Medicinal and aromatic plants: A case example of evolving secondary metabolome and biochemical pathway diversity

Suman P.S. Khanuja*, Tripta Jhang and Ajit Kumar Shasany*****

**Former Director, CIMAP (CSIR),*
C41-42, Double Storey, Ramesh Nagar, New Delhi - 110015, India
khanujazy@yahoo.com, spskhanuja@gmail.com
***Genetic Resources & Biotechnology Division,*
Central Institute of Medicinal and Aromatic Plants (CSIR),
Lucknow - 226015, India
jhangt@gmail.com,
****Senior Scientist, Genetic Resources & Biotechnology Division, Central Institute of Medicinal and Aromatic Plants (CSIR), Lucknow 226015*
akshasany@yahoo.com

Abstract: The unique capability of plants, in spite of being immobile in strict sense, to defend and respond precisely to environmental stresses whether biotic or abiotic is relatable to their ability to synthesize an array of phytochemicals as metabolites. We find these plants compounds useful for health care as nutraceuticals, drugs and medicines and even most attractive fragrances and flavors. This huge diversity of low molecular weight compounds is represented by secondary metabolites that confer the power of responding to stimuli in plants. Hence, the network of metabolic pathways in plant species represents the pool of functions and chemical diversity leading to biomolecules such as alkaloids, flavanoids, terpenoids, glycosides, etc. Although the structures of approximately 50,000 such metabolites have already been elucidated, there are probably hundreds of thousands of such compounds which we are not able to detect or decipher within the existing limitation of detection. Only a few of these are part of 'primary' metabolic pathways (those common to all organisms). The rest are termed 'secondary' metabolites; this term is historical and was initially associated with inessentiality but we know today their necessity in defence to signals and stimuli.

In the last decade, the research on plant secondary metabolism has been aiming to understand genetic basis from synthesis to regulation of plant compounds at molecular level. With these gene and genome studies, fascinating insights into the creation of genetic diversity of secondary metabolism have become evident. This has been leading to rising inquisition about the mechanisms of gene recruitment and diversification for novel functions. After more than 100 years of ignorance, biologists have reached a stage to recognize the causal connection between gene diversity and plasticity of secondary metabolism in its indispensable ecological role in the dynamic interactions of the plant

kingdom with its continuously changing environment. Nevertheless, the complexities are many folds. The multitude of metabolites found in living organisms and the calculated, unexpected small number of genes identified during genome sequencing projects pose more questions to the biologists. Several processes on the posttranscriptional and posttranslational level lead to the formation of enzyme diversity through structural and functional variant forms explaining partially this surprising situation. Further, lower enzyme specificity may also contribute to metabolome diversity. The bottom-line is the evolution of these genes which may provide multiple forms and thereby leading to metabolic diversity even for any cause including adaptation. The situation of pathway diversity across species and genera breaks the conventional taxonomic barriers and brings in chemo-taxonomic basis as new dimension to explain the ever evolving metabolome and hence vertical to horizontal genomic flow for the natural combinatorial biological chemistry working in plant kingdom as no surprise.

Keywords: Medicinal plants, Aromatic plants, Biochemical pathways, Secondary metabolism, Metabolome, Diversity, Evolution, Pathway genes, Metabolic enzymes

Introduction

Plants like other living forms in nature tend to conserve the energy in the metabolic activities but are known to produce and accumulate a large number of crucial and structurally very diverse molecules/metabolites. These metabolites directly may not serve as essential molecules of life for not being involved in the primary processes of growth and development. But such molecules are present in almost all plant forms from fungi to higher plants and are termed as secondary metabolites.[1] Secondary metabolites, however, are known to have crucial role in plant defence (against herbivores, microbes, viruses or competing plants) and signaling or to attract pollinating or seed dispersing animals. Thus, secondary metabolites represent plant's adaptive functions essential for its own survival and reproductive fitness aiding in its natural selection during evolution. The huge phytochemicals diversity found in plant kingdom is the result of different selection pressure conditions that plants have been successfully coping with through evolution.

Medicinal and aromatic plants have become most attractive category of plants both for researchers as well as industry mainly because of the abundant presence of these prized secondary metabolites which find diverse uses from drugs to industrial and agricultural or environmental applications. Structures of these secondary metabolites appear to have evolved during evolution in such a way that they can mimic the structures of endogenous substrates, hormones, neurotransmitters or other ligands. The distribution pattern of secondary metabolite in a given plant is complex but very dynamic as it precisely links with different tissues/developmental stages. Plant metabolites of this nature can exist as pro-active/pro-drug molecules that become activated upon wounding, infection or in

the body of a herbivore. The biosynthesis of some secondary metabolites is induced upon wounding or infection and is made de novo, for example, phytoalexins.[2] This quality of plants represents their capability too of producing enormous pool of molecules dynamically through the metabolic pathways diversity. Perhaps this diversity also is in an ever-evolving state through the continuum of genetic pool recombination, reshuffling and mutations happening and getting accumulated across the Darwinian forces and selection pressures. This state has been taken up here as the case study of evolving genomes in plants and discussed at length.

Evolving plant genomes: Emergence of duplicate genes, their clustering and divergence

Understanding the evolutionary process leading to the emergence, distribution, diversification and selection of genes involved in plants' metabolite biosyntheses is possible by visualising the ancestral forms vis-à-vis present phenotypes whether in primary or secondary metabolism. It is believed that in primary metabolism the new genes/alleles mainly arise by gene duplication followed by divergence.[3,4] This equips the organism with one gene that maintains the original function and a second copy that is not restricted by natural selection. This second copy can then accumulate mutations until, rarely, it has acquired a new function and might then become fixed in the population. So gene duplication is assumed to be the major driving force for diversification and gene recruitment.[1,5] If a gene that directs an essential function is duplicated, the duplicate is released from the stringent function of the mother gene; it may either be eliminated by inactivation (e.g. pseudogenization) or recruited for modified or new functions. There are at least two routes for duplicate genes to be recruited and functionalized. (1) Continuous modification of its function during plant speciation, as consequence of which large gene families with rich functional diversity arise and (2) the duplicate is recruited for a new stringent function in a novel biochemical environment, as consequence, a new single copy gene originates.[6] Further elaborating this rationale are the following examples and cases.

Amyrin synthases in oats are entirely distinct from other plants. The gene AsbAS1 seems to have arisen by duplication and divergence of a cycloartenol synthase-like gene, and later clustering with other genes required for distinct steps in avenacin biosynthesis in a region of the genome that is not conserved in other cereals. The components of this gene cluster are required for at least four clearly distinct enzymatic processes (2,3-oxidosqualene cyclization, amyrin oxidation, glycosylation and acylation), it is unlikely that the cluster has evolved as a consequence of duplication of a common ancestor. Although clusters of paralogous genes are common in plants (e.g. gene clusters for rRNA and specific disease resistance), reports of clusters of genes that do not share sequence relatedness and whose products contribute to a single selectable function are rare.[7] With a

series of mutants and their functional characterization, it has been shown that AsbAS1 might have evolved from an ancestral cycloartenol synthase-like gene cycloartenol by duplication and rapid sequence divergence. The close relatedness of AsbAS[1] to cycoartenol synthases is interesting because, although cycloartenol synthase and amyrin synthase both use, 2,3-oxidosqualene as a substrate, but the structures of the cyclization products generated are quite different. Although gene clusters for secondary metabolites are not well documented in plants, they are common in fungal biosynthetic components, specific pathway regulators and for auto-resistance to the end-product.[8] Transmission of these self-contained 'gene cassettes' by horizontal gene transfer has also been suggested as an explanation for the persistence of clustering in fungi although recent phylogenetic analyses underestimate the significance of vertical transmission.[9]

Clustering facilitates the inheritance of the genes for selective advantage as a single functional unit. Disruption of the gene cluster may lead to failure in producing desired pool of protective chemicals and further could result in the accumulation of deleterious intermediates.[7] Synthesis of avenacins, like many other plant secondary metabolites, is highly tissue specific and under strict developmental control. Such situations may also confer undefined selective advantages associated with physical proximity and position effects. Intimate coadaptation of all enzymes of a single pathway is likely to be important as an additional mechanism for strict control and containment of secondary metabolites and their pathway intermediates during synthesis. This coadaptation may extend to physical interactions among pathway components, which would aid the channelling of metabolic intermediates within multienzyme complexes.[10]

Genes for secondary metabolism may in turn be derived from genes for primary metabolism by gene duplication and divergence or possibly also by allelic polymorphism. Similarly, another group of the plant enzymes, terpene synthases (which collectively mediate production of a diverse class of natural products) are predicted to be derived from genes for primary metabolism by duplication and consequent divergence in structural and functional specialization. The example of the clustering of polyketide synthases (PKSs) genes from different plant species shows them to be into two categories. The two clusters represent chalcone synthases (CHSs) and nonchalcone-synthases of angiosperms. The latter apparently originated from ancient duplication of an ancestral CHS gene. During angiosperm speciation, it seems, one gene retained the essential CHS function while the duplicate underwent functional diversification. Under environmental selection pressure new polyketide synthases might have evolved, producing a rich diversity of polyketides that can be distinguished by their biogenetic starter units, the hierarchy through number of added C2-units and the mechanism involved in cyclization.[6]

The other example for the gene duplication worth mentioning is the evolution of benzoxazinone (DIMBOA) and indole.[7] Molecular characterization of DIMBOA pathway indicates that non-homologous genes are organized in a gene cluster. The first gene in the sequence (BX1) is believed to have originated from a duplication of the gene encoding the unknown subunit of ubiquitous tryptophan synthase followed by recruitment of the second gene duplicate for the emission of volatile indole. Thus, both gene duplicates acquire completely new functions. The first being involved in the biosynthesis of defence compounds under developmental control of young seedlings, the second, encoding indole-3-glycerol phosphate lyase (IGL) is induced in mature leaves in response to herbivore damage. IGL catalyzes the formation of indole as one of the volatiles in maize emitted as signals in tritrophic defence (the volatile signals allure parasitoids to their prey, i.e. the maize herbivores). Another example relates to homospermidine synthase (HSS), which has already been mentioned as the first pathway-specific enzyme in the biosynthesis of pyrrolizidine alkaloids. HSS evolved by duplication of the gene encoding desoxyhypusine synthase (DHS).[11] DHS takes part in the posttranslational activation of the eukaryotic initiation factor 5A (eIF5A). It is evident that HSS retained all kinetic and molecular properties of DHS except the ability to bind the eIF5A precursor protein.[12] The ability of synthesizing homospermidine from putrescine and spermidine, a side activity already existent in DHS, ultimately became the core-activity of HSS. Again the duplicate of an essential gene of primary metabolism here was recruited for an entirely divergent function in secondary metabolism. The third example has also been described by Hartmann[6] that concerns acyltransferases operating with 1-*O*-ß-acetalesters (1-*O*-glucose esters) as acyl donors instead of coenzymeA thio-esters. These enzymes, which play important role in plant phenylpropanoid metabolism, have most likely evolved from serine carboxy peptidases by gene duplication and new functionalization.[13,14]

Gene duplications are relatively frequent events within genomes and have a high impact on the evolution of new biological functions.[15,16] The most likely event after gene duplication is the production of pseudogenes from one of the gene copies by knockoff mutations. However, in rare cases, one copy may acquire a completely new function as a result of beneficial mutations within its regulatory and structural components. According to a model proposed by Hughes,[17] the evolution of functionally distinct daughter genes is preceded by a period in which the ancestral gene is bifunctional. This bifunctionality is accomplished by the deoxypussine synthase (DHS) protein that originally possesses homospermidine synthesizing activity in many separate angiosperm lineages, later becoming the exclusive activity of the other gene copy. The HSS-coding gene copy presumably lost the protein-modifying activity of DHS and escaped the strong selection pressure on this essential function of primary metabolism. Nevertheless, its remaining ability

to synthesize homospermidine became the object of selection pressure from herbivores enabling some plants to recruit the gene copy to establish the first step in the biosynthesis of defence compounds. There are several examples known in which genes have been independently recruited to a single function within a gene family. For instance, the resistance of insects to specific toxins seems to have been acquired through the independent recruitment of paralogous genes belonging to the cytochrome P450 superfamily.[18,19] Plant terpene synthases, such as limonene synthase, have been shown to be repeatedly recruited within their gene family from other terpene synthases[20] enabling some plants to recruit the gene copy to establish the first step in the biosynthesis of defence compounds.

In case of *Solanum habrochaites* (formerly *Lycopersicon hirsutum*) f. typicum LA1777 and cultivated tomato *Solanum lycopersicum* (formerly *Lycopersicon esculentum*), it has been observed that the glandular type VI trichomes present on the leaves and stems accumulate monoterpenes in *S. lycopersicum*[21] but sesquiterpenes, as insecticidal carboxylic acid derivatives, are accumulated in LA1777.[22] Enzymes related to germacrene C synthase mediate the accumulation of a group of structurally similar compounds termed class I sesquiterpenes (cI-Ss) in LA1777 and *S. lycopersicum*.[23] This represents the existence of a mechanism of secondary metabolites controlling the production of two distinct groups of sesquiterpenes from different precursor pools in the wild species. Such a partition in sesquiterpene biosynthesis could be the result of metabolite channeling, through two distinct farnesyldiphosphate (FPP) synthases in the cytoplasm for the secondary metabolites that are associated either with cI-S synthase or with plastidial transporters and the cII-S synthase.

Case examples of pathway genes diversity for secondary metabolism in plants

Secondary metabolites of high structural as well as functional similarities are well known to occur simultaneously across even unrelated families of the plant kingdom. For example, the anti-tumor alkaloid camptothecin (inhibitors DNA-topoisomerase) has been found in the following unrelated orders and families *Nothapodytes foetida* (Celastrales), *Pyrenacantha klaineana* (Icacinaceae), *Camptotheca acuminate* (Cornales: Nyssaceae), *Ophiorrhiza mungos*, *O. pumila, O. filistipula* (Rubiales: Rubiaceae), *Ervatamia heyneana* (Apocynaceae) and *Mostuea brunonis* (Loganiaceae). [2] Consequently, the co-occurrence of a structural class in two taxa could, but not necessarily, be an indication of a monophyletic relationship. This could be due to convergent evolution or differential gene expression wherein, it is likely that in some cases the genes that encode the enzymes for the production of a given structure or structural skeleton might have evolved early during evolution. These genes were not lost during phylogeny

but got switched off and on again at some later point.[24] An example of lupaine pathway shows that genes evolved early during evolution were turned on in some plants using the alkaloids as chemical defence substances but remained turned off in most instances.[25] Wink and Mohamed[26] to study the phylogenetics of the Leguminosae by selecting representative taxa covering a broad range of tribes established a large rbcL data set and using secondary metabolites as chemical defence traits, reducing the entire Leguminosae to 95 taxa.

Quinolizidine alkaloids are frequently present in all taxa of the subfamily Papilionoideae except the tribe Crotalarieae.[27] *Crotalaria* species sequester pyrrolizidine alkaloids and/or nonprotein amino acids but not Quinolizidine alkaloids. In the genus *Lotononis*, some taxa produce quinolizidine alkaloids and others produce pyrrolizidine alkaloids. Since *Crotalaria* and *Lotononis* have derived from same ancestors, producing quinolizidine alkaloids but not pyrrolizidine alkaloids, the genes encoding biosynthetic enzymes of quinolizidine alkaloid formation must still be present. More likely the quinolizidine alkaloid genes have been turned off in *Crotalaria* and partially in *Lotononis*. The formation of pyrrolizidine alkaloids (which are typical secondary metabolite of the Boraginaceae and some Asteraceae) instead appears to be a new acquisition for chemical defence, which probably evolved independently. The protease inhibitors (i.e. trypsin and chymotrypsin inhibitors) distribution pattern is also similar to quinolizidine alkaloids.[28] The members of the Caesalpinoideae and many Mimosoideae accumulate protease inhibitors in their seeds, where they serve concomitantly as chemical defence and nitrogen storage compounds. Within the Papilionoideae, protease inhibitors are prominent in the tribes Vicieae, Trifolieae, Cicereae, Abreae, Galegeae, Loteae, Phaseoleae, and Tephrosieae, but are not described in the Mirbelieae.

Withanolides represent a group of steroidal lactones with strong insecticidal properties which appear to be restricted to the family Solanaceae.[29,30] Withanolide producing genera are typical for the tribe Physaleae, but isolated occurrences have been reported for *Brugmansia* (Datureae), *Hyoscyamus* (Hyoscyameae), *Lycium* (Lycieae), *Jaborosa* (Jaboroseae), *Nicandra* (Nicandreae) and *Browallia* (Browallieae) serving as chemical defence compounds in the plants producing them. Therefore, these compounds constitute important fitness traits and represent adaptive characters with some, but usually have limited value as a taxonomic marker.[26] All members of a monophyletic group share a chemical characteristic; favouring their use as a taxonomic marker. In other instances, a particular secondary metabolite may occur in several unrelated clades and/or plant families.[26,28] The erratic secondary metabolite distribution can be due to simple convergence, where genes encoding a particular biosynthetic pathway evolved independently in several parts of a phylogeny. There is evidence however for an alternative explanation: In several cases, it is apparent that ancestral members of a group evolved the

biosynthetic capacity to produce a specific secondary metabolite. The absence of such a trait in phylogenetically derived groups is probably due to differential gene expression, where the corresponding genes are not lost but switched off. Since secondary metabolites play a vital role as defence and signal compounds, their occurrence apparently reflects adaptations and particular life strategies embedded in a particular phylogenetic framework.[2]

Khanuja *et al*.[31] while assessing the interspecific as well as intraspecific relationships in *Mentha* at molecular (DNA) level, observed that *M. gracilis* Sole Cardiaca showed a much higher similarity with *M. spicata* as well as *M. arvensis*, which amongst themselves showed rather a greater distance. This indicated that the species might have evolved as a natural hybrid between *M. arvensis* and *M. spicata*. The GC and GC/MS of the oil quality for metabolite components of 20 accessions of *M. piperita*[32] showed close relationship of mutant 'Kukrail' with Japanese oil, whereas other two accessions were similar to Chinese oil as analysed in the component plot. These genotypes were released as varieties by the names of 'CIM -Indus' (high pulegone and menthofuran content) and' CIM-Madhuras' (peppermint plant having aroma with high acceptability respectively by CIMAP. Similarly, the aromatic grass species *Cymbopogon*, in the analysis for essential oil biosynthetic components indicated remarkable variation among various species.[33] The major essential oil components citral 'a' and 'b', were detected in *Cymbopogon pendulus*, *C. flexuosus* and *C. citratus* with highest in *C. Citratus* where as it was not detected in *C.winteianus* (Jowii) The hierarchical cluster analysis based on essential oil composition placed *C. winterianus* distantly from all other taxa followed by the hybrid Jamrosa. This hybrid contained high geraniol (68%), low citral (less than 2%) and trace of citronellal (0.5%) in the essential oil composition. The genomic synteny in essential oil compositions along with the differences indicated gene duplication followed by variations to create the metabolite diversity during speciation.

Domain swapping and neo-funtionalization: Another level of functional divergence

Domain swapping represents an independent mechanism for the generation of new composite genes with or without prior gene duplications.[34] It is theoretically possible for a new allele in one of the plant's genetic loci to be selected for if it encodes the ability to make a new defence compound, whereas the older alleles still specify the synthesis of another defence compound that however may be no longer effective. Thus, in secondary metabolism, there is a potential for new genes to keep evolving without a prior gene duplication event. In such cases, these orthologous genes in related species might encode proteins with different functions[1]. It is widely believed that enzymes with more active and specialized

function evolve divergently from enzymes with promiscuous function. This process is thought to be closely associated with the evolution of metabolic pathways.[35–37] It has been observed that that the recruitment of single enzymes from other metabolic pathways might significantly drive the evolution of both enzymes and metabolic pathways.[35] It is believed that enzymes with promiscuous functions can be initially shared by two distinct metabolic pathways. Enzymes with promiscuous function might give organism novel metabolic capabilities and, thus, render them adaptable to different environments. When the gene is duplicated, one enzyme is free to abandon the role it had in the previous pathway and thus can specialize its function for the new pathway, and vice versa. This results in divergent molecular evolution of enzymes and a mosaic or patchwork evolution of metabolic pathways. If multiple steps in a metabolic network are catalyzed by a series of promiscuous enzymes, although inefficient, this network would be able to produce a large library of natural products. If any of these products were to be captured by positive selection, the metabolic network could then converge on the pathway through divergent evolution of each component enzyme.[38]

In a research study, a combination of domain swapping and reciprocal site-directed mutagenesis was carried out in grand fir between (–)-(4*S*)–limonene synthase (LS) and (–)-(4*S*)-limonene/ (–)-1S, 5*S*–α-pinene synthase (LPS) by Katoh *et al.*[39] Exchange of the predicted helix D through F region in LS gave rise to an LPS-like product outcome. Whereas reciprocal substitutions of four amino acids in LPS (two in the predicted helix D and two in the predicted helix F) altered the product distribution to that intermediate between LS and LPS. This resulted in a 5-fold increase in relative activity. Based on the results of these chimeric studies, reciprocal point mutations were made in each parent monoterpene synthase to provide five single and 33 multiple site mutants of LS, and five single and 12 multiple site mutants of LPS, and were successfully expressed and evaluated. The most effective in altering monoterpene olefin distribution was V384L, a position predicted by modeling to reside in helix D. In this mutant, the proportions of generated α- and β-pinene were reversed and the level of β-phellandrene was found to be nearly doubled. These experiments were conducted, in conjunction with modeling of the two enzymes to know the critical amino acids for product determination. But indirectly this denotes the change in domains among highly homologous genes can lead to change the product distribution itself.

In another study, domain swapping experiments between Cl(-) γ PINS (γ -pinene synthase) and Cl γ TS (γ -terpinene synthase) and between Cl(+)LIMS2 (limonene synthase) and Cl γ TS were conducted[40] to identify domains within the monoterpene synthase enzymes determining the product specificity. Similarly, a sesquiterpenes cyclase (CASC2) showing 77% amino acid identity with the previously cloned sesquiterpene cyclase CASC1 of *Capsicum annuum* failed to express in *Escherichia coli*. However, the chimeric construct of CASC2 in which

the amino terminal 164 amino acid was substituted by the equivalent portion of either CASC1 or tobacco sesquiterpene, the cyclase was capable of expressing the functional sesquiterpene cyclase activities.[41] Relatively closer similarity of GES (geraniol Synthase) and LIS ((R)-linalool synthase protein pairs in *Ocimum basilicum* indicates that further terpene biosynthetic diversity is continuing to be generated in the basil lineage by gene duplication and divergence.[42] This two genes have been found to produce only a single product, either linalool or geraniol, but not both. But the results of domain-swapping experiments by the same group also demonstrated that it is possible to generate a monoterpene synthase that can synthesize both geraniol and linalool. Here, multiple amino acids contribute to such a dual selectivity but such an enzyme is not found in nature till date.

Another phenomenon of importance observed relates to the occult biosynthetic capacities of plant either constitutive or induced. This confers them an unlimited potential to produce a large array of different compounds that can be activated when novel substrates become available.[43] Through biotechnological route, over-expression of *Clarkia breweri* floral gene linalool synthase (LIS), an enzyme that catalyzes the formation of (*S*)-linalool from the monoterpene precursor geranyl diphosphate was attempted in tomato fruit.[44] Interestingly, accumulation of (*S*)-linalool (expected) and 8-hydroxylinalool (unexpected) was observed. In this case, the availability of novel substrate due to expression of a foreign gene enabled endogenous 'occult' hydroxylase activity to act on (*S*)-linalool. When the Clarkia LIS gene was overexpressed in carnation flowers, linalool in the transgenic flowers further metabolized to linalool oxides. Thus, the overexpression of an identical gene in different target tissues and organisms could gave rise to distinct phenotypes, according to the metabolism present or induced in the target organism that might interact with the novel products generated.

Similarly, from lemon basil (*O. basilicum L. cv. Sweet Dani*) geraniol synthase (*GES*) gene was overexpressed in tomato fruits to modify the aroma and flavor.[45] Besides high levels of geraniol, eleven novel additional metabolites sharing a common chemical backbone accumulated. These were derived from geraniol i.e the monoterpene alcohols nerol and citronellol. This way, monoterpene aldehydes geranial, neral and citronellal, the monoterpenol esters geranyl, neryl and citronellyl acetate, geranic and neric acid and rose oxide accumulated in such transgenic tomatoes. In another study,[46] the sweet basil type that possesses OMT activities converting methylate chavicol to estragole, was compared with a distinct basil type that accumulates only methyl eugenol (a 3 methoxylated estragole derivative). When chavicol was provided as a substrate it was methylated to estragole in cell-free extracts, although the plant itself does not contain estragole. The "lemon-scented" basil line 197 is also known to contain citral (a mixture of the monoterpene aldehydes geranial and neral) but lacks volatile phenylpropenes in its essential oil. And yet cell-free extracts derived from this lemon-scented line

could readily convert *O*-methylate chavicol to estragole, eugenol to methyl eugenol or isoeugenol to methyl-isoeugenol although none of these compounds are present in the plant. This further indicates the presence of 'silent' *O*-methyltransferase activities in basil lines that may readily accept novel substrates if the ability to produce those substrates is acquired by breeding or mutation, easily yielding to novel chemotypes.

Biosynthesis of tryptophan, a precursor for indole alkaloids, is well known for the feedback inhibition mechanism to regulate its own production. Hence, its availability can be well understood as a limiting factor in biosynthesis of indole alkaloids. However, the introduction of *Arabidopsis thaliana* feedback-resistant anthranilate synthase (AtAS) and induction of tryptophan decarboxylase (TDC) in *Catharanthus roseus* hairy roots did not significantly improve downstream alkaloids even though the levels of early alkaloid precursors tryptamine and tryptophan increased.[47] This suggests that the availability of the early amino acid precursor is not limiting for TIA (Terpenoid indole alkaloids) biosynthesis, which confirms the finding that the availability of secologanin was the important rate-limiting step in TIA biosynthesis.[48]

A superfamily is supposed to be a group of enzymes related by divergent evolution.[49] Among these, the plant terpene synthase superfamily is interesting and the most cited example to explain divergent molecular evolution. These enzymes share a strikingly similar active site scaffold comprising several α-helices and catalyze the formation of diverse terpenes from several different classes of prenyl diphosphates through wide varieties of carbocation rearrangements. Terpene synthase subfamilies within angiosperms are more closely related to each other than are members in the same subfamily from gymnosperms. In each subfamily, terpene synthases from the same or related species are more closely related to each other than are ones from different species with similar catalytic mechanism based secondary metabolites.[20] These observations indicate that divergence in terpene synthase subfamilies evolved after angiosperms and gymnosperms separated and that of terpene synthases within each subfamily arose after a series of subsequent speciation events.[38] All terpene synthases so far described show promiscuous function. Among those, γ-humulene and dselinene synthases are very promiscuous sesquiterpene synthases that are constitutively expressed in *Abies grandis*, each catalyzing the formation of at least 52 and 36 sesquiterpenes, respectively. In addition, these enzymes can use geranyl diphosphate as a substrate and catalyze the formation of monoterpenes. Although the specific roles of these enzymes have not been identified, it is thought that they might create chemical libraries that are important in general defence against microbial invasion. By contrast, many other terpene synthases have highly specialized functions and are often found to have very specific roles in the formation of bioactive metabolites. For example, (+)-D-cadinene, vetispiradiene and 5-epi-aristolochene synthases catalyze the

first reaction step of phytoalexin (anti-fungal agents) production in various plant species and yield their respective sesquiterpenes with more than 98, 90 and 70% selectivity, respectively. Thus, it has been suggested that terpene synthases with specific functions have evolved from ones with promiscuous function after having been captured by positive selection.[38]

Acknowledgments: Prof. Sushil Kumar, former Director, CIMAP who happens to be our teacher showing the path through the pathways in plants and life inspiring this writing attempt.

References

1. Pichersky E, Gang DR (2000) Genetics and biochemistry of secondary metabolites in plants: An evolutionary perspective. Trend Plant Sci 5:439–445
2. Wink M (2003) Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. Phytochem 64:3–19
3. Ohno S (1990) In: *Evolution by Gene Duplication*. Springer, Verlag
4. Pichersky E (1990) Nomad DNA – a model formovement and duplication of DNA sequences implant genomes. Plant Mol Biol 15:437–448
5. Ober D (2005) Seeing double-gene duplication and diversification inplant secondary metabolism. Trends Plant Sci 10:444–449
6. Hartmann T (2007) From waste products to ecochemicals: Fifty years research of plant secondary metabolism. Phytochemistry 68:2831–2846
7. Gierl A, Frey M (2001) Evolution of benzoxazinone biosynthesis andindole production in maize. Planta 213:493–498
8. Ward TJ, Bielawski JP, Kistler HC, Sullivan E, O'Donnell K (2002) Ancestral polymorphism and adaptive evolution in the trichothecene mycotoxin gene cluster of phytopathogenic Fusarium. Proc Natl Acad Sci USA 99:9278–9283
9. Kroken S, Glass NL, Taylor JW, Yoder OC, Turgeon BG (2003) Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. Proc Natl Acad Sci USA 100:15670–15675
10. Burbulis IE, Winkel-Shirley B (1999) Interactions among enzymes of the Arabidopsis flavonoid biosynthetic pathway. Proc Natl Acad Sci USA 96:12929–12934
11. Ober D, Hartmann T (1999) Homospermidine synthase, the firstpathway-specific enzyme of pyrrolizidine alkaloid biosynthesis,evolved from deoxyhypusine synthase. Proc Natl Acad Sci USA 96:14777–14782
12. Ober D, Harms R, Witte L, Hartmann T (2002) Molecular evolution by change of function: alkaloid-specific homospermidine synthase retained all properties of deoxyhypusine synthase except binding theeIF5A precursor protein. J Biol Chem 278:12805–12815
13. Milkowski C, Strack D (2004) Serine carboxypeptidase-like a cyltransferases. Phytochemistry 65:517–524
14. Stehle F, Brandt W, Milkowski C, Strac D (2006) Structure determinants and substrate recognition of serine carboxypeptidase-like acyltransferases from plant secondary metabolism. FEBS Lett 580:6366–6374

15. Wagner A (1998) The fate of duplicated genes: Loss or new function? Bioessays 20: 785–788
16. Kondrashov FA, Rogozin IB, Wolf YI, Koonin EV (2002) Selection in the evolution of gene duplications. Genome Biol 3:1–9
17. Hughes AL (1994) The evolution of functionally novel proteins after gene duplication. Proc R Soc Lond B Biol Sci 256:119–124
18. Scott JG, Wen Z (2001) Cytochromes P450 of insects: The tip of the iceberg. Pest Manag Sci 57:958–967
19. Wilson TG (2001) Resistance of Drosophila to toxins. Annu Rev Entomol 46:545–571
20. Bohlmann J, Meyer-Gauen G, Croteau R (1998) Plant terpenoid synthases: Molecular biology and phylogenetic analysis. Proc Natl Acad Sci USA 95:4126–4133
21. Besser K, Harper A, Welsby N, Schauvinhold I, Slocombe S, Li Y, Dixon RA, Broun P (2009) Divergent Regulation of Terpenoid Metabolism in the Trichomes of Wild and Cultivated Tomato Species. Plant Physiol 149:499–514
22. Li L, Zhao Y, McCaig BC, Wingerd BA, Wang J, Whalon ME, Pichersky E, Howe GA (2004) The tomato homolog of Coronatine-Insensitive1 is required for the maternal control of seed maturation, jasmonate-signaleddefense responses, and glandular trichome development. Plant Cell 16:126–143
23. van der Hoeven RS, Monforte AJ, Breeden D, Tanksley SD, Steffens JC (2000) Genetic control and evolution of sesquiterpene biosynthesis in *Lycopersicon esculentum* and *L. hirsutum*. Plant Cell 12:2283–2294
24. Wink M, Witte L (1983) Evidence for a wide spread occurrence of the genes of quinolizidine alkaloid biosynthesis. Induction of alkaloid accumulation in cell suspension cultures of alkaloid-free species. FEBS Letters 159:196–200
25. Wink M (1992) The role of quinolizidine alkaloids in plant insect interactions. In: *Insect–Plant Interactions*. Bernays EA (Ed.), CRC Press, Boca Rato
26. Wink M, Mohamed GIA (2003) Evolution of chemical defence traits in the Leguminosae: Mapping of distribution patterns ofsecondary metabolites on a molecular phylogeny inferred fromnucleotide sequences of the rbcL gene. Biochem. Syst Ecol 31:897–917
27. Kinghorn AD, Balandrin MF (1984) Quinolizidine alkaloids of the Leguminosae: Structural types, analysis, chemotaxonomy, and biological activities. In: *Alkaloids, Chemical and Biological Perspectives*. Pelletier WS (Ed.), Wiley, New York
28. Wink M, Waterman P (1999) Chemotaxonomy in relation to molecular phylogeny of plants. In: *Biochemistry of Plant Secondary Metabolism, Annual Plant Reviews*. Wink M (Ed.), Sheffield Academic Press and CRC Press
29. Hegnauer R (1973) Chemotaxonomie der Pflanzen. Birkha user Verlag, Basel
30. Harborne, JB, Baxter H (1993) Phytochemical Dictionary. In: *A Handbook of Bioactive Compounds from Plants*. Taylor & Francis, London
31. Khanuja SPS, Shasany AK, Srivastav A, Kumar S (2000) Assessment of genetic relationships in Mentha species. Euphytica 111:121–125
32. Shasany AK, Gupta S, Gupta MK, Singh AK, Naqvi AA, Khanuja SPS (2007) Chemotypic comparison of AFLP analyzed Indian Peppermint germplasm to selected peppermint oils of other countries. J Essent Oil Res 19:138–145
33. Khanuja SPS, Shasany AK, Pawar A, Lal RK, Darokar MP, Naqvi AA, Rajkumar S, Sundaresan V, Lal N, Kumar S (2005) Essential oil constituents and RAPD markers to

establish species relationship in Cymbopogon Spreng (Poaceae). Biochem Syst Ecol 33:171–186

34. Doolittle RF (1995) The multiplicity of domains in proteins. Ann Rev Biochem 64: 287–314
35. Schmidt S, Sunyaev S, Bork P, Dandekar T (2003) Metabolites: A helping hand for pathway evolution? Trends Biochem Sci 28:336–341
36. O'Brien PJ, Herschlag D (1999) Catalytic promiscuity and the evolution of new enzymatic activities. Chem Biol 6:R91–R105
37. Jensen RA (1976) Enzyme recruitment in evolution of new function. Ann Rev Microbiol 30:409–425
38. Yoshikuni Y, Keasling JD (2007) Pathway engineering by designed divergent evolution. Curr Opin Chem Biol 11:233–239
39. Katoh S, Hyatt D, Croteau R (2004) Altering product outcome in Abies grandis (−)-limonene synthase and (−)-limonene/(−)-α-pinene synthase by domain swapping and directed mutagenesis. Arch Biochem Biophys 425:65–76
40. Tamer MKE, Lucker J, Bosch D, Verhoeven HA, Francel WAV, Schwab W, Tunen AJV, Voragen AGJ, Maagd RA, Bouwmeester H (2003) Domain swapping of Citrus limon monoterpene synthases: Impact on enzymatic activity and product specificity. Arch Biochem Biophys 411:196–203
41. Back KNJ, Lee SB, Song JH, Shin DH , Kim HY (2000) Cloning of a sesquiterpene cyclase and its functional expression by domain swapping strategy. Mol Cells 10:220–225
42. Iijima Y, Davidovich-Rikanati R, Fridman E, Gang DR, Bar E, Lewinsohn E, Pichersky E (2004) The biochemical and molecular basis for the divergent patterns in the biosynthesis of terpenes and phenylpropenes in the peltate glands of three cultivars of basil. Plant Physiol 136:3724–3736
43. Lewinsohn EG (2009) Phytochemical diversity: The sounds of silent metabolism. Plant Sci 176:161–169
44. Lucker J, Bouwmeester HJ, Schwab W, Blaas J, van der Plas LH, Verhoeven LH (2001) Expression of Clarkia S-linalool synthase in transgenic petunia plants results in the accumulation of S-linalyl-beta-d-glucopyranoside. Plant J 27:315–324
45. Davidovich-Rikanati R, Tadmor S, Iijima Y, Bilenko N, Bar E, Carmona B, Fallik E, Dudai N, Simon JE, Pichersky E, Lewinsohn E (2007) Enrichment of tomato flavor by diversion of the early plastidial terpenoid pathway. Nat Biotechnol 25:899–901
46. Xie ZJ, Kapteyn DR, Gang A (2008) Systems biology investigation of the MEP/terpenoid and shikimate/phenylpropanoid pathways points to multiple levels of metabolic control in sweet basil glandular trichomes. Plant J 54:349–361
47. Leonard E, Weerawat R, O'Connor S, Prather KJ (2006) Opportunities in metabolic engineering to facilitate scalable alkaloid production. Nat Chem Biol 5:292–272
48. Shukla AK, Shasany AK, Gupta MM, Khanuja SPS (2006) Transcriptome analysis in *Catharanthus roseus* leaves and roots for comparative terpenoid indole alkaloid profiles. J Experimen Botany 57:3921–3932
49. Glasner ME, Gerlt JA, Babbitt PC (2006) Evolution of enzyme superfamilies. Curr Opin Chem Biol 10:492–497