BOT 6516 Plant Metabolism

Lecture 22

Natural Products

Slide sets available at:

http://hort.ifas.ufl.edu/teach/guyweb/bot6516/index.html
Why is natural product and secondary metabolism so widely variable?

Plants must evolve and exploit metabolic systems to create complex natural products that have adaptive benefit for survival.

Many of natural products provide adaptive advantage to the plant as these compounds can act as chemical cues for plants during their on-going interactions with physical and biotic factors in their environments.

In addition, plant natural products have the unintended consequences of positive and negative impacts on human and animal health and nutrition.
The complex structures of many natural products can require 20–30 enzymes and steps for their biosynthesis. Additionally, each biosynthetic pathway can be branched into a metabolic network that leads to a multiplicity of structurally related chemical structures.

Natural product biosynthesis and storage of plant compounds such as phenylpropanoids, terpenoids and alkaloids are complex processes that involve multiple subcellular compartments and specific cell types.

The question of how reactivity, regio-chemistry and stereo-chemistry are combined in multi-step conversion of substrates into products looms large for natural products.

How does the catalytic landscape change as an enzyme family acquires new substrate and/or product specificities?

Structural biology provides a means to uncover information about the structure–function relationships of metabolic enzymes at the molecular and atomic levels.

When combined with genomic and biochemical approaches, structural biology can provide an appreciation of molecular evolution and structural enzymology that have promise to help us better understand a wide number of questions relating to specialized secondary metabolism.
Schematic of natural product pathways. The pathways are simplified, and do not show all intermediate or end products. Pink panel: phenylpropanoid biosynthesis. Yellow panel: biosynthesis of oxygenated carotenoids (astaxanthin) or monoterpenes. Dark blue panel: formation of glycinebetaine. Light blue panel: biosynthesis of pyridine (nicotine) or benzopanthenanthridine alkaloids. Tan panel: biosynthesis of cyanogenic glucosides and glucosinolates. Enzymes shown are: ANR, anthocyanidin reductase; ANS, anthocyanidin synthase; BBE, berberine bridge enzyme; C4H, cinnamate 4-hydroxylase; CHI, chalcone isomerase; CHS, chalcone synthase; CYP, cytochrome P450; DFR, dihydroflavonol reductase; F3H, flavanone 3-hydroxylase; F3′H, flavonoid 3′-hydroxylase; FLS, flavonol synthase; HQT; hydroxycinnamoyl CoA quinate hydroxycinnamoyl transferase; IFS, isoflavone synthase; LH, limonene hydroxylase; γ-TS, γ-terpinene synthase; LS, limonene synthase; NMCH, N-methylcoclaurine 3′-hydroxylase; PAL, L-phenylalanine ammonia-lyase; PENMT, phosphoethanolamine N-methyltransferase; β-PS, β-pinene synthase; PMT, putrescine N-methyltransferase; TR, tropinone reductase; UGT, uridine diphosphate glucosyltransferase; 3,3′-β-H, carotenoid 3,3′-β-hydroxylase; 4CL, 4-coumarate CoA ligase; 4,4′-β-O, carotenoid 4,4′-β-oxygenase. From Dixon, 2005.
Will it be possible one day for plant biologist to rationally engineer biosynthesis pathways to create novel “natural” products for human use?

A fundamental weakness is the present lack of understanding of the full metabolic potential of any plant species, even Arabidopsis.

Genome sequencing is revealing real and potential metabolic pathways that will require experimental confirmation. For known metabolic pathways, multi-gene families for many enzymes make it difficult to know which gene(s) to target in metabolic engineering;

Points of flux control are virtually unknown for most plant natural product pathways; this question has mostly been addressed by trial-and-error transgenic modulation of predicted ‘rate-limiting’ enzymes.

Much can be learned using metabolite and transcript profiling of transgenic plants, the organization of pathways or sub-pathways in metabolic channels may result in unexpected patterns of metabolite accumulation and pathways.

Failures’ in pathway engineering may reflect a lack of understanding of transport and accumulation processes rather than of biosynthesis per se.
Clearly engineering natural product pathways for plant improvement is limited by a lack of a detailed understanding of the biochemistry of the biosynthetic pathways, and by the need for coordinate and integrated regulation of multiple gene activities.

New approaches are facilitating both the discovery of genes that encode natural products and pathway engineering. Notable successes have been reported in altering complex pathways to improve plant quality and resistance to biotic and abiotic stresses.

Predictive metabolic engineering describes the manipulation of a biosynthetic pathway based on the application of systems biology approaches (i.e. integrated metabolomics, proteomics and transcriptomics).

It requires reiterative testing of models until an approximation of the true regulatory network is described.

New informatics tools and databases of metabolic pathways and enzymes and new approaches that integrate analytical and informatic information for comparing gene, enzyme and metabolite levels in metabolically perturbed systems, are helping aid gene discovery and to provide a more integrated understanding of pathway regulation.
Metabolic engineering of plant secondary metabolism checklist. The plant image shows an example of secondary metabolite pathway engineering that did not yield the expected outcome. Ectopic co-expression in Arabidopsis of three transcription factors for proanthocyanidin led to accumulation of anthocyanins and proanthocyanidins but also caused severe stunting and early death of the plants. From Dixon 2005.
In the case of phenylpropanoids, a number of the enzymes in these multiple-branched biosynthetic networks belong to large super-families of biosynthetic enzymes that catalyze a core set of chemical transformations that extend biosynthetic capabilities beyond the basic phenylpropanoid (C6–C3) building block.

Recall that three enzymes are required to transform phenylalanine into the Coenzyme A (CoA)-activated hydroxycinnamoyl ester. The deamination of by PAL produces cinnamic acid, which then serves as the precursor for all phenylpropanoid secondary metabolism.

Cinnamic acid 4-hydroxylase (C4H) catalyzes the addition of a hydroxyl group at the para position of the phenyl ring of cinnamic acid, producing coumaric acid. The carboxyl group of coumaric acid is then activated by the formation of a thioester bond with CoA, a process catalyzed by hydroxycinnamate CoA ligase (4CL).

Some grasses and fungi have a dual-specificity ammonia lyase (PAL/tyrosine ammonia lyase [TAL]) that can use tyrosine as a substrate, reducing the number of enzymes that are essential for the production of p-coumaroyl-CoA from three in the general phenylpropanoid pathway to two.
Chalcone synthase, stilbene synthase belong to an expanding family of plant and microbial Polyketide Synthases, which give rise to chemical diversity and ultimately physiological and ecological diversity in their host organisms.

Recent evidence is emerging that consecutive enzymes of phenylpropanoid and flavonoid biosynthesis are organized into macromolecular complexes associated in some cases with endomembranes.

FRET, a non-invasive procedure for monitoring the protein–protein interactions of protein molecules in vivo have indicated a weak protein–protein interaction between PAL and C4H at the surface of the endomembrane suggesting that metabolic channeling is a characteristic of the early phenylpropanoid pathway.

Further, the interaction of chalcone synthase (CHS), chalcone isomerase (CHI) and dihydroflavonol 4-reductase (DFR) in Arabidopsis has been demonstrated using the yeast two-hybrid system. In addition, the interaction appears to be directional: as CHS interacts with DFR, CHI with CHS and DFR with CHI. Recent evidence suggests that flavonoid enzymes may form a globular complex rather than a sequential linear array.
(a) Localization of PAL– and C4H–eGFP fusions in leaf epidermis cells of tobacco. (1) Free eGFP fluorescing throughout the cytoplasm and nucleus. (2) eGFP–HDEL (harboring a carboxy-terminal ER-retention signal) fluorescing in a reticulate pattern indicative of ER. (3) C4H–MA–eGFP (C4H membrane anchor [MA] fused to eGFP) showing the same reticulate pattern of fluorescence as eGFP–HDEL. (4) PAL1–eGFP showing reticulate and cytoplasmic fluorescence. (5) PAL2–eGFP showing cytoplasmic fluorescence.

(b) Immunolocalization of CHS and CHI in Arabidopsis root cells by confocal microscopy. Whole-mount seedlings double-labeled with anti-CHS and anti-CHI antibodies: CHI, red fluorescence; CHS, green fluorescence; nuclei pseudocolored blue; co-localization of CHS and CHI is indicated by yellow color resulting from the merged red and green images. The arrowhead points towards a root epidermal cell containing CHS and CHI. The arrow points towards the co-localization of CHS and CHI at the apical end of a cortex cell. From Kutchan 2005.
Monoterpenoid biosynthetic enzymes are exclusively localized to highly specialized glandular trichome secretory cells.

Monoterpenoid indole- and morphinan alkaloids require a combination of a number of cell types for synthesis and storage; phloem parenchyma, laticifers and epidermal cells.

Intra- and intercellular translocation must be considered to fully understand how plants can regulate the formation and accumulation of complex natural products.
Monoterpene biosynthesis in plants.

(a) Schematic of a peltate glandular trichome of peppermint. Monoterpene biosynthetic enzymes have been biochemically and immunocytochemically localized to the secretory cells of the gland.

(b) The monoterpene biosynthetic pathway in peppermint and spearmint. The enzymes indicated are the small-subunit (SSU) and large subunit (LSU) of geranyl diphosphate synthase (GPSS); (−)-limonene synthase (LS); (−)-limonene 3-hydroxylase (peppermint) (L3OH); (−)-limonene 6-hydroxylase (spearmint) (L3OH); (−)-trans-isopiperitenol dehydrogenase (peppermint) (IPD); (−)-trans-carveol dehydrogenase (CD); and (+)-pulegone reductase (PR). The localization of the enzymes of monoterpene biosynthesis in mint species, as revealed by immunogold cytochemistry is indicated by black spots. LP, leucoplast; M, mitochondrion; N, nucleus. From Kutchan, 2005.
Schematic of the biosynthetic pathway that leads to monoterpenoid indole alkaloids in Madagascar periwinkle.

The three non-mevalonate (MEP) pathway genes DXS, DXR and MECS, and G10H are expressed in internal phloem parenchyma cells. Further along the biosynthetic pathway, the TDC, SLS and STR genes are expressed in epidermis and are directly involved in the formation of the central monoterpenoid indole alkaloid biosynthetic intermediate 3α(S)-strictosidine. The D4H and DAT genes encode the final enzymes of vindoline biosynthesis and are expressed in laticifers and idioblasts. Open arrowheads, idioblasts; closed arrowheads, laticifers. From Kutchan 2005.


(Commentary from previous slides were derived from the three references above as coded by the color.)
Some Useful Metabolism Website Tools

http://www.arabidopsis.org/biocyc/index.jsp
Pathway Tools Query Page

This form provides several different mechanisms for querying Pathway/Genome Databases.

**Select a dataset:** Arabidopsis thaliana.COL

- **Query**
  
  To retrieve objects by name, first select the type of object you wish to retrieve, then enter the name of the object and click Submit. All objects containing that name as a substring will be returned. You may also enter multiple names or EC numbers, separating them with commas.

- **Browse Ontology:**
  
  Each dataset contains classification hierarchies for pathways, for reactions (the enzyme nomenclature system), for compounds, and for genes. Select a classification system to browse.

- **Choose from a list of all**

- **Links to summary information about the selected organism:**
  
  - Summary page for dataset
  - Cellular Overview Diagram/Omics Viewer (not available for MetaCyc)
  - History of updates to this dataset
  - PathoLogic Pathway Analysis (not available for E. coli or MetaCyc)
  
- **Comparative Analysis**
The query "Pathways" matched 281 objects:

- α-amyrin biosynthesis
- β-alanine biosynthesis
- β-alanine biosynthesis
- β-alanine biosynthesis
- (deoxy)ribose phosphate degradation
- 13-LOX and 13-HPL pathway
- cis-zeatin biosynthesis
- de novo biosynthesis of purine nucleotides
- de novo biosynthesis of purine nucleotides
- ent-kaurene biosynthesis
- trans-zeatin biosynthesis
- abscisic acid biosynthesis
- abscisic acid glucose ester biosynthesis
- acetyl-CoA biosynthesis (from citrate)
- acetyl-CoA biosynthesis (from pyruvate)
- acrylonitrile degradation
- acyl-ACP desaturation pathway
- acyl-ACP thioesterase pathway
- acyl-CoA synthetase pathway
- acyl-CoA thioesterase pathway
- adenosylmethionine biosynthesis
- aerobic respiration
- ajugose biosynthesis (galactinol-dependent)
- alanine biosynthesis
- alanine degradation
- aldoxime degradation
- ammonia assimilation cycle
KEGG: Kyoto Encyclopedia of Genes and Genomes

A grand challenge in the post-genomic era is a complete computer representation of the cell, the organism, and the biosphere, which will enable computational prediction of higher-level complexity of cellular processes and organism behaviors from genomic and molecular information. Towards this end, we have been developing a bioinformatics resource named KEGG as part of the research projects of the Kanehisa Laboratories in the Bioinformatics Center of Kyoto University and the Human Genome Center of the University of Tokyo.

Main entry points to the KEGG web service

<table>
<thead>
<tr>
<th>KEKG2</th>
<th>KEGG Table of Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>PATHWAY</td>
<td>KEGG pathway maps for biological processes</td>
</tr>
<tr>
<td>BRITE</td>
<td>Functional hierarchies of biological systems</td>
</tr>
<tr>
<td>GENES</td>
<td>Gene catalogs and ortholog relations in complete genomes</td>
</tr>
<tr>
<td>LIGAND</td>
<td>Chemical compounds, drugs, glycans, and reactions</td>
</tr>
</tbody>
</table>
# Arabidopsis thaliana (thale cress)

<table>
<thead>
<tr>
<th>Genome info</th>
<th>Pathway maps</th>
<th>Gene catalogs</th>
<th>Genome map</th>
<th>Organism list</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Organism</strong></td>
<td>ath</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Name</strong></td>
<td>A.thaliana, ARATH, 3702</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Full name</strong></td>
<td>Arabidopsis thaliana (thale cress)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Definition</strong></td>
<td>Arabidopsis thaliana (thale cress)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Taxonomy</strong></td>
<td>TAX: 3702</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Lineage</strong></td>
<td>Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data source</strong></td>
<td>RefSeq</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Original DB</strong></td>
<td>TIGR MIPS Kazusa TAIR</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| **Statistics** | Number of protein genes: 26735  
Number of RNA genes: 68 |               |            |               |
| **Reference** | PMID: 11130712 (chromosome 1)  
Theologis A, et al.  
Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana.  
| **Authors**   |                |               |            |               |
| **Reference** | PMID: 10617197 (chromosome 2)  
Lin X, et al.  
Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
### KEGG Pathway maps

#### 01110 Carbohydrate Metabolism

- 00110 Glycolysis / Gluconeogenesis
- 00210 Citrate cycle (TCA cycle)
- 00310 Pentose phosphate pathway
- 00410 Pentose and glucuronate interconversions
- 00510 Fructose and mannose metabolism
- 00520 Galactose metabolism
- 00530 Ascorbate and aldarate metabolism
- 00540 Starch and sucrose metabolism
- 00550 Aminosugars metabolism
- 00560 Nucleotide sugars metabolism
- 00620 Pyruvate metabolism
- 00630 Glyoxylate and dicarboxylate metabolism
- 00640 Propanoate metabolism
- 00650 Butanoate metabolism
- 00670 Inositol metabolism
- 00680 Inositol phosphate metabolism

#### 01120 Energy Metabolism

- 00120 Oxidative phosphorylation
- 00150 Photosynthesis
- 00160 Photosynthesis - antenna proteins
- 00210 Carbon fixation
- 00220 Reductive carboxylate cycle (CO2 fixation)
- 00680 Methane metabolism
- 00910 Nitrogen metabolism
- 00920 Sulfur metabolism
That’s all folks!