

Biosynthesis

6

OM6.1 ACETYL-CoA CARBOXYLASES DIFFER BETWEEN PLASTIDS AND CYTOSOL, AND BETWEEN EUDICOTS AND MONOCOTS

Plants have plastidial and cytosolic forms of acetyl-CoA carboxylase. In eudicots and monocots other than grasses, the plastidial enzyme is a complex of four different proteins (**Figure OM6.1-1**), three encoded in the nuclear genome and one in the plastid genome. It is similar in structure to bacterial forms of the enzyme, reflecting the evolutionary origin of plastids (see Section 1.2). The cytosolic form of acetyl-CoA carboxylase consists of a single protein with four different regions (domains) that catalyze the separate reactions carried out by the four proteins in the plastidial complex. This type of monomeric enzyme is also found in animals and fungi, reflecting the eukaryotic origins of most of the non-plastidial parts of the plant cell. Malonyl-CoA synthesized by the plastidial acetyl-CoA carboxylase enters fatty acid synthesis whereas malonyl-CoA synthesized by the cytosolic enzyme is used for elongation of fatty acid chains in the endoplasmic reticulum (see Section 5.3), and for the synthesis of flavonoids (see Section 6.5).

Fatty acid synthesis in the grass family differs from that in eudicots and non-grass monocots. In the grass family, the plastidial as well as the cytosolic forms of acetyl-CoA carboxylase are monomeric. The monomeric plastidial enzyme is inhibited by synthetic chemicals that have no effect on the multimeric plastidial enzyme of other plants. This has allowed the development of grass-specific herbicides that kill grasses by preventing fatty acid synthesis but have no effect on other plants.

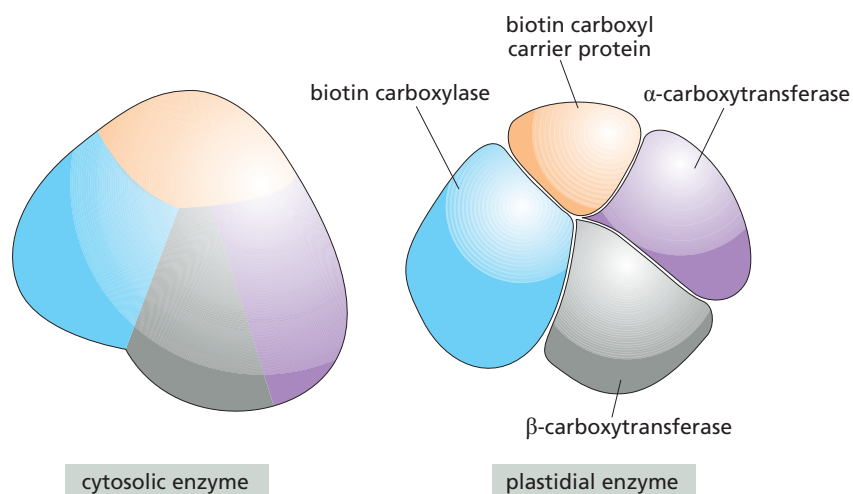


Figure OM6.1-1 Structure of cytosolic and plastidial acetyl-CoA carboxylases. The carboxylation of acetyl-CoA begins with the carboxylation (with bicarbonate as reactant) of a biotin cofactor on the enzyme, an ATP-requiring reaction. The carboxyl group is then transferred to acetyl-CoA to form malonyl-CoA. The plastidial enzyme in most plants consists of four separate proteins (*right*): biotin carboxyl carrier protein, to which biotin is attached; biotin carboxylase; and two carboxytransferase subunits, which transfer the carboxyl group from biotin to acetyl-CoA. The cytosolic enzyme (*left*) and the plastidial enzyme in grasses are single, multidomain proteins in which different domains carry out these functions. The cytosolic enzyme is not involved in fatty acid synthesis; instead it provides malonyl-CoA for related two-carbon molecular elongation reactions outside the plastid.

OM6.2 MANIPULATION OF THE LYSINE CONTENT OF MAIZE KERNELS

Most attempts to produce high-lysine plants have centered on abolishing feedback regulation, so that the accumulation of lysine no longer restricts the rate of its synthesis and the content of lysine-rich proteins is increased. The research was done initially on *Arabidopsis*, then strategies successful in this model plant were tested in the crop plant maize. The most successful approach has been to manipulate dihydrodipicolinate synthase to make it feedback-insensitive. This has been achieved in two ways. First, populations of mutated *Arabidopsis* plants were screened for mutants resistant to a toxic analog of lysine (the analog is toxic because it can be incorporated into proteins in place of lysine, rendering the proteins dysfunctional). Mutants that are resistant to this analog survive because they have an altered dihydrodipicolinate synthase that is not subject to feedback inhibition by lysine. The resulting high lysine levels in the mutant reduce the frequency of incorporation of the toxic lysine analog into protein, so the plant has sufficient normal protein to survive. Second, bacterial genes encoding dihydrodipicolinate synthase were introduced into *Arabidopsis* plants. The bacterial enzyme is much less sensitive to lysine inhibition than the plant enzyme, so this strategy resulted in elevated lysine levels in the transgenic plants (Figure OM6.2-1).

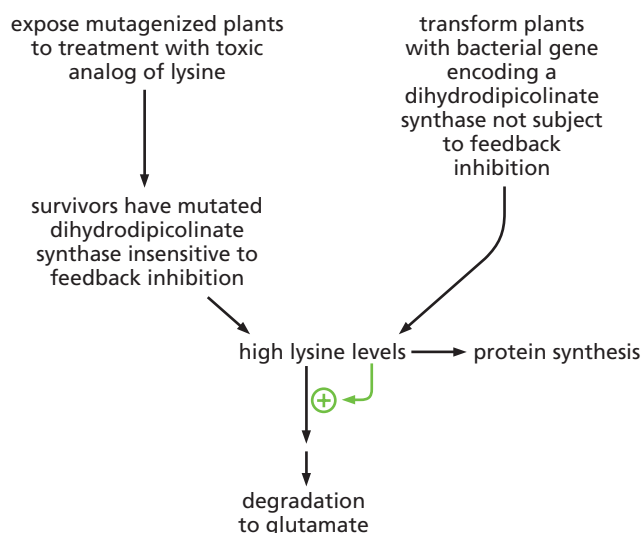


Figure OM6.2-1 Two strategies for increasing lysine levels in plants. Both strategies targeted dihydrodipicolinate synthase – the first enzyme in the aspartate to lysine pathway after the branch point to lysine (see Figure 6.9) – with the aim of introducing a dihydrodipicolinate synthase that is insensitive to feedback inhibition by lysine. Both strategies were successful in that the feedback insensitivity of the introduced dihydrodipicolinate synthase did allow higher rates of lysine synthesis. However, the increases in the level of free lysine (and of lysine-containing storage proteins) were lower than expected, because high lysine levels induced the activity of an enzyme of lysine degradation. In other words, lysine was involved in feed-forward promotion of its own degradation.

Unfortunately, increasing lysine synthesis by reducing feedback inhibition did not immediately have effects likely to improve the nutritional value of feed crops like maize. Most of the extra lysine in *Arabidopsis* was not incorporated into proteins; instead it was degraded to glutamate and other products including pipercolate and acetyl-CoA by a pathway in which expression of the first enzyme, lysine-ketoglutarate reductase, is induced by lysine. To find a solution to this problem, researchers introduced into *Arabidopsis* both a feedback-insensitive dihydrodipicolinate synthase and a mutation that blocks lysine-ketoglutarate reductase (Figure OM6.2-1). The resulting transgenic plants had very much higher levels of lysine than plants that contained the feedback-insensitive dihydrodipicolinate reductase alone. The strategy has now been successfully deployed in maize: the two manipulations successfully increase the content of lysine-rich proteins in the kernel. These experiments illustrate that lysine levels in plant cells are determined by the rate of lysine synthesis and its incorporation into protein, and also by the rate of lysine degradation. The levels of many other amino acids may also be determined by interactions between synthesis, utilization for protein synthesis or other metabolic pathways, and degradation.

An alternative strategy to increase the lysine contents of crops has been to screen for mutants that have elevated lysine contents. Kernels of the *opaque2* mutant of maize have almost twice the lysine content of normal maize, the additional lysine being in proteins. *Opaque2* is a transcription factor that affects expression of genes encoding some storage proteins. Its loss in the *opaque2* mutant reduces the amounts of several lysine-poor storage proteins, collectively called zeins, with a concomitant increase in the amounts of other, non-zein proteins with higher lysine contents. However, kernels

of the original *opaque2* mutant also had an undesirable soft and floury texture. Zeins in normal kernels occur in protein bodies (protein storage compartments), which are tightly packed together with starch granules to form a hard (vitreous) kernel matrix. Loss of zeins reduces the size and number of protein bodies, resulting in a matrix with a soft texture. Further breeding for improved texture has resulted in new **cultivars** that have both a high lysine content and the required vitreous matrix (**Figure OM6.2-2**).

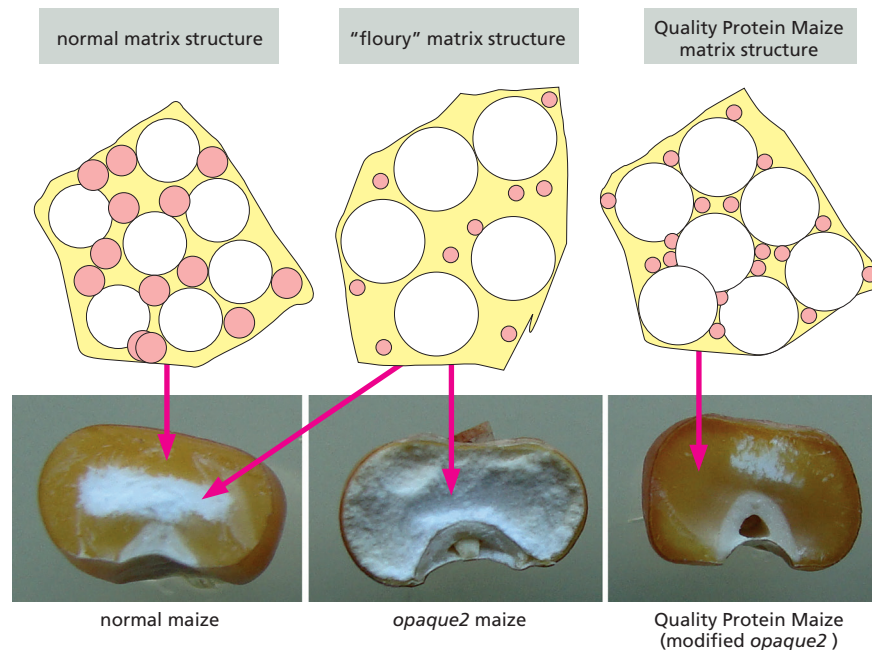


Figure OM6.2-2 Development of high-lysine maize. The endosperm of normal maize consists of starch granules (*white*) embedded in a protein matrix rich in protein bodies (*pink*). It has levels of lysine that are too low for optimal human and animal nutrition. The *opaque2* mutant has much higher lysine levels; however, this improvement is brought about by loss of particular classes of lysine-poor zein proteins that promote protein body formation. The *opaque2* endosperm has few protein bodies and hence an undesirably soft and floury matrix. A breeding program in the late 20th century selected for increased levels of proteins that could replace the function of the lost zeins, restoring protein body formation and hence a hard matrix. The resulting quality protein maize (QPM) cultivars are widely grown and have had a positive impact on human health in countries with protein- and lysine-poor diets. The researchers who introduced QPM (Surinder Vasal and Evangelina Villegas, working at the International Maize and Wheat Improvement Center (CIMMYT) in Mexico) were awarded the World Food Prize in 2000. (From Holding DR (2014) *Front Plant Sci* 5:276. Distributed under a Creative Commons Attribution License: courtesy of David R. Holding.)

OM6.3 UBIQUITIN-MEDIATED PROTEIN DEGRADATION

The left side of **Figure 6.3-1** provides a simplified cartoon of the pathway. In the first phase (1), free ubiquitin is activated and bound to the E1 protein through an adenylate group, then attached via a thioesterase linkage to a cysteine residue of the E1 protein. The ubiquitin moiety is next transferred to a cysteine moiety of the E2 protein, or ubiquitin carrier enzyme. The ubiquitinated E2 protein associates with an E3 ubiquitin ligase that is already associated at a different site with a specific target protein (2). The ubiquitin moiety from the E2 ligase is then transferred to the target protein (3).

The circled cartoon on the right of the figure shows a more detailed representation of the E3 ligase at this stage of the proteolytic process. The main protein components are the scaffold protein CULLIN that serves as a platform for other components; the Ring Box (RBX) protein that binds ubiquitinated E2, the SKP1 protein, and the **F-box protein** that binds the target protein.

A further E2 ubiquitin is recruited, and transferred to the ubiquitin already in place on the target protein (4). This sequence of events is repeated, resulting in a chain of ubiquitin attached to the target protein. The process terminates with detachment of E2 and the E3 ubiquitin ligase from the ubiquitinated target protein and its transfer to the proteasome (5).

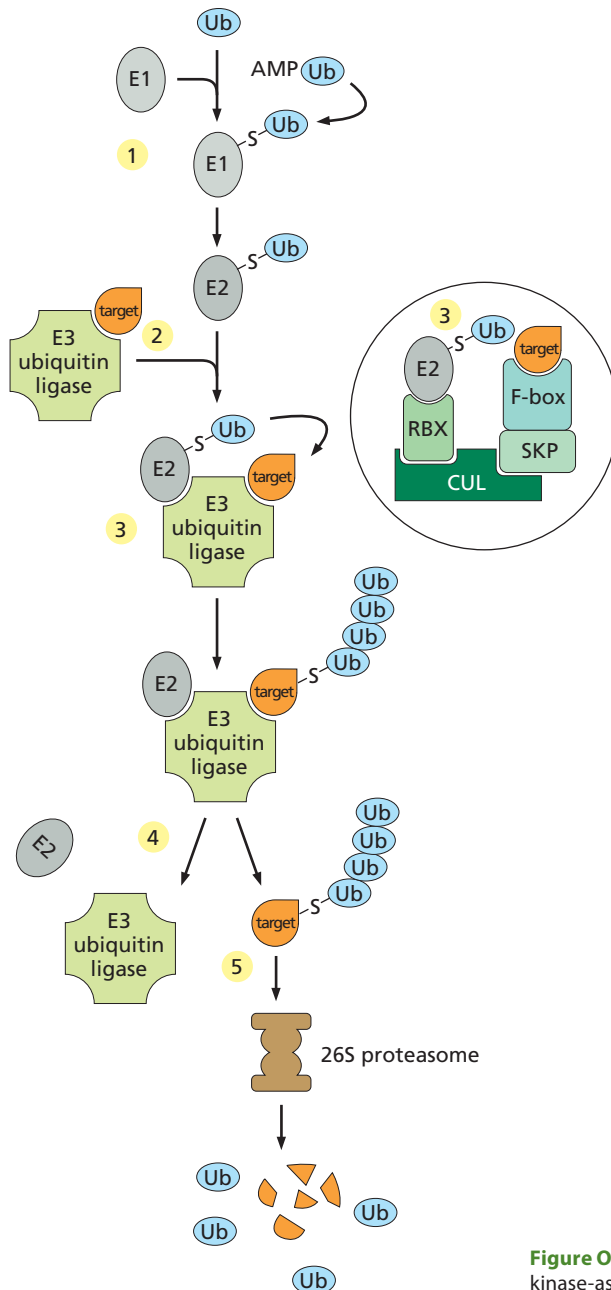


Figure OM6.3-1 The SCF E3 ubiquitin ligase pathway. SCF, S-phase kinase-associated protein1 (SKP1)/CULLIN/F-box.

OM6.4 NON-PROTEIN AMINO ACIDS

Non-protein amino acids and analogs of protein amino acids may accumulate to high levels, and tend to be characteristic of an individual species or group of related species. Examples include the accumulation of canavanine (similar to arginine) and homoserine (an intermediate in threonine, isoleucine, and methionine synthesis) in the seeds of legumes; azetidine-2-carboxylic acid (similar to proline) in lily-of-the-valley (*Convallaria majalis*); hypoglycine in the unripe fruits of *Blighia sapida* (akee); mimosine (similar to tyrosine) in the leaves of the mimosa relative *Leucaenea leucocephala*; and the glutamate analog oxalyldiaminopropionic acid (ODAP) in the seeds of grass pea (*Lathyrus sativus*). These unusual amino acids may serve as nitrogen reserves, but may also act as deterrents against predation and competition from other plants because of their toxicity.

The reasons for toxicity of non-protein amino acids vary widely. As close analogs of the amino acids of proteins, some non-protein amino acids including canavanine are toxic because tRNAs of herbivores recognize and incorporate them into proteins in place of the correct amino acid, thus disrupting protein synthesis. Some beetle herbivores are able to consume canavanine-rich legume seeds because their arginyl-tRNAs can discriminate against canavanine. Hypoglycine is toxic to humans because it is converted in the body to a metabolite that inhibits fatty acid oxidation, preventing the use of fat reserves. Glycogen reserves are depleted, leading to hypoglycemia. Mimosine is toxic to mammals because it is an inhibitor of DNA replication. Some ruminants can safely consume leaves containing mimosine because it is metabolized to harmless compounds by bacteria in the rumen. ODAP is a neurotoxin: it binds to and thus over-excites glutamate receptors in the human nervous system. Grass pea is a potentially nutritious crop with exceptional drought, flood, and salinity tolerance, widely grown in East Africa and the Indian subcontinent. It can be eaten safely as part of a wider diet, but excessive consumption during drought when other crops have failed results in the disabling disease neuropathy. International collaboration is underway to breed new grass pea varieties with reduced levels of ODAP.

OM6.5 A GENE CLUSTER ENCODING THE ENZYMES OF NOSCAPINE SYNTHESIS

Several of the important drugs synthesized in the developing seed capsules of opium poppies are derived from reticuline, which is itself formed by condensation and modification of two molecules of the aromatic amino acid tyrosine. Reticuline is the substrate for the synthesis of noscapine and two other opiates of the same class (isoquinoline alkaloids), morphine and codeine. The pathway to noscapine is specified by genes that include a cluster of ten contiguous genes encoding five different classes of enzymes. This cluster is unique to noscapine-synthesizing varieties of poppy. **Figure OM6.5-1** shows the ten genes involved and the classes of enzymes they encode, and for six of the genes indicates where on the pathway their products act.

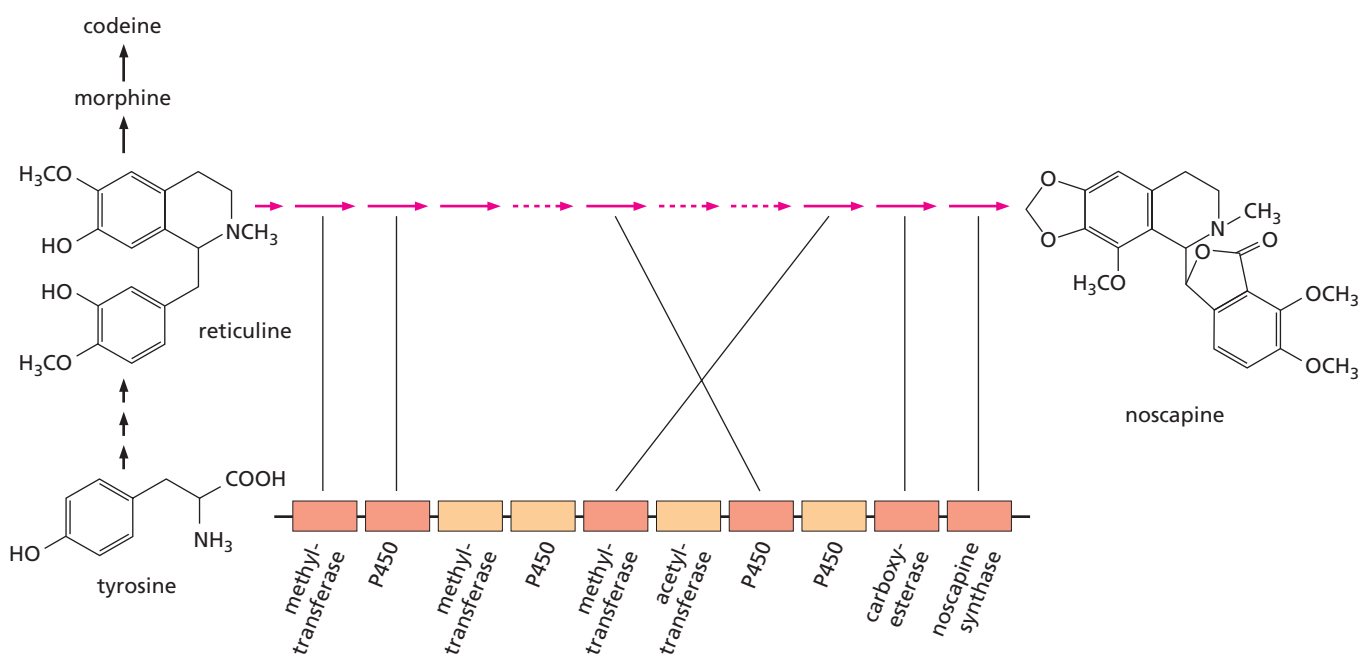


Figure OM6.5-1 A gene cluster encoding the enzymes of noscapine synthesis in *Papaver somniferum*. Several of the important drugs synthesized in the developing seed capsules of opium poppies are derived from reticuline, which is itself formed by condensation and modification of two molecules of the aromatic amino acid tyrosine. Reticuline is modified to make noscapine by a series of reactions (pink arrows) involving five different classes of enzymes. At least six of the enzymes are encoded in a ten-gene cluster (represented as boxes) that is unique to noscapine-synthesizing varieties of poppy. Other genes in the cluster (orange boxes) encode enzymes that may well catalyze uncharacterized steps on the pathway (dashed arrows).

OM6.6 PRODUCTS DERIVED FROM MONOLIGNOLS

The three precursor monolignols give rise to three structural units in lignin polymers: *p*-hydroxyphenyl (H; from coumaryl alcohol), guaiacyl (G; from coniferyl alcohol), and sinapyl (S; from sinapyl alcohol) moieties. The composition of lignin varies considerably between species. In ferns and gymnosperms, only *p*-coumaryl and coniferyl alcohols are synthesized, hence the lignin is composed of H and G units only. In angiosperms, sinapyl alcohol is also synthesized, but most angiosperms do not use *p*-coumaryl alcohol for lignin synthesis. Thus, the wood of angiosperm trees is composed of S and G units only. Grasses use all three alcohols in lignin synthesis, so lignin in this plant family contains S, H, and G units. Lignin composition also varies between cell types and in different regions of the cell wall in the trunks of trees. For example, in black spruce (*Picea mariana*), a gymnosperm tree, lignin in the secondary cell wall of water-conducting vessels (tracheids) is comprised predominantly of G units, whereas lignin found in the middle lamella – the zone separating the cell walls of adjacent cells – consists predominantly of H units. In birch (*Betula papyrifera*), an angiosperm tree, the lignin of fiber cells in the wood contains both G and S units, but the lignin of the water-conducting vessels in the wood consists predominantly of G units.

The monolignol pathway also provides precursors for other phenolic compounds, including some important volatile compounds. Many monolignol derivatives are valuable as spices and flavorings, and are synthesized by specific groups or species of plants (**Figure OM6.6-1**). Vanillin, the aroma compound produced in the seed pods of climbing orchids of the genus *Vanilla* as well as in other plant species, is a precursor in plants for synthesis of piperonal, which is a component of the aroma of dill (*Anethum graveolens*), violets (*Viola* spp.), and black pepper (fruits of *Piper nigrum*). Eugenol, a major component in the aroma of cloves (flower buds of the tree *Syzygium aromaticum*), basil (*Ocimum* spp.), nutmeg (seed of the tree *Myristica fragrans*), and bay (*Laurus nobilis*), is derived from coniferyl alcohol. Further elaborations of eugenol give rise to myristicin, a component of the aroma of nutmeg, parsley (*Petroselinum crispum*), and dill.



Figure OM6.6-1 Herbs and spices in which the main aroma compounds are monolignol derivatives. (A) Three common culinary spices: black peppercorns (top left), cloves (top right), and nutmeg (bottom). (B) The plant parts harvested to provide the spices shown in (A). (Left) Unripe fruits of the vine *Piper nigrum*. The fruits are dried to provide black peppercorns. (Center) Developing fruits of the nutmeg tree *Myristica fragrans*. Intact fruits are at the top; at the bottom is a fruit split open to show the nutmeg surrounded by a red covering (an aril) which is harvested as the spice mace. Nutmeg and mace have similar profiles of aroma compounds derived from monolignols. (Right) Developing buds of the clove tree *Syzygium aromaticum*. The buds are harvested and dried to provide the spice cloves. (C) Flowers and dried seed pods of a vanilla orchid. *Vanilla planifolia* pods are the commercial source of vanillin (vanilla extract). (D) Flowers of the herb dill (*Anethum graveolens*). Mature seeds and leaves are used in cookery. (E) Leaves of the herb parsley (*Petroselinum crispum*). This curly-leaved variety and the more strongly flavored flat-leaved parsley are widely used in cookery and as garnishes. (B left, from "Black Pepper Plant". Licensed under CC BY SA 4.0 via Wikimedia Commons. B center top, from "Nutmeg on Tree" by Joe Ravi. Licensed under CC BY SA 3.0 via Wikimedia Commons. B center bottom, from "Nutmeg fruit seed and aril" by Slashme. Licensed under CC BY SA 4.0 via Wikimedia Commons. B right, from "The flowers of clove trees in Pemba Island" by Prof. Chen Hualin. Licensed under CC BY SA 4.0 via Wikimedia Commons. C left, from "Vanilla 6beans" by B. Navez. Licensed under CC BY SA 3.0 via Wikimedia Commons. C right, "Vanilla planifolia" by Malcolm Manners, Lakeland, FL, USA. Licensed under CC BY 2.0 via Wikimedia Commons. D, from "20141003Anethum graveolens" by AnRo0002. E, from "Parsley Curled" by Donovan Govan. Licensed under CC BY SA 3.0 via Wikimedia Commons.)