South African Journal of Botany 89 (2013) 164–175

Contents lists available at ScienceDirect

South African Journal of Botany

journal homepage: www.elsevier.com/locate/sajb

Evolution of secondary metabolites in legumes (Fabaceae)

M. Wink *

Heidelberg University, Institute of Pharmacy and Molecular Biotechnology, INF 364, D-69120 Heidelberg, Germany

ARTICLE INFO

Available online 11 July 2013

Edited by B-E Van Wyk

Keywords:
Horizontal gene transfer
Evolution of secondary metabolisms
Molecular phylogeny
Chemotaxonomy
Function of secondary metabolites
Fabaceae
Leguminosae

ABSTRACT

Legumes produce a high diversity of secondary metabolites which serve as defence compounds against herbivores and microbes, but also as signal compounds to attract pollinating and fruit-dispersing animals. As nitrogen-fixing organisms, legumes produce more nitrogen containing secondary metabolites than other plant families. Compounds with nitrogen include alkaloids and amines (quinolizidine, pyrrolizidine, indolizidine, piperidine, pyridine, pyrrolidine, simple indole, *Erythrina*, simple isoquinoline, and imidazole alkaloids; polyamines, phenylethylamine, tyramine, and tryptamine derivatives), non-protein amino acids (NPAA), cyanogenic glucosides, and peptides (lectins, trypsin inhibitors, antimicrobial peptides, cyclotides). Secondary metabolites without nitrogen are phenolics (phenylpropanoids, flavonoids, isoflavones, catechins, anthocyanins, tannins, lignans, coumarins and furanocoumarins), polyketides (anthraquinones), and terpenoids (especially triterpenoid, steroidal saponins, tetraterpenes). While some secondary metabolites have a wide distribution (flavonoids, triterpenes, pinitol), however, others occur in a limited number of taxa. The distributions of secondary metabolites with an irregular occurrence are mapped on a molecular phylogeny of the Fabaceae, reconstructed from a combined data set of nucleotide sequences from *rbc*L, matK and ITS genes. In most cases, the distribution patterns of secondary metabolites do not agree with the phylogeny of the plants producing them. In contrary, the distribution of many secondary metabolites is patchy and irregular. Thus, the use of phytochemical data to reconstruct a phylogeny of plants is often not informative and can be misleading. The patchy distribution may be due to convergent evolution, a contribution of endophytic fungi or more likely, to an early acquisition of the key genes of secondary metabolism in the evolution of land plants among others by horizontal gene transfer from bacteria. Thus it would be a matter of gene regulation whether these genes are active in some but not all taxa.

1. Introduction

Phytochemical investigations have revealed a high structural diversity of plant secondary metabolites, comprising more than 21,000 alkaloids, 700 non-protein amino acids (NPAA), 200 cyanogenic glucosides and glucosinolates, >20,000 terpenoids, >10,000 polyphenols, >1500 polyacetylenes and fatty acids, 750 polyketides, and 200 carbohydrates (reviewed in Bell and Charlwood, 1980; Conn, 1981; Harborne, 1993; Roberts and Wink, 1998; Seigler, 1998; Dewick, 2002; DNP, 1996; Wink, 2008a, 2010a,b).

The synthesis and storage of secondary metabolites can be regarded as a strategy of plants for defence and communication. Plants are sessile and cannot run away when attacked by herbivores nor do they have the complex immune system of animals against bacteria, fungi, viruses and parasites. In order to defend themselves against herbivores, competing plants and pathogens, plants have evolved a diversity of secondary metabolites with a wide range of pharmacological and toxicological properties (reviewed in Fraenkel, 1959; Levin, 1976; Swain, 1977; Rosenthal and Berenbaum, 1991; Brown and Trigo, 1995; Wink, 1988, 1993a, 2007, 2008a; Wink and Schimmer, 2010). In addition, plants employ secondary metabolites for communication as signal compounds to attract pollinating insects, fruit dispersing animals, or rhizobial bacteria (reviewed in Cipollini and Levey, 1997; Wink, 2008a). Some compounds serve for nitrogen storage, UV protection and as antioxidative agents (reviewed in Hartmann, 2007; Wink, 1988, 2003, 2008a,b).

Plants not only synthesize the defence compounds but store them in high concentrations in the vacuole (in the case of hydrophilic compounds), resin ducts, trichomes, laticifers or cuticle (for lipophilic compounds) (reviewed in Wink, 1993c, 1997, 2010a,b), where they do not interfere with the plant's own metabolism. Some secondary metabolites are made de novo in case of an herbivore or pathogen attack (sometimes called phytoalexins). Plants always produce a complex mixture of secondary metabolites which usually consists of members from different groups, such as polyphenols, terpenoids and others (Wink, 2008b). There is experimental evidence that a synergistic potentiation of biological activities is achieved by combinations of individual defence compounds in a mixture (Wink, 2008b). The composition of individual compounds and their concentrations is not static but differs from organ to organ, within a developmental cycle of a plant and furthermore,

Abbreviations: CG, cardiac glycoside; FC, furanocoumarin; NPAA, Non-protein amino acid; QA, Quinolizidine alkaloids; PA, Pyrrolizidine alkaloids; $5HT_2$ R, 5-hydroxytryptamine receptor; mAChR, muscarinic acetylcholine receptor; nAChR, nicotinic acetylcholine receptor; GABA-R, GABA receptor; D_2, dopamine receptor D_2; α_2, adrenergic receptor α_2; AChE, acetylcholine esterase.

* Tel.: +49 6221 544881.
E-mail address: wink@uni-hd.de.

0254-6299/$ – see front matter © 2013 SAAB. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.sajb.2013.06.006

within and between populations. This variation, which leads to complex mixtures of secondary metabolites, is probably a strategy against the selection of specialised herbivores or pathogens (reviewed in Wink, 2003, 2008a,b). When antibiotics were applied in medicine as single entities, many bacteria have evolved resistance against them: If mixtures of antibiotics, which attack differing molecular targets in microbes, would have been employed instead, such a development, that generates a severe medical problem at present, could probably have been prevented.

2. Occurrence of secondary metabolites in Fabaceae

What had been discussed before for the occurrence and function of secondary metabolites in plants in general, more or less applies for members of the family Fabaceae. With 745 genera and over 19,500 species legumes represent the third largest plant family (reviewed in Lewis et al., 2005; www.mobot.org). Being such a large family, the enormous diversity of legume secondary metabolites does not surprise. Because legumes can fix atmospheric nitrogen (most members of Papilionoideae and Mimosoideae, but only 25% of Caesalpinoideae; Sprent and McKey, 1994), legumes can produce more nitrogen-containing secondary metabolites (especially, NPAAs, glucosinolates, amines, and alkaloids) than other non-nitrogen fixing plants. Interestingly, legumes produce fewer mono-, sesqui- and diterpenes than other plants (e.g., Asteraceae, Lamiaceae, Rutaceae). The nitrogen-containing defence compounds often accumulate in seeds where they serve a dual function: In addition to being toxic they are used as nitrogen storage compounds which are remobilized during germination and seedling development (Wink and Witte, 1984).

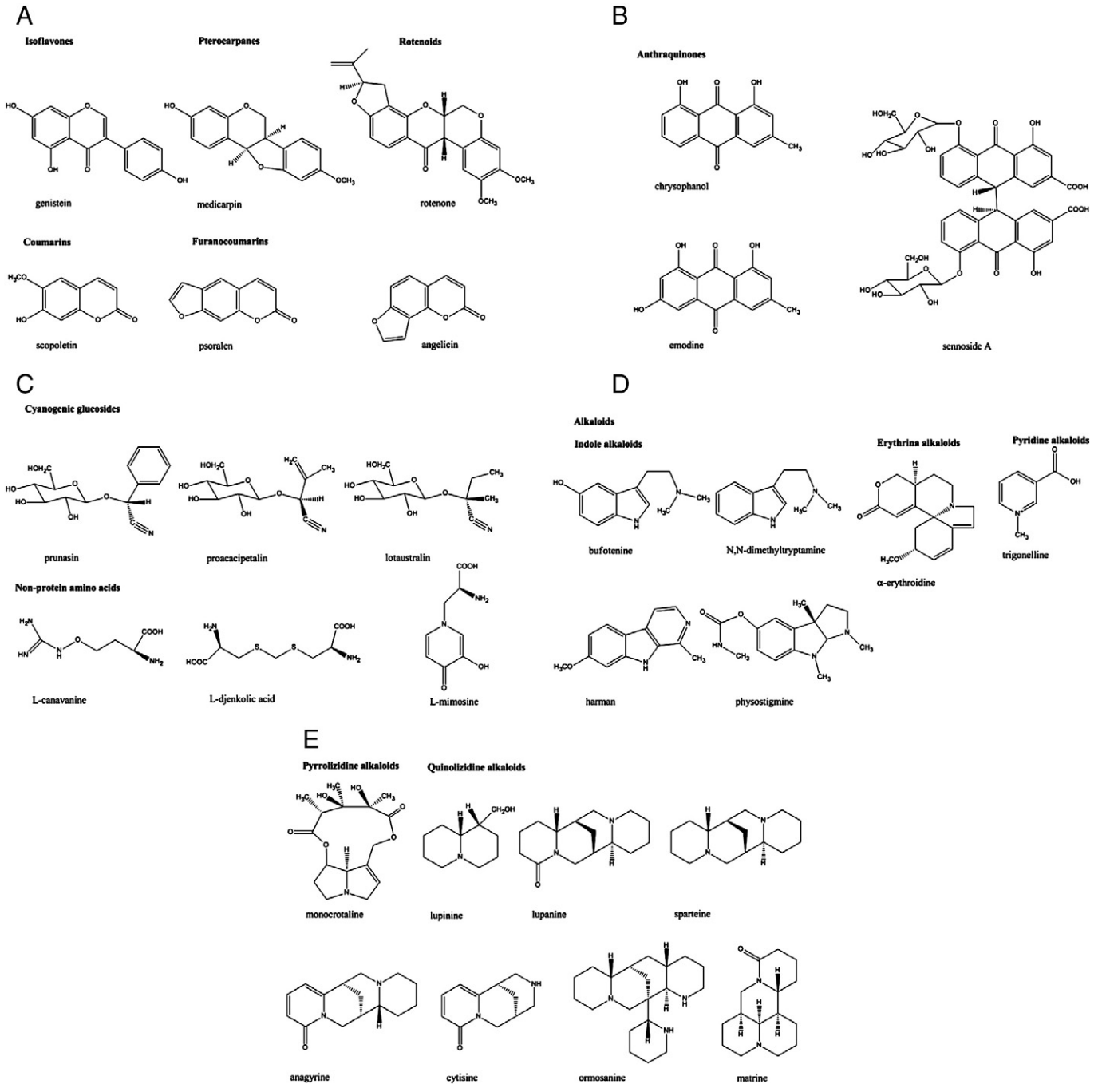

Fig. 1. Structures of some secondary metabolites of legumes. A. Isoflavones and coumarins; B. anthraquinones, C. cyanogenic glucosides and non-protein amino acids, D. simple indole, *Erythrina* and pyridine alkaloids, E. pyrrolizidine and quinolizidine alkaloids.

Table 1
Overview of secondary metabolites in the Fabaceae, their occurrence, pharmacological and toxicological properties.

Secondary metabolite	Examples from Fabaceae	Main occurrence	Pharmacological and toxicological activities
Alkaloids and amines			
Quinolizidine alkaloids (QA)	Sparteine, lupanine, anagyrine, cytisine, matrine, lupinine	Genistoid clade; *Ormosia* clade; *Sophora secundiflora*; *Calia, Bolusanthus*	Neurotoxins; modulation of nAChR and mAChR; Na^+ channel blocker
Pyrrolizidine alkaloids (PA)	Monocrotaline, senecionine	*Crotalaria*; *Lotononis*	Mutagenic and carcinogenic; modulators of several neuroreceptors, including $5HT_2$, mACh, GABA, D_2 and α_2.
Indolizidine alkaloids	Swainsonine, castanospermine	Astragaleae; *Castanospermum*	Inhibitor of endoplasmic hydrolases
Piperidine alkaloids	Ammodendrine 2-Piperidine carboxylic acid, 4-hydroxy-2-piperidine carboxylic acid	Genistoid clade Many members of all three subfamilies	Causes malformations in embryos
Pyridine alkaloids	Trigonelline	In all subfamilies; abundant in IRLC clade and Phaseoleae sens. lat.	Antimicrobial
β-Carboline alkaloids	Harman, harmalan, tetrahydroharman, leptocladine	*Petalostyles labicheoides*, *Acacia complanata*, *Burkea africana*, *Prosopis nigra*, *Desmodium gangeticum*	MAO inhibitor; serotonin receptor agonist; DNA intercalation; mutagenic
Simple indole alkaloids	Physostigmine	*Physostigma venenosum*, *Dioclea* spp.	AChE inhibitor
Erythrina alkaloids	Erysodine, erysopine, erythraline, erythroidine,	*Erythrina*	Neuromuscular blocking agent
Simple isoquinoline alkaloids	Salsoline, salsolidine	*Desmodium*, *Alhagi*, *Cytisus*, *Dendrolobium*	Dopamine antagonist
Imidazole alkaloids	Cynodine, cynometrine	*Cynometra* spp.	
Polyamines	Spermine, spermidine	Mainly Phaseoleae	Growth regulator
Phenylethylamines	*N*-Methyl phenylethylamine;	Mainly *Acacia* spp.; Caesalpininoideae	Psychoactive
Tyramines	*N*-Methyltyramine; hordenine; *N*-methylmescaline	Mimosoideae, Desmodieae	Psychoactive; insect feeding inhibitor
Tryptamines	*N,N*-Dimethyltryptamine; bufotenin; *N*-methyltryptamine	Mimosoideae	Serotonin receptor agonist; hallucinogenic
Histamine	*N*-Cinnamoylhistamine	*Acacia* spp., *Spartidium*	
NPAA	Canavanine, albiziine, carboxyethylcysteine, 2-acetamido-2-aminoproanoic acid, djenkolic acid, willardiine, homoarginine, mimosine, 4-hydroxypipecolic acid	Widely present in all tribes (except those with alkaloids)	Antimetabolites; anti-herbivore and antimicrobial activities
Cyanogenic glucosides	Prunasin, linamarin, lotaustralin, proacacipetalin	*Acacia* spp.; *Holocalyx balansae, Lotononis*spp., *Lotus* spp., *Ornithopus* spp., *Trifolium repens*, *Phaseolus lunatus*	Release HCN; inhibitor of respiratory chain; strong animal poison
Peptides			
Lectins	Abrin, robin	*Abrus precatorius*, *Robinia*	Inhibitors of ribosomal protein biosynthesis
Protease inhibitors	Trypsin inhibitors	Several Fabaceae	Inhibition of trypsin in herbivores
Antimicrobial peptides (AMP)	ApDef1	*Adenanthera* spp.	Potent antimicrobial
	Cyclotides	*Clitoria*	Antimicrobial
Phenolics			
Simple phenols	Vanillin, syringic acid, ferulic acid, gentisic acid, gallic acid, p-hydroxybenzaldehyde	Widely distributed	Antioxidants; antimicrobial
Flavonoids	Quercetin, kaempferol, etc.	Widely distributed, in all tribes	Antimicrobial and anti-herbivore activities; antioxidants
Isoflavones	Genistein, daidzein, formononetin,	Widely distributed only in SF Papilionoideae	Antioxidants; phytoestrogens; antimicrobial
Pterocarpans	Maackiain, glycinol, acanthocarpan, cristacarpin, glyceollin, medicarpin, phaseollin, pisatin, variabilin	Several members of Papilionoideae	Antimicrobial; antifungal; phytoalexin; antioxidants
Rotenoids	Rotenone	Amorpheae, Dalbergioids, Phaseoleae, Millettioids sens. strict.	Fish poison, insecticide; inhibits mitochondrial respiratory chain
Catechin	Catechin, epicatechin, catechin gallate, epigallocatechin gallate (EGCG)	Mostly trees	Antimicrobial and anti-herbivore activities; antioxidants
Anthocyanins	Delphinidin, peonidin, cyanidin	Widely distributed	Antioxidants; flower pigments; attraction of pollinators
Tannins	Mostly catechin type	Mostly trees	Antimicrobial and anti-herbivore activities
Lignans	Syringaresinol; hydnocarpin	A few taxa in Cercideae, Cassiaeae, Mimosoideae, *Chamaecrista*	Cytotoxic
Coumarins and furanocoumarins	Umbelliferone, scopolin, psoralen, bergapten, xanthotoxin	FC: mostly in tribe Psoraleae	FC: DNA intercalation and DNA alkylation; mutagenic; antimicrobial
Polyketides	Anthraquinones: chrysophanol, emodin, rhein	Few species in Cassiinae, *Ormosia* clade, Millettioid sens. strict.	DNA intercalator; mutagenic; causes drastic diarrhoea
Terpenoids			
Monoterpenes	Linalool, citronellol, limonene	Few species with fragrant flowers	Attracting pollinating insects; antimicrobial
Triterpenoid saponins	Ring skeleton: oleanane, lupane, and ursane	Widely distributed in all tribes	Interaction with biomembranes; cell lysis; cytotoxic, antimicrobial; antifungal
Steroids	Campesterol, β-sitosterol	Widely distributed in all tribes	Intercalate biomembranes
Cardenolides	Corotoxigenin, scorpioside, frugoside, hyrcanoside	A few taxa of *Coronilla* and *Securigera*	Inhibitors of Na^+/K^+ ATPase; strong poisons
Terpenoids			

Table 1 (*continued*)

Secondary metabolite	Examples from Fabaceae	Main occurrence	Pharmacological and toxicological activities
Tetraterpenes	Carotenoids	Widely distributed	Antioxidants; attraction of pollinating and fruit-dispersing animals
Carbohydrates	Pinitol (a methoxy inositol)	Widely distributed	Osmoticum
Organic acids	Fluoroacetic acid	In a few Australian taxa; *Gastrolobium*, *Gompholobium*, *Oxylobium*, and *Acacia*	Inhibitor of citric acid cycle; metabolic poison

The main secondary metabolites of legumes, which include alkaloids, NPPA, cyanogens, peptides, phenolics, polyketides, and terpenoids are summarized in Fig. 1 and Table 1 (reviewed in Harborne et al., 1971; Hegnauer and Hegnauer, 1994, 1996, 2001; Kinghorn and Balandrin, 1984; Seigler, 1998; Southon, 1994; Wink, 1993b; Veitch, 2010). Some secondary metabolites have a wide distribution (flavonoids, triterpenes, pinitol), however, others occur in a limited number of taxa (Table 1). It should be kept in mind that our information on the occurrence and distribution of secondary metabolites in legumes is incomplete because several taxa have not been studied so far. In other instances, phytochemists were rather interested to publish the finding of new compounds and not to report the detection of known metabolites. Furthermore, a central data base does not exist for secondary metabolites, which would include new phytochemical findings published later than the pioneering work of Hegnauer, Harborne and Southon (Harborne et al., 1971; Hegnauer and Hegnauer, 1994, 1996, 2001; Southon, 1994). Unfortunately, the existing literature also holds records of wrong identifications of secondary metabolites and of legumes, a fact which can distort the distribution patterns, described in this review.

Most of the secondary metabolites exhibit some biological, pharmacological or toxicological activity (Table 1) (Wink et al., 1998; reviewed in Teuscher and Lindequist, 2010). Many of the alkaloids are neurotoxins or neuromodulators (reviewed in Wink, 1992, 1993a, 2000, 2007; Wink and Schimmer, 2010) and probably evolved for defence against herbivores. Pyrrolizidine alkaloids become activated in the liver of herbivores; they then alkylate DNA which leads to mutations and even cancer in animals and humans (reviewed in McLean, 1970; Hartmann and Witte, 1995; Wink and Schimmer, 2010). A few alkaloids and amines of legumes exhibit psychotropic activities, such as bufotenine, *N*, *N*-demethyltryptamine, methylmescaline and β-carboline alkaloids

Fig. 2. Phylogeny of Fabaceae (cladogram) reconstructed from a 50% consensus tree based on an ML analysis of *rbc*L, matK and ITS nucleotide data from over 1276 legume taxa.

(summary in Wink, 2000; Wink and Van Wyk, 2008). Some of these legumes have a famous history as hallucinogenic drugs. Some toxins have immediate effects and thus directly work against herbivores. Others have longer term consequences. These compounds act indirectly by decreasing the longterm survival and reproductive fitness of herbivores.

Upon wounding cyanogenic glucosides release HCN after enzymatic hydrolysis. HCN is a respiratory poison as it blocks the mitochondrial respiratory chain. It is a deadly poison for most animals (reviewed in Seigler, 1998; Wink and Van Wyk, 2008). NPAA are analogues of one of the 20 proteinogenic amino acids. When they are incorporated into proteins, these proteins fold in a different way, leading to inactive or wrongly active proteins. NPAA can be regarded as toxic antimetabolites (Rosenthal, 1982) which affect herbivores, bacteria, fungi and viruses.

Polyphenols (including tannins) can form several hydrogen bonds and even ionic bonds (when their phenolic hydroxyl groups dissociate) with most proteins and even DNA-bases. They thus modulate the activity of many proteins, involving enzymes, ion channels, transporters, transcription factors, motor proteins, and cytoskeletal proteins. As a consequence many polyphenols are pharmacologically active, being among others antioxidant, anti-inflammatory, antibacterial, antifungal, and antiviral (Wink, 2008b).

Furanocoumarins (FC) are lipophilic and can diffuse easily into cells where they intercalate DNA. When activated by UV light, they can form covalent bonds with adjacent pyrimidine bases (such as cytosine or thymine) (Wink and Schimmer, 2010). FC treatment leads to apoptotic cell death. In the liver FC are converted into epoxides which can alkylate DNA. When skin comes into contact with FC, severe inflammation can result (resembling strong sunburn). Thus FC are mutagenic and do not only interact with animal targets, but also with DNA of bacteria, fungi and viruses. Therefore, FC have pronounced antimicrobial properties (Wink and Van Wyk, 2008).

In legumes, most saponins are of the triterpene type, with steroidal saponins being rare. Saponins interfere with biomembranes of most species, where they form complexes with cholesterol. Whereas the lipophilic core of monodesmosidic saponins intercalates biomembranes, their hydrophilic sugar side chains remain outside the cells and interacts with glycoproteins or glycolipids. Saponins form pores in membranes and can even lyse cells at higher concentration (Wink, 2008b). The bitter-tasting saponins are apparently directed against herbivores and microbes, especially fungi and viruses.

In the class of steroidal compounds, cardiac glycosides (CG), which inhibit the Na^+/K^+ ATPase in animals and are thus strongly poisonous (reviewed in Wink and Van Wyk, 2008), are rare in legumes. Only a few taxa of the tribe Loteae, such as *Coronilla* and *Securigera*, produce cardiotonic cardenolides (corotoxigenin, glaucotoxigenin, scorpioside, frugoside, hyrcanoside).

3. Molecular phylogeny and chemotaxonomy

Phytochemists had observed early on that a number of secondary compounds are not widely distributed in the plant kingdom but restricted to smaller related groups, such as families, tribes or even genera. As a consequence the discipline “chemotaxonomy” emerged (summarized in Swain, 1963, 1966; Smith, 1976; Bell et al., 1978; Harborne and Turner, 1984). For legumes, corresponding comprehensive

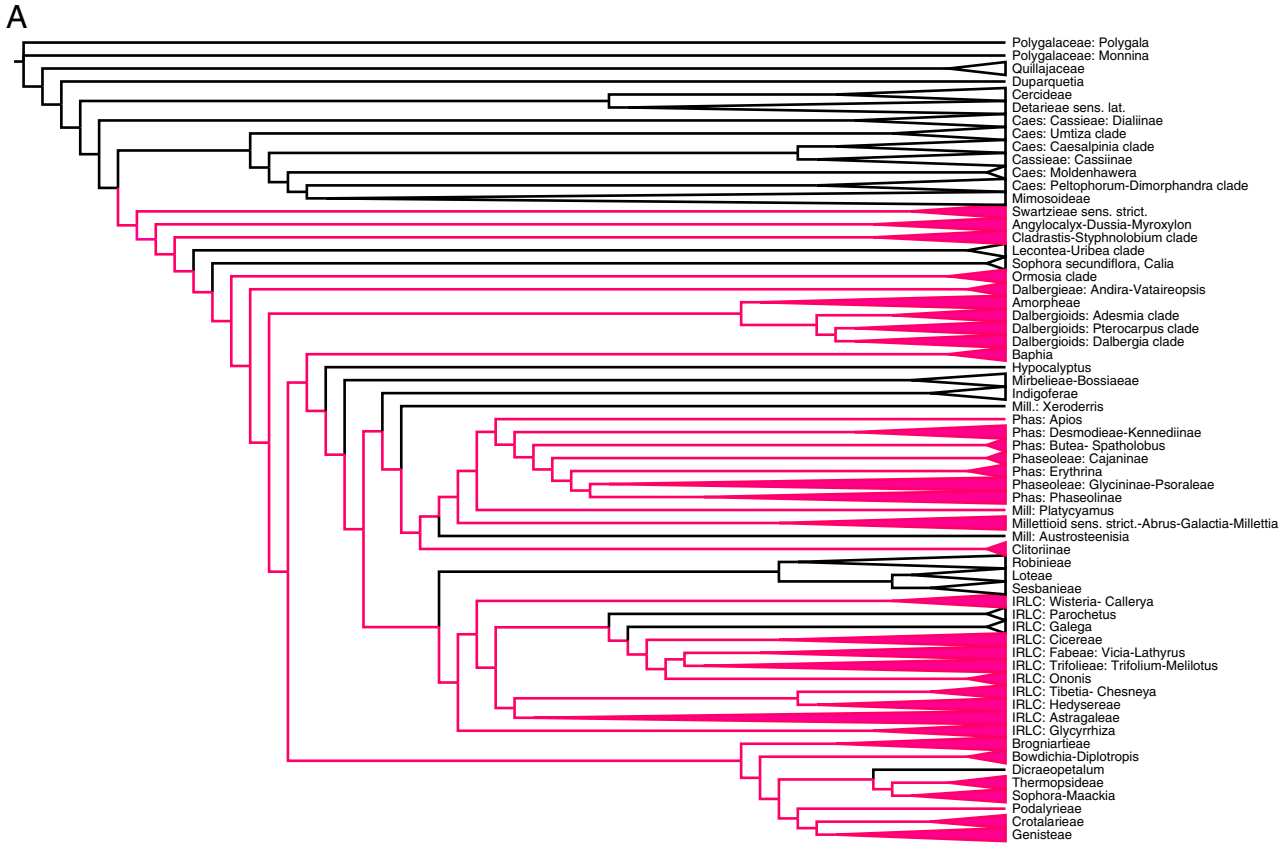

Fig. 3. Distribution of secondary metabolites in legumes. Clades which produce a certain class of secondary metabolites are highlighted in the cladogram. 3A. Isoflavones, B. rotenone (blue) and bufotenine (red), C. coumarins and furanocoumarins, D. cyanogenic glucosides (red) and anthraquinones (blue), E. the NPAA canavanine, F. alkaloids: quinolizidine alkaloids (red), pyrrolizidine alkaloids (green), *Erythrina* alkaloids (blue), physostigmine (yellow), and β-carboline alkaloids (black).

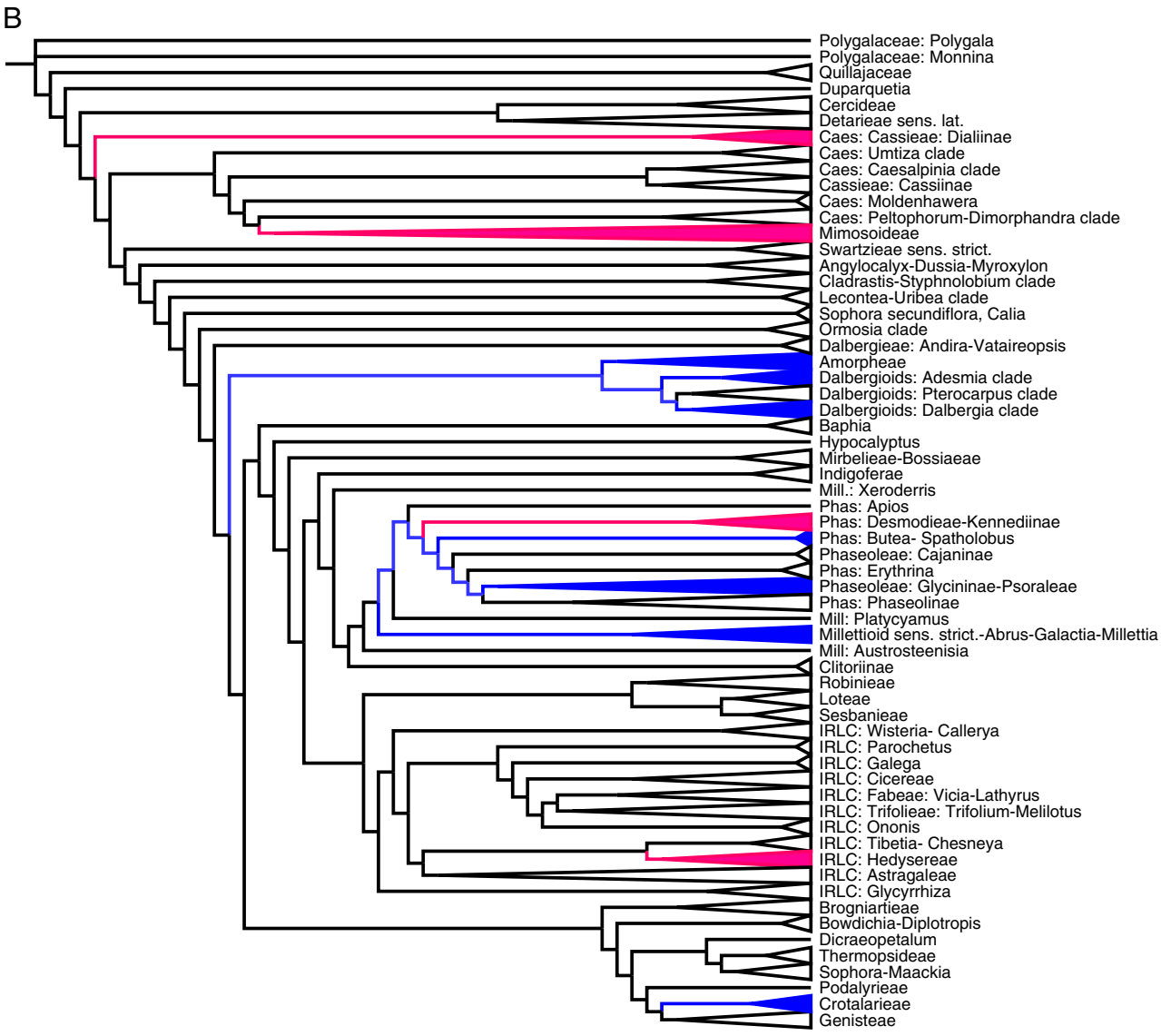

Fig. 3 (*continued*).

treatments include Harborne et al. (1971) and Hegnauer and Hegnauer (1994, 1996, 2001).

One of the basic ideas was that secondary metabolites had no function (some regarded them as waste products) (Hartmann, 2007) and were thus considered as neutral markers which were not subjected to adaptive evolution as known for morphological traits. As we know today, this assumption was wrong; as discussed above (1. Introduction), secondary metabolites have important functions for plants and are important for their ecological fitness. Therefore, secondary metabolites must be considered as adaptive traits and convergent evolution a rule rather than an exception.

The concept of chemotaxonomy proposed to use chemical traits to establish a systematic framework that should reflect phylogeny. It was assumed that all taxa, which produce a secondary metabolite with a limited distribution, should be closely related. Even in the early days of chemotaxonomy, it was observed that a certain secondary metabolite could be found in a particular taxon (such as a genus or tribe) but that not all members of this taxon actually produced it. Since a reliable phylogeny did not exist until 10 to 15 years ago, chemotaxonomists could always place taxa together on account of their common chemical traits and postulate that this would represent their true phylogeny.

This situation changed completely when molecular systematicists started to use nucleotide sequence data from chloroplast and nuclear marker genes (summary in Lewis et al., 2005). It was one of the goals of the 6th International Legume Conference 2013 to reconstruct a new molecular phylogeny of legumes and to redefine subfamily and tribe circumscriptions.

For this paper a molecular phylogeny of more than 1276 species was reconstructed from cpDNA (*rbc*L, matK) and ncDNA (ITS) (provided by M. Wojciechowski). The complete 50% consensus file of a Maximum likelihood analysis was too large to be useful for this evaluation. Therefore, I used the programme FigTree to reduce the tree to a size which would fit on a single page. The programme allows for the collapse of clades with closely related taxa to groups which are in the correct phylogenetic context (Fig. 2). The resulting phylogeny contains a few topologies which are not consistent with a recently published legume phylogeny (LPWG, 2013). These differences are however of no practical consequences for purposes of this paper. The phylogeny shown in Fig. 2 confirms many earlier findings from other molecular studies: Whereas the subfamily (SF) Papilionoideae is monophyletic, the SF Caesalpinioideae appears to be paraphyletic because the SF Mimosoideae is embedded in the Caesalpinioideae and clusters

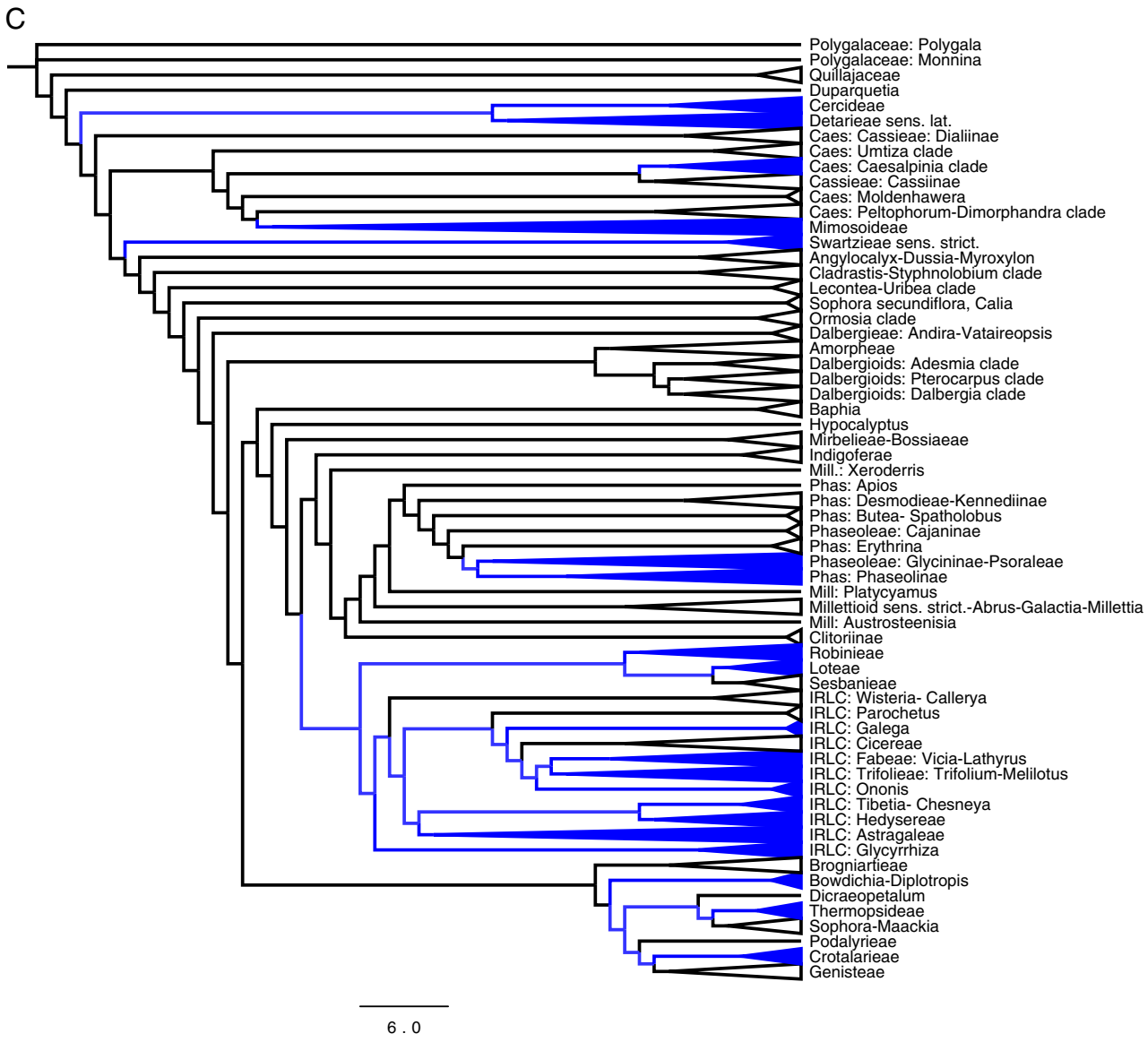

Fig. 3 (*continued*).

as a sister to the *Peltophorum–Dimorphandra* clade. Within the SF Papilionoideae several of the earlier recognized tribes appear to be polyphyletic (e.g. Sophoreae, Dalbergieae) and need to be reorganized into monophyletic taxa. In this publication, the approach of Lewis et al. (2005) is followed by not renaming tribes but by referring to clades (Fig. 2).

When the first phylogenetic trees of Fabaceae became available it was possible to map the distribution of secondary metabolites on the trees (Wink and Waterman, 1999; Wink and Mohamed, 2003; Wink et al., 2010). In this publication, this approach has been repeated by using a better molecular phylogeny and by extending the phytochemical data: (Southon, 1994; ILDIS and Chapman & Hall data base) and several individual publications.

From the class of phenolics, isoflavones and rotenone were selected, because they show a restricted distribution pattern (Fig. 3A,B). Other phenolics, such as flavonoids, simple phenolics or anthocyanins, are widely distributed and present in most taxa. Isoflavones are restricted to members of the SF Papilionoideae. However, a few clades appear not to produce them, such as Robinieae, Loteae, Sesbanieae, Indigofereae and the Mirbelieae–Bossiaeeae clade (Fig. 3A). In other cases, isoflavones might have escaped detection. Rotenone, which represents a toxic isoflavone, only occurs in Papilionoideae, with a predominance in the tribes Amorpheae, some Dalbergioids (except *Pterocarpus* clade), some Phaseoleae and Millettioids (Fig. 3B). Catechins and catechin-derived tannins are abundant in Mimosoideae and Caesalpinioideae; this trait appears to be related to the growth type of legumes in that trees more often produce them than herbs.

Coumarins and furanocoumarins show a patchy distribution (Table 1) with some isolated occurrences in the Mimosoideae and Caesalpinioideae. They are more frequent in the Papilionoideae, especially in the IRLC clade, Loteae, Robinieae, and some Phaseoleae (Fig. 3C). Furanocoumarins are typical for members of the Psoraleae, but not all of them produce them. Anthraquinones have a very limited distribution in Fabaceae, especially in Cassiinae, the *Ormosia* clade and some Millettioids (Fig. 3D).

Terpenoids were not selected for this study, because most of them are widely distributed among legumes, such as triterpenoid saponins, steroids and carotenoids. Only cardenolides are restricted to a few (but not all) members of the genera *Coronilla* and *Securigera* (Table 1).

Among nitrogen-containing secondary metabolites a few classes were analysed in this context, such as cyanogenic glucosides, canavanine

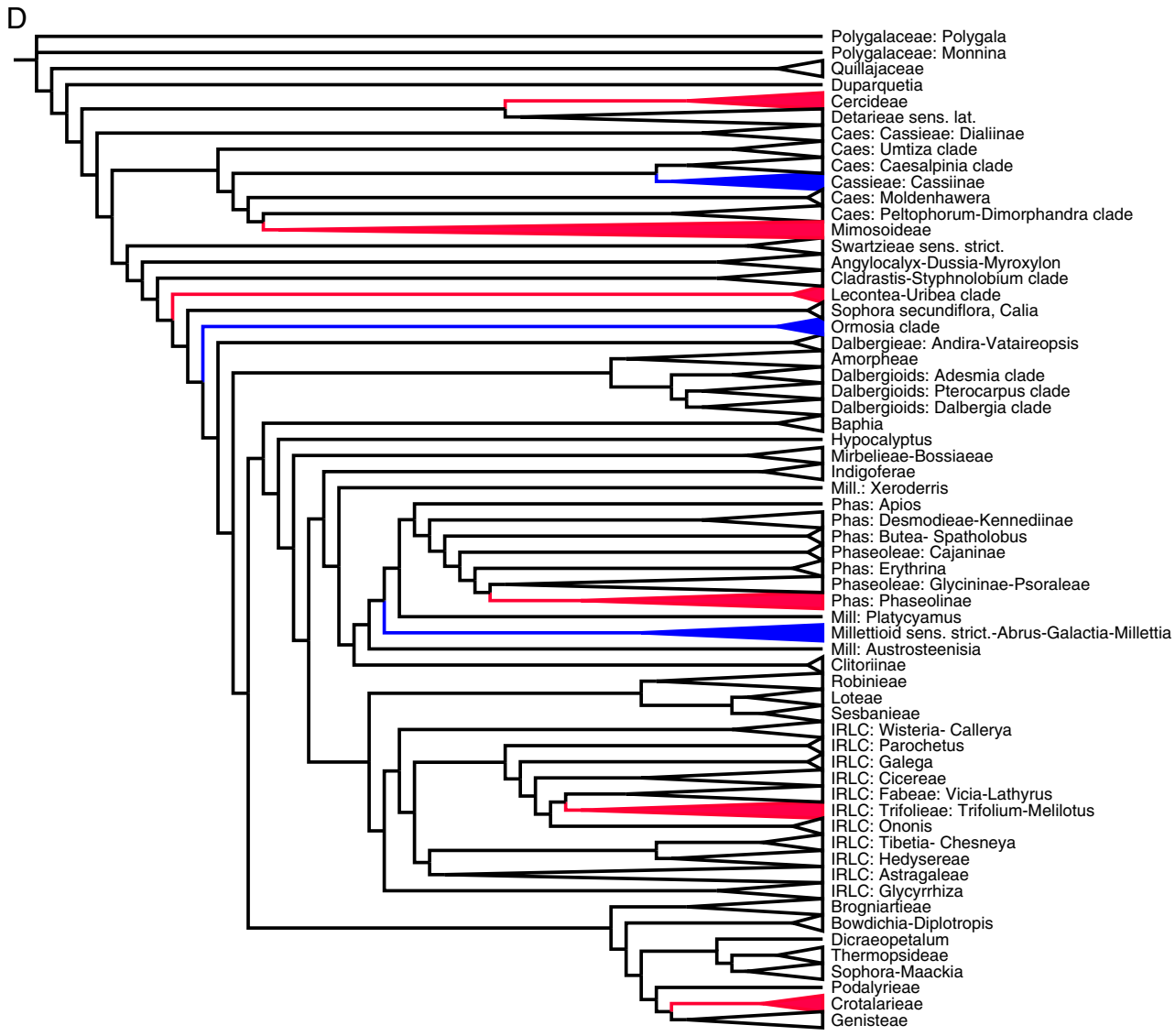

Fig. 3 (*continued*).

(as a representative for NPAA), bufotenin, and alkaloids (QA, PA, β-carbolines, physostigmine and *Erythrina* alkaloids).

Cyanogenic glucosides occur in a restricted number of species of the three subfamilies, which appear largely non-related (Fig. 3D), such as *Acacia* spp., *Holocalyx balansae*, *Lotononis* spp., *Lotus* spp., *Ornithopus* spp., *Trifolium repens*, and *Phaseolus lunatus*.

NPAAs are abundant in tribes which do not sequester alkaloids (Fig. 3F). Canavanine, a NPAA which was intensely studied (Bell et al., 1978) is restricted to the SF Papilionoideae, except the genistoids, dalbergioids, and the Swartzieae–Sophoreae (sens. lat.) complex. Whereas most members of the IRLC clade sequester canavanine, this NPAA has not been found in the Cicereae.

Quinolizidine alkaloids occur in almost all taxa of the genistoid clade, except for *Crotalaria* and *Lotononis* sens. strict., which sequester the biosynthetically unrelated pyrrolizidine alkaloids (Robins, 1993; Hartmann and Witte, 1995) (Fig. 3F). QA are also found in some basal branches of the Papilionoids, such as *Sophora secundiflora*, *Calia*, *Bolusanthus* and the *Ormosia* clade (Fig. 3F) which are distantly related to the genistoids.

β-Carboline alkaloids have been detected in a few species of the Mimosoideae and Caesalpinioideae and in *Desmodium gangeticum* (Table 1; Fig. 3F). The simple indole alkaloid physostigmine is restricted to the genera *Physostigma* and *Dioclea* (Fig. 3F). Erythrina-type alkaloids only occur in the genus *Erythrina*.

Indolizidine alkaloids have been detected in several species of the Astragaleae and in *Castanospermum*. There is evidence, that *Astragalus* and *Oxytropis* harbour an endophytic fungus which is able to produce indolizidine alkaloids such as swainsonine (Ralphs et al., 2008). The piperidine alkaloid ammodendrine often co-occurs with QA in the genistoid clade (Wink, 1993b), whereas 2-piperidine carboxylic acid and related compounds were discovered in all three subfamilies (Table 1). The pyridine alkaloid trigonelline is abundant in members of the IRLC clade, but has also been found in other taxa of the three subfamilies. Among simple amines, the psychoactive tryptamines (such as bufotenin; Fig. 3B) have been detected in a few species of the Mimosoideae, Cassiaeae, Desmodieae and Hedysareae sometimes together with β-carboline alkaloids.

If the distribution data would be analysed strictly cladistically, the resulting cladograms would be largely incongruent with molecular phylogenies. Thus, phytochemical data cannot be used as a direct taxonomic marker in most instances (as discussed in Wink and Waterman, 1999). The same is true for morphological traits which are highly adaptive. However, secondary metabolites nevertheless represent interesting traits which help to understand the evolution

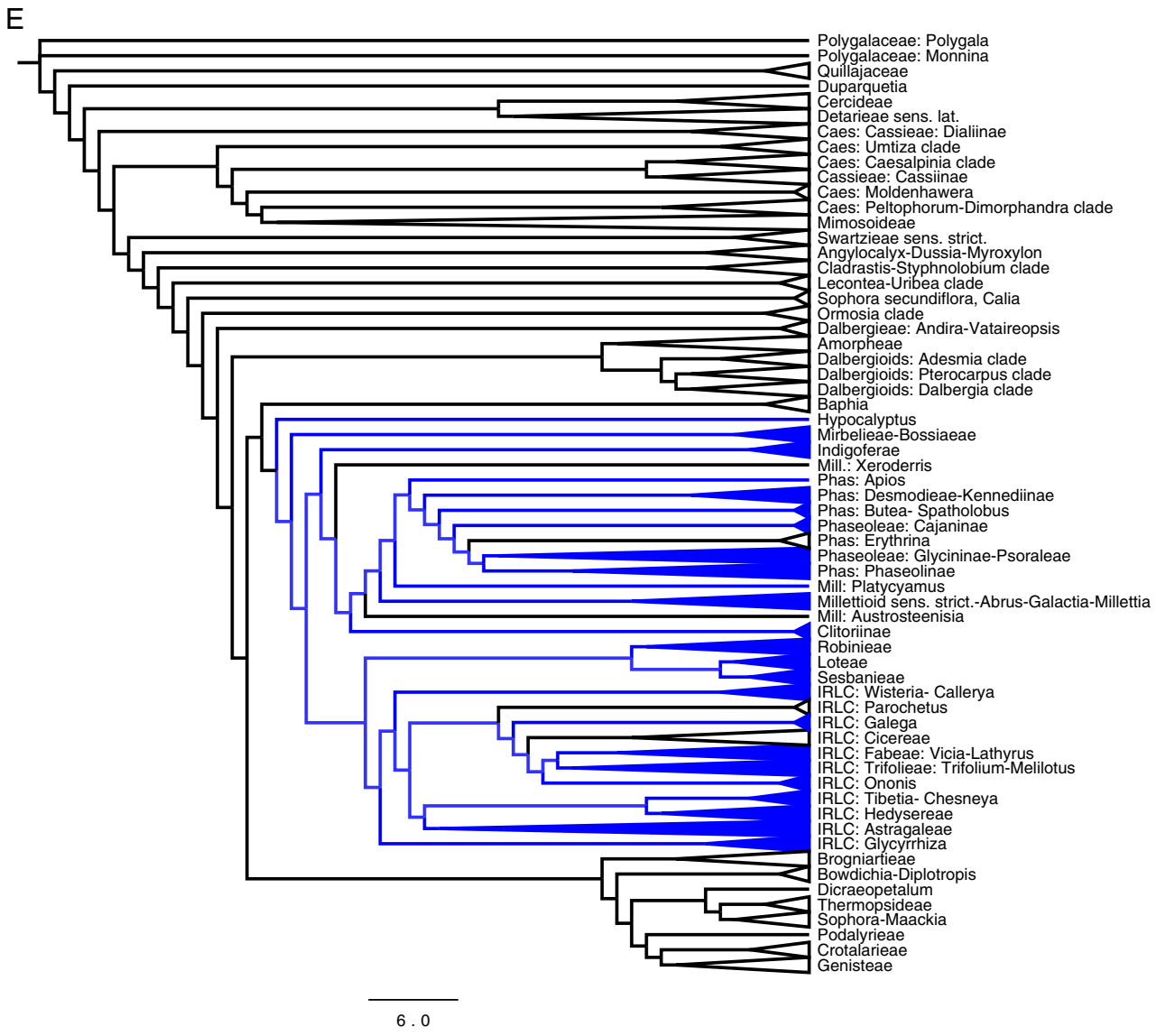

Fig. 3 (*continued*).

of secondary metabolite phenotypes and the adaptation of plants towards a world of dangerous herbivores and microbes.

4. Reasons for the patchy distribution of secondary metabolite

The present evaluation corresponds with the main findings of earlier publications (Wink and Waterman, 1999; Wink and Mohamed, 2003; Wink et al., 2010), indicating that the distribution of several secondary metabolites (Fig. 3) is only partially congruent with the phylogeny of the corresponding groups. The result is a patchy distribution pattern. In this analysis, we have not looked into the distribution of particular secondary metabolites within a genus. As exemplified in earlier publications (Wink and Waterman, 1999; Wink and Mohamed, 2003; Wink et al., 1995, 2010; and Table 1), even within a genus, we often find that some unrelated members produce a certain metabolite and others not. How can we explain such patchy distribution patterns? A few possibilities are outlined in Fig. 4:

1. The phytochemical analysis is far from complete in legumes; therefore, gaps and patchy distribution pattern could reflect missing data. This statement might be true for some compounds and rare legumes but several legume clades have been extensively studied, which makes such an omission less likely as a general explanation.
2. Alternatively, one could assume that the occurrence of particular metabolites in non-related legume taxa is based on convergent evolution, suggesting that the biosynthetic pathways evolved independently and repeatedly in the Fabaceae. The production of cardenolides in a few members of *Coronilla* and *Securigera* might represent such a convergent trait, because CG occurs island-like in many unrelated plant families (Wink, 2003).
3. A convergent trait could also be due to endophytic fungi, which produce a number of secondary metabolites on their own (review in Wink, 2008a). As shown for indolizidine alkaloids, the patchy distribution of these alkaloids might actually depend on such an infection (Ralphs et al., 2008). Thus horizontal gene transfer could be another source for distributional diversity of secondary metabolites in plants.
4. QA are a typical trait of the genistoids suggesting that their ancestors already had the genes for QA synthesis and QA storage which became distributed during phylogeny among all members. As mentioned before, members of the genera *Crotalaria* and *Lotononis* produce PA instead of QA (Fig. 3F). As the Crotalarieae are deeply embedded within the genistoids (Fig. 2), their founders must have obtained the QA genes from their ancestors. I suggest that the

Fig. 3 (*continued*).

QA genes were either permanently inactivated in the Crotalarieae or were just turned off. As seen in Fig. 3F some of the early branches of Papilionoids already produce QA, suggesting that the corresponding genes could have been present early on, but that they became inactivated in many papilionoid tribes which produce NPAA instead (Fig. 3E).

5. PA derive from a completely different biosynthetic pathway (reviewed in Robins, 1993; Hartmann and Witte, 1995); the PA pathway either evolved de novo in the Crotalarieae or derived from an early ancestor common for both the Crotalarieae and the Asteraceae. The complex PA senecionine occurs with an identical stereochemistry in *Senecio* and other Asteraceae. The Asteraceae belong to the asterids whereas legumes are rosids according to APG3. Similar to the situation of QA genes in genistoids, one could imagine that the early ancestors of core dicots already had evolved the genes for the PA pathway and that it was turned on in restricted places only (Fig. 4). Thus, PA occurrence would be rather a matter of gene regulation.

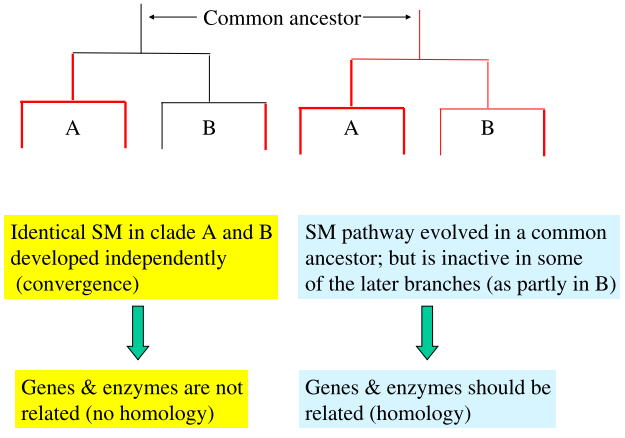

Fig. 4. Scheme to explain the patchy distribution of secondary metabolites.

As shown in Fig. 4, we thus face the alternative of convergent evolution versus gene regulation and inheritance by descendants not only in the PA/QA example, but in all groups illustrated in Fig. 3. How can this problem be solved? In the days of genomics more and more genomes become available for comparison. As discussed in Wink et al. (2010) there is evidence that the genes which encode key enzymes of biosynthesis of flavonoids, indole and isoquinoline alkaloids are present not only in taxa which actually produce such compounds, but in most plant taxa (e.g. even in *Arabidopsis thaliana*

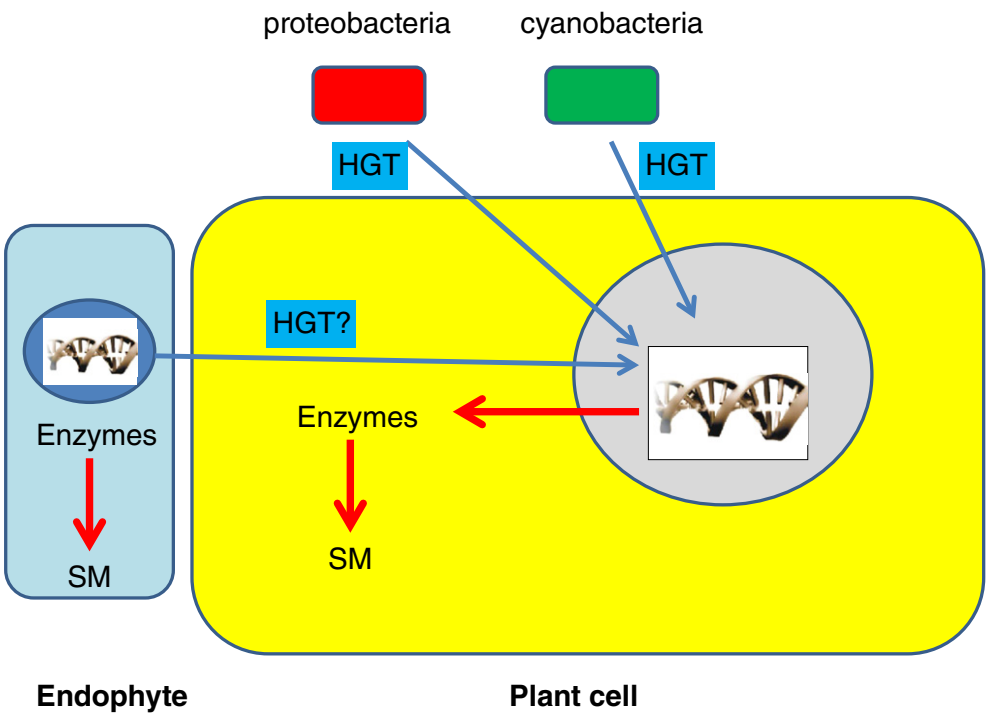

Fig. 5. Model for the evolution of plant secondary metabolism.

which does not make alkaloids; Facchini et al., 2004). In several cases related genes/proteins could be discovered in bacteria and fungi, suggesting that these genes had evolved much earlier in evolution (Wink, 2003). The genes might have found their way into plant genomes by distant horizontal gene transfer (HGT) from endosymbiotic bacteria from which mitochondria and chloroplasts had derived (Wink, 2008a,b; Wink et al., 2010). Horizontal gene transfer could also have taken place in case of endophytic and viral infections (Fig. 5). When land plants evolved about 400 million years ago, they had to deal with herbivores and microbes. It is likely that the early plants used terpenoids and phenolics for defence. When angiosperms which attract pollinating and seed dispersing animals evolved in the Cretaceous, more powerful anti-herbivore defences were needed. The dominance of alkaloids and other nitrogen-containing secondary metabolites in angiosperms must be regarded in this context (Wink, 2003, 2008a; Wink et al., 2010). As a consequence, secondary metabolism appears to be an early rather than a recent innovation of plants.

For future work, we need more data on the genes responsible for the biosynthesis and storage of several secondary metabolites in legumes discussed in Fig. 4 before deciding on the issue whether convergent evolution or phylogenetic transmission is the underlying mechanism. Although there is evidence that plant genomes contain hundreds of genes for the biosynthesis of secondary metabolites, there is as yet no information on whether all these genes are still functional or not. One way to solve this problem would be to clone and express the genes, e.g. from *Arabidopsis*, and analyse whether specific alkaloids could be synthesized by recombinant enzymes.

Acknowledgements

A tree file for the Fabaceae, which was used to reconstruct the trees used in this paper, was kindly supplied by Martin F. Wojciechowski (Arizona State University; USA).

References

Bell, E.A., Charlwood, B.V., 1980. Secondary Plant Products. Springer, Heidelberg.
Bell, E.A., Lackey, J.A., Polhill, R.M., 1978. Systematic significance of canavanine in the Papilionoideae. Biochemical Systematics and Ecology 6, 201–212.
Brown, K., Trigo, J.R., 1995. The ecological activity of alkaloids. In: Cordell, G.A. (Ed.), The Alkaloids, vol. 47, pp. 227–354.
Cipollini, M.L., Levey, D.J., 1997. Secondary metabolites of fleshy vertebrate-dispersed fruits: adaptive hypotheses and implications for seed dispersal. American Naturalist 150, 346–373.
Conn, E.E., 1981. Secondary plant products. The Biochemistry of Plants, vol. 7. Academic Press, New York.
Dewick, P.M., 2002. Medicinal natural products, A Biosynthetic Approach, 2nd ed. Wiley, New York.
DNP, 1996. Dictionary of Natural Products. CD-ROM Version 5:1. Chapman and Hall, London.
Facchini, P.J., Bird, D.A., St-Pierre, B., 2004. Can *Arabidopsis* make complex alkaloids? Trends in Plant Science 9, 116–122.
Fraenkel, G., 1959. The raison d'être of secondary substances. Science 129, 1466–1470.
Harborne, J.B., 1993. Introduction to Ecological Biochemistry, 4th ed. Academic Press, London.
Harborne, J.B., Turner, B.L., 1984. Plant Chemosystematics. Academic Press, London.
Harborne, J.B., Boulter, D., Turner, B.L., 1971. Chemotaxonomy of the Leguminosae. Academic Press, London.
Hartmann, T., 2007. From waste products to ecochemicals: fifty years research of plant secondary metabolism. Phytochemistry 68, 2831–2846.
Hartmann, T., Witte, L., 1995. Chemistry, biology and chemoecology of the pyrrolizidine alkaloids. In: Pellitier, S.W. (Ed.), Alkaloids: Chemical and Biological Perspectives, vol. 9. Pergamon, Oxford, pp. 155–233.
Hegnauer, R., Hegnauer, M., 1994. Chemotaxonomie der Pflanzen, vol. 11a. Birkhäuser, Basle.
Hegnauer, R., Hegnauer, M., 1996. Chemotaxonomie der Pflanzen, vol. 11b–1. Birkhäuser, Basle.
Hegnauer, R., Hegnauer, M., 2001. Chemotaxonomie der Pflanzen, vol. 11b–2. Birkhäuser, Basle.
Kinghorn, A.D., Balandrin, M.F., 1984. Quinolizidine alkaloids of the Leguminosae: structural types, analysis, chemotaxonomy and biological activities. In: Pelletier, W.S. (Ed.), Alkaloids: Chemical and Biological Perspectives. Wiley, New York, pp. 105–148.
Levin, D.A., 1976. The chemical defences of plants to pathogens and herbivores. Annual Review of Ecology and Systematics 7, 121–159.
Lewis, G., Schrire, B., Mackinder, B., Lock, M., 2005. Legumes of the World. The Royal Botanical Gardens, Kew.
LPWG, 2013. Global legume diversity assessment: concepts, key indicators, and strategies. Taxon 62, 249–266.
McLean, E., 1970. The toxic actions of pyrrolizidine (*Senecio*) alkaloids. Pharmacology Reviews 22, 430–483.
Ralphs, M.H., Creamer, R., Baucom, D., Gardner, D.R., Welsh, S.L., Graham, J.D., Hart, C., Cook, D., Stegelmeier, B.L., 2008. Relationship between the endophyte *Embellisia* spp. and the toxic alkaloid swainsonine in major locoweed species (*Astragalus* and *Oxytropis*). Journal of Chemical Ecology 34, 32–38.
Roberts, M.F., Wink, M., 1998. Alkaloids: Biochemistry, Ecological Functions and Medical Applications. Plenum, New York.

Robins, D.J., 1993. Pyrrolizidine alkaloids. In: Waterman, P.G. (Ed.), Methods in Plant Biochemistry, vol. 8. Academic Press, London, pp. 175–195.
Rosenthal, G.A., 1982. Plant Nonprotein Amino and Imino Acids. Academic Press, New York.
Rosenthal, G.A., Berenbaum, M.R., 1991. Herbivores: their interactions with secondary plant metabolites. The Chemical Participants, vol. 1. Academic Press, San Diego.
Seigler, D.S., 1998. Plant Secondary Metabolism. Kluwer, Norwell.
Smith, P.M., 1976. The Chemotaxonomy of Plants. E. Arnold Ltd., London.
Southon, I.W., 1994. Phytochemical Dictionary of the Leguminosae. Chapman & Hall, London.
Sprent, J.I., McKey, D., 1994. Advances in Legume Systematics, Part 5, The Nitrogen Factor. The Royal Botanical Gardens, Kew.
Swain, T., 1963. Chemical Plant Taxonomy. Academic Press, London.
Swain, T., 1966. Comparative Phytochemistry. Academic Press, London.
Swain, T., 1977. Secondary compounds as protective agents. Annual Review of Plant Physiology 28, 479–501.
Teuscher, E., Lindequist, U., 2010. Biogene Gifte. Biologie, Chemie, Pharmakologie, Toxikologie. Wissenschaftliche Verlagsgesellschaft, Stuttgart.
Veitch, N.C., 2010. Flavonoid chemistry of the Leguminosae. In: Santos-Buelga, C., Escribano-Baillon, M.T., Lattanzio, V. (Eds.), Recent Advances in Polyphenol Research. Wiley-Blackwell, Chichester, pp. 23–58.
Wink, M., 1988. Plant breeding: importance of plant secondary metabolites for protection against pathogens and herbivores. Theoretical and Applied Genetics 75, 225–233.
Wink, M., 1992. The role of quinolizidine alkaloids in plant insect interactions. In: Bernays, E.A. (Ed.), Insect–Plant Interactions, vol. IV. CRC Press, Boca Raton, pp. 133–169.
Wink, M., 1993a. Allelochemical properties and the raison d'être of alkaloids. In: Cordell, G. (Ed.), The Alkaloids, vol. 43. Academic Press, Orlando, pp. 1–118.
Wink, M., 1993b. Quinolizidine alkaloids. In: Waterman, P.G. (Ed.), Methods in Plant Biochemistry. Academic Press, London, pp. 197–239.
Wink, M., 1993c. The plant vacuole: a multifunctional compartment. Journal of Experimental Botany 44, 231–246.
Wink, M., 1997. Compartmentation of secondary metabolites and xenobiotics in plant vacuoles. Advances in Botanical Research 25, 141–169.
Wink, M., 2000. Interference of alkaloids with neuroreceptors and ion channels. In: Atta-Ur-Rahman (Ed.), Bioactive Natural Products. Elsevier, New York, pp. 1–127.
Wink, M., 2003. Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. Phytochemistry 64, 3–19.
Wink, M., 2007. Molecular modes of action of cytotoxic alkaloids—from DNA intercalation, spindle poisoning, topoisomerase inhibition to apoptosis and multiple drug resistance. In: Cordell, G. (Ed.), The Alkaloids, vol. 64. Academic press, San Diego, pp. 1–48.
Wink, M., 2008a. Plant secondary metabolism: diversity, function and its evolution. Natural Products Communications 3, 1205–1216.
Wink, M., 2008b. Evolutionary advantage and molecular modes of action of multi-component mixtures used in phytomedicine. Current Drug Metabolism 9, 996–1009.
Wink, M., 2010a. Functions and biotechnology of plant secondary metabolites. Annual Plant Reviews, vol. 39. Wiley-Blackwell, Oxford.
Wink, M., 2010b. Biochemistry of plant secondary metabolism. Annual Plant Reviews, vol. 40. Wiley-Blackwell, Oxford.
Wink, M., Mohamed, G.I.A., 2003. Evolution of chemical defense traits in the Leguminosae: mapping of distribution patterns of secondary metabolites on a molecular phylogeny inferred from nucleotide sequences of the rbcL gene. Biochemical Systematics and Ecology 31, 897–917.
Wink, M., Schimmer, O., 2010. Molecular modes of action of defensive secondary metabolites. In: Wink, M. (Ed.), Functions and Biotechnology of plant secondary metabolites. Annual Plant Reviews, vol. 39. Wiley-Blackwell, Oxford, pp. 21–161.
Wink, M., Van Wyk, B.-E., 2008. Mind-altering and Poisonous Plants of the World. Briza Publications, Pretoria.
Wink, M., Waterman, P., 1999. Chemotaxonomy in relation to molecular phylogeny of plants. In: Wink, M. (Ed.), Biochemistry of Plant Secondary Metabolism. Annual Plant Reviews, vol. 2. Sheffield Academic Press, pp. 300–341.
Wink, M., Witte, L., 1984. Turnover and transport of quinolizidine alkaloids: diurnal variation of lupanine in the phloem sap, leaves and fruits of *Lupinus albus* L. Planta 161, 519–524.
Wink, M., Meiβner, C., Witte, L., 1995. Patterns of quinolizidine alkaloids in 56 species of the genus *Lupinus*. Phytochemistry 38, 139–153.
Wink, M., Schmeller, T., Latz-Brüning, B., 1998. Modes of action of allelochemical alkaloids: interaction with neuroreceptors, DNA and other molecular targets. Journal of Chemical Ecology 24, 1881–1937.
Wink, M., Botschen, F., Gosmann, C., Schäfer, H., Waterman, P.G., 2010. Chemotaxonomy seen from a phylogenetic perspective and evolution of secondary metabolism. In: Wink, M. (Ed.), Biochemistry of Plant Secondary Metabolism. Annual Plant Reviews, vol. 40. Wiley-Blackwell, Oxford.