substrates [18]. The *C3H* gene was also isolated by positional cloning from the *ref8* mutant, which has reduced epidermal fluorescence under UV light because of a reduced level of sinapoylmalate [19]. The previously widely accepted view did not attribute the 3-hydroxylation to a cytochrome P450 activity, and envisaged free *p*-coumarate as the substrate.
Phenylpropanoid metabolism has long been thought to involve a series of reactions in which degradation of phenylalanine and a series of hydroxylation and O-methylation reactions generate hydroxycinnamic acids, which are converted to their corresponding hydroxycinnamaldehydes and hydroxycinnamyl alcohols. The ref1 mutant shows reduced epidermal fluorescence under UV light because of a reduced content of sinapic acid and ferulic acid. The REF1 gene has been found to encode an aldehyde dehydrogenase, which can convert coniferaldehyde and sinapaldehyde to ferulate and sinapate, respectively [20]. Therefore, this finding has caused a complete reversal in the previously accepted view of intermediate to end-product relationships in phenylpropanoid metabolism. In the current revised view, ferulate and sinapate feature as end products rather than as intermediates.

Condensed tannins or proanthocyanidins are flavonoid oligomers that are synthesized by a branch of the flavonoid pathway [3]. The brown color of the seed coat of Arabidopsis seeds is determined by these compounds; in the banysuls (ban) mutant, the seeds have a transparent testa (tt). The BAN enzyme was proposed to be a leucoanthocyanidin reductase [21]. Using BAN from Arabidopsis and Medicago truncatula, Xie et al. [22] demonstrated that BAN is in fact a reductase of anthocyanidin compounds located one step downstream from the leucoanthocyanidins. In accordance with this finding, an Arabidopsis mutant in which anthocyanidin synthase is defective lacks proanthocyanidins in the seed coat. The accumulation of proanthocyanidins in the Arabidopsis seed coat is also dependent on the MYB transcription factor TT2 [23], the basic-helix–loop–helix transcription factor TT8 [24], and the WD-repeat protein TRANSPARENT TESTA GLABRA1 (TTG1) [24]. By expressing glucocorticoid receptor fusion proteins in the corresponding mutant backgrounds, Baudry et al. [25] demonstrated that TT2, TT8 and TTG1 regulate expression of the BAN gene by direct interaction with its promoter. Yeast two- and three-hybrid experiments indicate that these proteins form a complex on the BAN promoter. Analysis of the new tt19 mutant allowed the isolation of a glutathione S-transferase-encoding gene, which is involved in anthocyanin and proanthocyanidin accumulation [26].

Phenotype-driven genetics in non-model species

Attempts to dissect secondary metabolism using classical genetics in non-model plant species are still scarce. In one example, a mutant screen was conducted in opium poppy following EMS mutagenesis [27]. Analysis of the latex of 8000 plants derived from EMS-treated seed led to the identification of the thebaine oripavine poppy1 (top1) mutant, which does not contain morphine or codeine but instead has elevated levels of the precursors thebaine and oripavine. Identification of the single mutated gene was attempted by analyzing global changes in gene expression with a 17 000-gene microarray, which led to the identification of 10 genes that were underexpressed in the top1 mutant. None of these genes was an obvious candidate for being related to the top1 phenotype. This demonstrates the difficulties in connecting a mutant phenotype with a gene in plant species for which genetic tools are lacking.

A powerful tool in phenotype-driven genetics is gain-of-function mutagenesis with a strong constitutive promoter that is carried on an insertion element such as Agrobacterium tumefaciens T-DNA. T-DNA activation tagging causes dominant gain-of-function mutations that are caused by the overexpression of the genes that flank the T-DNA insertion. It allows the recovery of mutant phenotypes that are conferred by functionally redundant genes. In addition, the T-DNA tag allows gene recovery from the primary transformant without a need for genetic crosses. T-DNA activation tagging has been successfully used in the model species Arabidopsis to identify the transcriptional regulator PAP1 (PRODUCTION OF ANTHOCYANIN PIGMENT1) [28]. Overexpression of this MYB-type protein upregulates multiple genes that are involved in anthocyanin biosynthesis and gives the transformant a purple color.

The T-DNA activation tagging technique has also been successfully applied to unravel secondary metabolism in non-model species. In tomato, T-DNA activation tagging allowed the identification of the MYB-type transcriptional regulator ANTHOCYANIN1 (ANT1). Overexpression of ANT1 upregulated multiple genes that are involved in the early and later steps of anthocyanidin biosynthesis, as well as genes that are involved in glycosylation and transport into the vacuole [29]. This overexpression caused an intense purple pigmentation throughout development [29*].

In Catharanthus roseus cells, T-DNA activation tagging of the octadecanoid-derivative-responsive Catharanthus APETALA2 (AP2)-domain3 (ORCA3) transcription factor conferred resistance to a toxic derivative of the metabolic intermediate tryptamine in the terpenoid indole alkaloid (TIA) pathway, which is detoxified by the TIA biosynthesis enzyme tryptophan decarboxylase [30]. Overexpression of ORCA3 upregulated several other genes in TIA biosynthesis besides that encoding tryptophan decarboxylase, as well as genes in precursor pathways. Additional evidence indicates that ORCA3 normally controls the jasmatic acid (JA)-responsive expression of TIA biosynthesis genes [31].

Reverse genetics in non-model species

Reverse genetics forms an integral part of the normal classical genetic cycle needed to confirm the gene-
phenotype relationship by complementing a mutant phenotype with the cloned candidate gene or by recapitulating a phenotype with a gene cloned from a T-DNA activation-tagged mutant. In the absence of mutants, however, reverse genetics is useful in uncovering processes in which a cloned gene is involved. One reverse genetic approach is the post-transcriptional silencing of gene expression using a variety of techniques, of which the most recent and successful is RNAi [4].

Reverse genetic approaches have been used to change the nicotine content in *Nicotiana attenuata*. One study involved virus-induced gene silencing in which constructs that expressed sense, antisense or inverted-repeat fragments corresponding to conserved parts of the two *PMT* genes, which encode putrescine methyltransferase, were introduced by *Agrobacterium*-mediated inoculation of the leaves. This led to a reduction in the corresponding transcript levels in the roots (the main site of nicotine biosynthesis) and a reduction of the basal and methyl jasmonate-induced nicotine content in the plants [32]. In a second study, silencing of *PMT* genes by a similar inverted repeat expressed from the CaMV 35S promoter led to reduced transcript levels in roots and reduced nicotine contents in leaves [33]. In contrast to wildtype plants, the silenced plants contained anabatine, which is presumably formed because of the accumulation of nicotine acid, which is normally coupled to a downstream product of *PMT* to form nicotine. A similar observation of elevated levels of anabatine at the expense of nicotine was made in *Nicotiana tabacum* plants expressing anti-sense *PMT* [34].

In opium poppy, codeinone reductase (COR), a terminal enzyme in morphine biosynthesis, has been knocked out by RNAi using a chimeric inverted repeat that targets all seven members of the gene family, leading to a significant reduction in the morphine and codeine levels of the seven intermediates and enzyme activities that are involved in general phenylpropanoid metabolism. Some progress has been achieved in understanding the basic network of intermediates and enzyme activities in this pathway. Enormous advances have also been made in understanding the basic network of glucosinolate biosynthesis, and its relationship to auxin homeostasis, in *Arabidopsis*.

In non-model plant species, the use of genetic approaches (with the exception of some reverse-genetic RNAi approaches) appears to have come to an almost complete standstill. It has been demonstrated that some genetic approaches, such as T-DNA activation tagging, can be successful with non-model species but this has not led to the expected explosion of similar approaches in a variety of plant species. It is possible that the timeframe is too short to expect results from such approaches, and that eventually such studies will be conducted successfully.

Understanding the basic network of metabolic intermediates and enzymes is a first step in unraveling secondary metabolism. Subsequent steps should aim at dissecting the regulatory mechanisms that control the enzyme-encoding genes, the mechanisms that control the activities of transcriptional regulators, and the mechanisms of intra- and inter-cellular transport of metabolites and enzymes. Some progress has been achieved in understanding the transcriptional regulation of the phenylpropanoid and flavonoid pathways. Transcriptional regulators have also been identified in glucosinolate and TIA biosynthesis. The picture is far from complete, however, and how the expression and the activities of these transcription factors are regulated remains largely mysterious [36]. The mechanisms of intra- and inter-cellular transport of metabolic intermediates and enzymes remain poorly understood aspects of secondary metabolism, but *Arabidopsis* genetics and genetic tools clearly offer promising opportunities for studying this neglected part of plant secondary metabolism.

Acknowledgements

I apologize to all those whose contributions I did not include in this review because of space constraints, personal preferences or simple oversight.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

* of special interest
** of outstanding interest


Further analysis of ATR1 function shows that this transcription factor regulates not only genes in tryptophan biosynthesis but also several genes in indole glucosinolate biosynthesis.


Analysis of the ref1 mutant overturns existing views on the transcriptional regulation of putrescine N-methyltransferase activity in transgenic Nicotiana tabacum L. can lead to elevated levels of anatamine at the expense of nicotine. Plant Mol Biol 2005, 59:87-105.

This paper provides a nice example of how drastic changes in metabolite patterns in a non-model species can be achieved using an RNAi approach. It also uncovers a previously unknown aspect of morphine biosynthesis: there appears to be a metabolic coupling between the terminal seven reaction steps.