Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

ANNUAL REVIEWS

Annual Review of Plant Biology

Evolutionary History of Plant Metabolism

Hiroshi A. Maeda[1] and Alisdair R. Fernie[2]

[1]Department of Botany, University of Wisconsin–Madison, Madison, Wisconsin 53706, USA; email: maeda2@wisc.edu

[2]Max-Planck-Institut für Molekulare Pflanzenphysiologie, 14476 Potsdam-Golm, Germany; email: fernie@mpimp-golm.mpg.de

Annu. Rev. Plant Biol. 2021. 72:185–216

First published as a Review in Advance on April 13, 2021

The *Annual Review of Plant Biology* is online at plant.annualreviews.org

https://doi.org/10.1146/annurev-arplant-080620-031054

Copyright © 2021 by Annual Reviews. All rights reserved

ANNUAL REVIEWS CONNECT

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Keywords

evolution of metabolism, specialized metabolism, endosymbiotic gene transfers, horizontal gene transfers, gene duplication, domestication

Abstract

Tremendous chemical diversity is the hallmark of plants and is supported by highly complex biochemical machinery. Plant metabolic enzymes originated and were transferred from eukaryotic and prokaryotic ancestors and further diversified by the unprecedented rates of gene duplication and functionalization experienced in land plants. Unlike microbes, which have frequent horizontal gene transfer events and multiple inputs of energy and organic carbon, land plants predominantly rely on organic carbon generated from CO_2 and have experienced very few, if any, gene transfers during their recent evolutionary history. As such, plant metabolic networks have evolved in a stepwise manner and on existing networks under various evolutionary constraints. This review aims to take a broader view of plant metabolic evolution and lay a framework to further explore evolutionary mechanisms of the complex metabolic network. Understanding the underlying metabolic and genetic constraints is also an empirical prerequisite for rational engineering and redesigning of plant metabolic pathways.

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

Contents

1. INTRODUCTION

A fundamental long-term goal of biological research has been to understand how the complex biomolecular networks underpinning life took on the form(s) that we observe today (16, 50). Cellular metabolism is arguably one of the earliest such networks (166) and therefore represents a wonderful model system for studying network evolution. In contrast to the evolution of development (evo-devo), which has received considerable research attention (135), the evolution of metabolism has been subjected to far less scrutiny. However, metabolism underlies developmental processes (e.g., synthesis of proteins, membranes, and cell walls), and hence metabolic evolution is indeed a part of evo-devo. Within the metabolic network, many thousands of diverse biochemical processes are linked in highly tailored systems that have been acted upon by billions of years of evolution. Understanding the stepwise development of the metabolic network also highlights how historical constraints impact the current network functions and provides strategies about how to improve them. This is particularly prominent in land plants that rely on a single carbon source, photosynthetic CO_2 fixation, and have had limited gene transfers in their recent history after terrestrialization.

Although recent diversification of plant specialized metabolic pathways has been extensively reviewed (115, 178), to our knowledge, the overall evolution of plant metabolism has not been reviewed to date. To capture a broader evolutionary history of the plant metabolic network from the beginning to today, this article starts from (❶) the origin of metabolism, followed by (❷) the evolution of plant primary metabolism, (❸) the emergence of plant-specific metabolism, and (❹) the diversification of plant primary metabolism and (❺) specialized metabolism, and ends with (❻) metabolic alterations introduced during crop domestication. In each section, we highlight various modes of metabolic evolution that primarily contributed at different stages of plant metabolic evolution (**Figure 1**). Overall, we provide a framework to further explore the evolutionary history of plant metabolism, which in turn enables the effective redesign of plant metabolic networks.

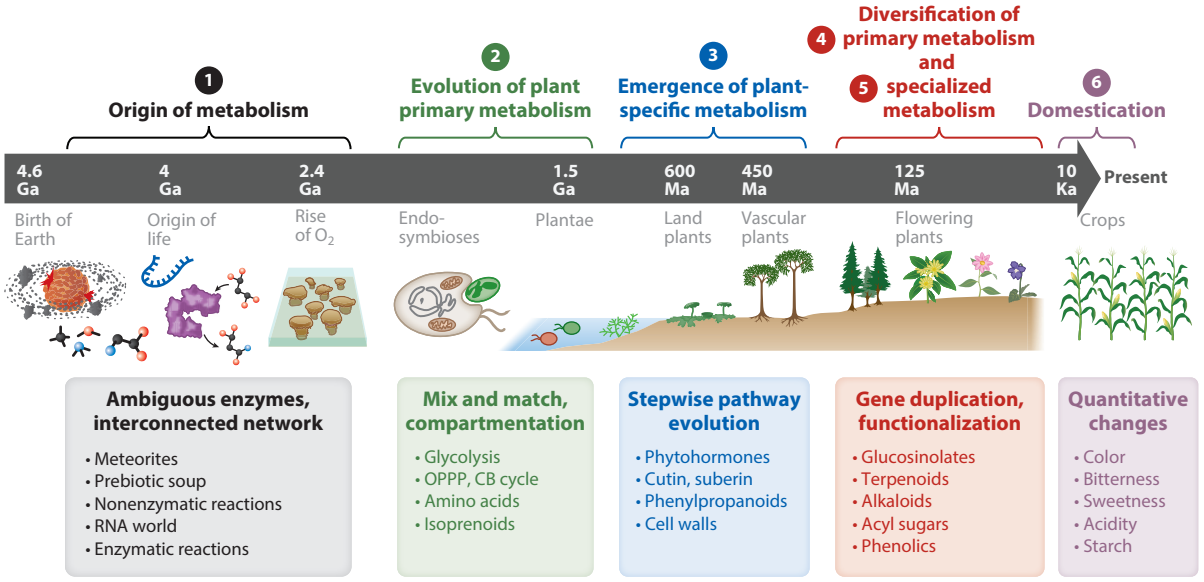


Figure 1

Schematic diagram of the overall evolutionary history of plant metabolism along with the estimated timelines of major Earth history and events indicated in Ga, Ma, or Ka. Various stages of plant metabolic evolution described in different sections of this article are colored and numbered from (❶) to (❻). Various modes of metabolic evolution that primarily contributed in each stage, as well as key metabolites or pathways that emerged or developed, are indicated in boxes. Timeline is not drawn to scale. Abbreviations: CB cycle, Calvin-Benson cycle; Ga, billion years ago; Ka, thousand years ago; Ma, million years ago; OPPP, oxidative pentose phosphate pathway.

2. ORIGIN OF CORE METABOLISM

When life began on Earth around four billion years ago (151), metabolism is believed to have been centered around very few chemistries, which probably formed in emergent replicating or organismal units and later developed as primitive cells (16, 152). Ancestral life forms likely inhabited an environment, known as the primordial soup, rich in spontaneously formed organic compounds (e.g., amino acids) of the primordial world (50). Additionally, meteoric evidence suggests that the presence of ribose and other simple sugars was not confined to the Earth (30). As more organisms occupied the primordial environment, less-abundant compounds became depleted, which imposed increased selective pressure for synthesizing them (50) through various modes of metabolic pathway evolution (see the sidebar titled Hypotheses of Early Metabolic Pathway Evolution) (58, 65, 74). Recent studies also revealed that nonenzymatic metabolism (e.g., of glycolysis) can operate under conditions similar to primordial Archean environments (82, 83, 137). These chemical reaction sequences, which were constrained by cofactors and conditions available at that time, could have been incorporated into cellular metabolism. Such a network was later facilitated by the rise of enzymes, perhaps initially by RNA (or ribozymes) and then by proteins—RNA world and RNA-protein world—that likely led to replication cycles (128). A simple H_2-driven CO_2 fixation pathway (177), known as the acetyl-CoA pathway (or the Wood-Ljungdahl pathway) (96), leading to the formation of pyruvate and fatty acids (11, 124), was also proposed to be one of the earliest metabolic pathways based on the discovery of alkaline hydrothermal vents in the deep ocean (84, 158). Such microenvironments harbor chemistries reminiscent of the primitive Earth characterized by steep gradients of heat, pH, and reduction potential across thin inorganic barriers. Today, these vents host a rich microbial community that represents some of the deepest branches of the

Primordial soup: mixtures of organic compounds in the early Earth environment, which were generated through abiotic reactions or provided via meteorites that bombarded the early Earth

Nonenzymatic metabolism: similar to the metabolism that takes place in organisms but mediated by chemical reactions without enzymes

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

HYPOTHESES OF EARLY METABOLIC PATHWAY EVOLUTION

Several hypotheses have been proposed for the evolution of metabolic pathways. The backward (or retrograde) hypothesis describes a stepwise attainment of individual reactions in a backward manner in regard to the reaction sequence—the acquisition of the final step, followed by the upstream steps in reverse order (65). However, intermediates and precursors, which themselves have no use, must be available in the primordial environment for this process to take place, meaning that new metabolites do not emerge. The forward hypothesis proceeds from functional intermediates and precursors, which are further converted into new products, leading to a stepwise development of a metabolic pathway in a forward direction (58). Once enzymes' repertoires have expanded through gene duplications, enzymes that have ambiguous substrate specificity can be recruited to form new pathways in a patchwork manner, known as the patchwork hypothesis (74, 93). It is important to note that these hypotheses are not mutually exclusive and likely support the evolution of different parts of metabolic pathways and networks.

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

tree of life (i.e., acetogenic bacteria and methanogenic archaea) and likely dates back to the last universal common ancestor (LUCA).

Initially, a limited number of starter enzymes with broad substrate specificity likely catalyzed multiple reactions (74); therefore, primordial metabolic pathways were highly interconnected (50) (**Figure 2*a***). These starter enzymes then underwent duplication and functional divergence to expand the number of enzymes, often with more specialized functions, leading to more distinct metabolic pathways where individual reactions were increasingly carried out by designated enzymes (74) (**Figure 2*b***). The emergence of new enzymes also expanded the overall metabolic network by further converting existing metabolites and producing new metabolites, which created new metabolic connections through metabolite–enzyme coevolution (122). Additionally, certain reaction modules are repeatedly found in metabolic networks (74, 122) (**Figure 2*c***), suggesting that the repurposing of not only enzymes but also reaction modules often took place through divergent and also possibly convergent evolution. Notably, promiscuous enzymes and multifunctional enzymes are still widespread in extant organisms (e.g., detoxification enzymes and aminotransferases) (75, 85), which support metabolic plasticity through so-called underground metabolism (31, 123) and also provide starting points for the evolution of new enzymes. Here, we describe the properties and origins of core metabolic pathways, highlighting various modes of pathway evolution and discussing how some of the evolutionary origins may still impact modern-day metabolism.

RNA world: a hypothetical evolutionary stage in which self-replicating RNA molecules evolved before proteins or DNA and likely catalyzed some chemical reactions

Last universal common ancestor (LUCA): the most recent ancestral organismal form that likely existed between 3.5 to 4 billion years ago, from which all organisms evolved

Starter enzyme: a primordial enzyme that emerged independently and served as a precursor for the later evolution of a variety of enzymes

Promiscuous enzyme: an enzyme that has activities beyond its primary catalytic function and hence exhibits additional but often weak side reactions that were not selected during evolution

2.1. Amino Acid Biosynthesis

While simple and stable amino acids, such as aspartate and alanine, were relatively abundant in the primordial soup, some amino acids (e.g., methionine, histidine, and tyrosine) became depleted early and needed to be synthesized (93). It has been proposed that threonine and methionine biosynthetic pathways potentially formed in a backward manner in respect to their reaction sequences (see the sidebar titled Hypotheses of Early Metabolic Pathway Evolution) (65); threonine and methionine were initially depleted and synthesized from homoserine and then from aspartate (**Figure 2*c***), as both were likely available in primordial environments.

A more complicated evolutionary path has been proposed for branched chain amino acid (BCAA) biosynthesis because the present acetolactate pathway derived from pyruvate and α-ketobutyrate (157) cannot develop in a backward manner, due to unstable intermediates (e.g., β-keto acid acetolactate) that would not have accumulated in the primordial soup. Instead, branched short chain fatty acids, which were abundant in meteorites, likely underwent

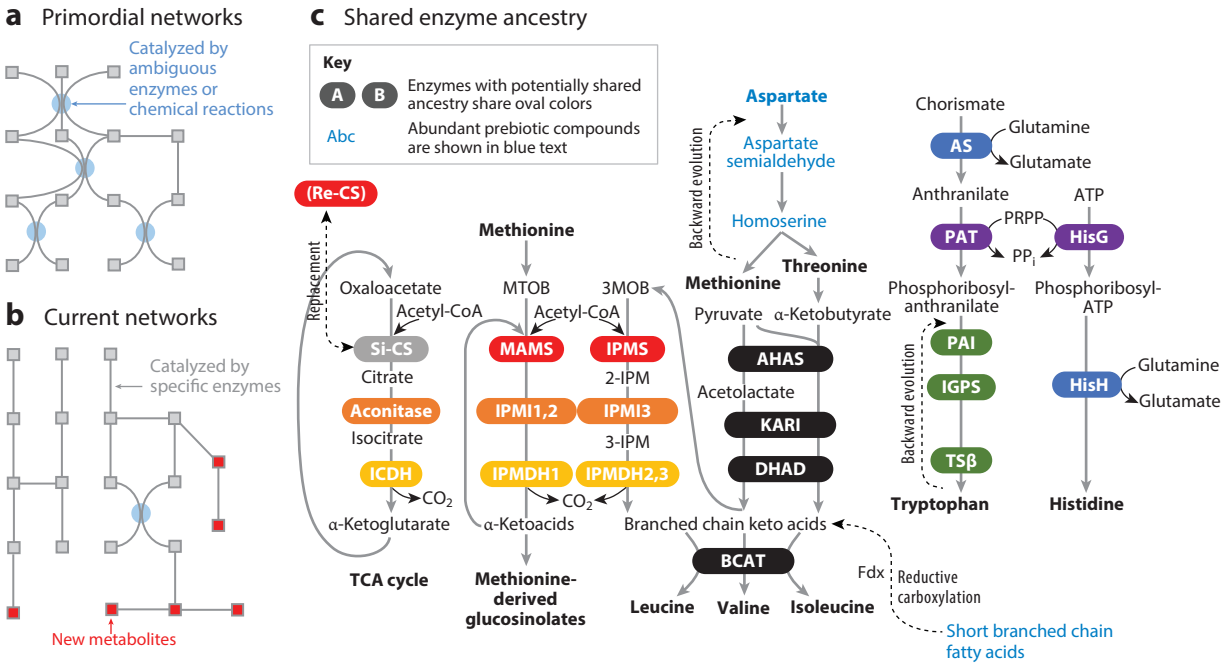


Figure 2

Metabolic enzymes and modules with potentially shared common ancestry. Compounds (*square nodes*) are converted via chemical or enzymatic reactions (*gray lines*) in (*a*) primordial and (*b*) current metabolic networks. Primordial networks were likely highly interconnected via ambiguous enzymes or chemical reactions (*blue circles*), whereas current networks are often catalyzed by designated enzymes and are further expanded to produce new metabolites. (*c*) Some enzymes and reaction modules from different pathways potentially have shared ancestry, as indicated by boxes with the same colors, though some of them might have evolved convergently. Black boxes over multiple arrows are multifunctional enzymes that are still involved in multiple pathways today. Earlier evolutionary processes of some pathways from abundant prebiotic compounds (*blue letters*) are shown in dotted arrows. Re-CS and Si-CS are stereospecific citrate synthases found in some anaerobes (thus shown in parenthesis) and other organisms including plants, respectively. Abbreviations: 3MOB, 3-methyl-2-oxobutanoate; AHAS, acetohydroxyacid synthase; AS, anthranilate synthase; ATP, adenosine triphosphate; BCAT, branched chain amino acid aminotransferase; DHAD, dihydroxyacid dehydratase; Fdx, ferredoxin; HisG, ATP phosphoribosyltransferase; HisH, the glutamine amidotransferase domain of imidazole glycerol-phosphate synthase; ICDH, isocitrate dehydrogenase; IGPS, indole-3-glycerol phosphate synthase; IPM, isopropylmalate; IPMDH, isopropylmalate dehydrogenase; IPMI, isopropylmalate isomerase; IPMS, isopropylmalate synthase; KARI, ketol-acid reductoisomerase; MAMS, methylthioalkylmalate synthase; MTOB, 4-methylthio-2-oxobutyrate; PAI, phosphoribosylanthranilate isomerase; PAT, phosphoribosylanthranilate transferase; PP_i, inorganic pyrophosphate; PRPP, 5-phosphoribosyl pyrophosphate; TSβ, tryptophan synthase β subunit.

reductive carboxylation, potentially mediated by ferredoxin, to produce keto acids, which were then transaminated to BCAAs (81) (**Figure 2*c***). Analogous reductive carboxylation pathways might have also contributed initially to aromatic amino acid synthesis and are still used in some rumen microbes (145). Later, the acetolactate pathway evolved and replaced the reductive carboxylation pathway for BCAA biosynthesis by the time LUCA emerged, as the acetolactate pathway is conserved among prokaryotes and eukaryotes. The early BCAA pathways were most certainly catalyzed by ambiguous enzymes, and the three consecutive steps of the isoleucine and valine/leucine pathways are still catalyzed by shared enzymes—acetohydroxyacid synthase (AHAS), ketol-acid reductoisomerase (KARI), and dihydroxyacid dehydratase (DHAD) (157) (**Figure 2*c***). Additionally, in several prokaryotes, isopropylmalate isomerase (IPMI) in leucine biosynthesis can also use homocitrate, an intermediate of lysine biosynthesis (189). Thus, even in the current metabolic network, multifunctional enzymes are still present and likely coordinate

Multifunctional enzyme: an enzyme that has more than one *functional* catalytic activity

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

different amino acid pathways, while feedback regulation of key enzymes, which likely evolved later, provides the regulation of individual pathways (74, 184).

Underground metabolism: a metabolic network formed via the promiscuity of enzymes and nonenzymatic reactions, which are beyond but connected to well-defined metabolic networks

The tryptophan and histidine biosynthetic pathways are directly interconnected to nucleotide biosynthesis and nitrogen assimilation, as both pathways require glutamine and 5-phosphoribosyl 1-pyrophosphate (PRPP) as substrates (72, 134). The anthranilate synthase (AS) enzyme and the glutamine amidotransferase (HisH) subunit of imidazoleglycerol-phosphate synthase incorporate the amino group of glutamine, while phosphoribosyl-anthranilate transferase (PAT) and ATP phosphoribosyltransferase (HisG) transfer the 5-phosphoribosyl group of PRPP into the tryptophan and histidine biosynthetic pathways, respectively (43) (**Figure 2*c***). These two corresponding enzymes likely share common starter enzymes between tryptophan and histidine biosynthesis. Both pathways are highly conserved across bacteria, archaea, and eukaryotes and thus were probably already present in LUCA (2, 183). Interestingly, the last three tryptophan pathway enzymes—phosphoribosylanthranilate isomerase (PAI), indole-3-glycerol phosphate synthase (IGPS), and tryptophan synthase α subunit (TSα)—catalyze different reactions but all have a $(\beta\alpha)_8$-barrel fold and a common binding site for the 5-phosphoribosyl moiety of their corresponding substrates (181). Indeed, PAI activity could be obtained through directed evolution of TSα (40). Thus, these three steps likely evolved in a backward manner through a series of ancient gene duplications and neofunctionalization (**Figure 2*c***).

In contrast to the highly conserved tryptophan and histidine pathways, lysine biosynthesis appears to have evolved multiple times, as there are α-aminoadipic acid and some variations of the diaminopimelic acid (DAP) pathways (172). Many enzymes of the DAP lysine pathway are homologous to those of arginine biosynthesis, suggesting that they were originally involved in both pathways catalyzed by ambiguous enzymes. Similarities of some codons for lysine (AAA, AAG) and arginine (AGA, AGG) also support that these two basic amino acids were initially used ambiguously in protein synthesis (182). Plants and cyanobacteria use a unique DAP pathway for lysine biosynthesis, which is mediated by DAP aminotransferase that bypasses three enzymatic steps and acyl intermediates of eubacterial DAP pathways (70).

2.2. Glycolysis and Pentose Phosphate Pathways

The Ralser group (83) showed that nonenzymatic pentose phosphate–like reactions can take place under conditions that are similar to the Archean ocean and contain reduced Fe(II). These reactions include the formation and interconversion of glucose, pyruvate, ribose 5-phosphate, and erythrose 4-phosphate, thus antedating the reaction sequences similar to those of extant enzyme-based metabolic pathways. Intriguingly, pH gradients alter the metabolic network, leading to a shift in the relative activities of glycolysis and the oxidative pentose phosphate pathway (83). While a number of questions remain regarding the evolution of glycolysis (e.g., the primordial source of glucose 6-phosphate) (82), the presence of nonenzymatic reactions and thereby the availability of metabolic precursors render the innovation of enzymatic metabolism a stepwise problem (136). Indeed, it has been shown that the simple amino acid glycine was able to improve the nonenzymatic catalysis for the formation of fructose 1,6-bisphosphate (110). Whatever the origins of the pathway, the evolution of glycolysis is impressive—with in silico evaluations suggesting that the lower part of glycolysis carries a higher flux than any biochemically possible alternative (28)—and even the less-efficient Entner-Doudoroff pathway was shown to represent a trade-off between energy yield and protein cost (49).

The reductive pentose phosphate pathway, also known as the Calvin-Benson cycle, is the most widely distributed CO_2 fixation pathway today but likely emerged relatively late. Among at least five alternative modes of CO_2 fixation pathways found in different organisms (51), the

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

reductive acetyl-CoA pathway driven by H_2 is likely the earliest mode of CO_2 fixation pathway (51, 177), though its prebiotic pathway likely used native transition metals (e.g., F^0, Ni^0) to reduce CO_2 into acetate and pyruvate (171). Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco), the first committed enzyme of the Calvin-Benson cycle, was likely involved initially in nucleotide metabolism, such as adenosine 5′-monophosphate metabolism as seen in type III Rubisco of some archaea (144, 168). Therefore, Rubisco likely evolved before oxygenic photosynthesis and hence still cannot discriminate O_2 from CO_2, which comes as a trade-off for reduced catalytic rate (39).

2.3. TCA Cycle

The evolutionary origins of the TCA cycle are still largely unclear (137), but the early TCA cycle was likely reductive—and still is in several species (14). Experimentally, several TCA cycle–like reactions were demonstrated in the presence of ultraviolet (UV) light and semiconductor particles (196), while a suite of reactions was shown to be promoted by Zn^{2+}, Cr^{3+}, and Fe^0 ions (118). More recently, a nonenzymatic acetate-driven TCA cycle, corresponding to the Wood pathway, was identified in bacteria (171). Not content with the demonstration that glycolysis could be nonenzymatically assembled, the Ralser group (82) also showed the same for the oxidative segment of the TCA cycle, this time in the presence of sulfate radicals, which were abundant in meteorites (27). Similar to this finding, the results of Springsteen et al. (163) indicate that oxidizing agents play an important role in the evolution of the TCA cycle. Interestingly, phylogenetic data indicate that the TCA cycle was, at least partially, present in LUCA (177), though the directionality of the cycle remains inconclusive. Indeed, this is compounded by the recent discovery of a reversible citrate synthase in a facultative chemolithoautotrophic thermophile (125). A more recent structural study further suggests that citrate synthase (the *Si*-citrate synthase type found in most organisms, including in plants) is derived from an ancestral citryl-CoA ligase that operates in the reverse TCA cycle (173), suggesting that the oxidative TCA cycle most likely did evolve from the reductive one.

Some TCA cycle enzymes might be derived from biosynthetic enzymes of leucine, which likely was exhausted from the primordial soup and needed to be synthesized very early (81). Examples include the potential recruitment of citrate synthase (i.e., the ancient *Re*-citrate synthase type found in anaerobic bacteria) from isopropylmalate synthase (IPMS), aconitase from IPMI, and isocitrate dehydrogenase from isopropylmalate dehydrogenase (IPMDH) (**Figure 2*c***). These three enzymes constitute a common reaction module of C1 elongation, which is also found in other pathways, such as α-aminoadipic acid lysine and methionine-derived glucosinolate biosynthesis (60, 122, 189).

2.4. Photosynthesis

Evolution of photosynthesis enabled the conversion of sunlight energy into chemical energy, which supports the life of photosynthetic organisms as well as many heterotrophs that consume them. Initially, H_2-dependent chemoautotrophy likely emerged and was followed by nonoxygenic phototrophy that used weak far-red light of hydrothermal vents and H_2S as an electron donor, as seen in purple bacteria (185). Later, oxygenic photosynthesis emerged in cyanobacteria, which utilize sunlight energy and water as the electron donor, generating O_2 (104). This was also accompanied by the evolution of photosynthetic pigments: Cobalamin (vitamin B_{12}, a cobalt-containing cofactor of DNA synthesis and amino acid metabolism) likely evolved to a zinc protoporphyrin IX for weak far-red light harvesting, followed by further conversions into magnesium-based chlorophylls for capturing sunlight (104). This represents an example of stepwise forward development of metabolic pathway evolution (58) (see the sidebar titled Hypotheses of Early Metabolic Pathway

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

Evolution). The core proteins of photosynthetic reaction centers are highly conserved structurally and likely share a common ancestor form, which underwent multiple duplication events to eventually form the heterodimeric structures [e.g., D1/D2 proteins of photosystem II (PSII)] (48). A nonphotosynthetic cyanobacteria ancestor then acquired both PSI and PSII, likely through horizontal (or lateral) gene transfer (HGT or LGT) events. PSI and PSII together enabled oxygenic photosynthesis (161), which drastically increased oxygen concentration and led to the Great Oxidation Event around 2.4 Ga (48). Due to the shifted redox state of Earth's atmosphere, many reactions of anoxic metabolism were replaced by aerobic reactions, including glucose oxidation through glycolysis and highly exergonic aerobic respiration (with O_2 as a terminal electron acceptor). Therefore, modern-day metabolic networks are most likely very different from those of LUCA, which was presumably anaerobic (177). This major shift in metabolic networks coincides with the emergence of complex eukaryotic life (138) (**Figure 1**).

Horizontal (or lateral) gene transfer (HGT or LGT): a transfer of a genetic material between organisms via a process other than vertical gene transfer (VGT)

Great Oxidation Event: the major atmospheric rise of molecular oxygen to appreciable levels around 2.4 Ga, mainly due to cyanobacterial oxygenic photosynthesis

Endosymbiotic gene transfer (EGT): a specific type of horizontal gene transfer by which a genetic material is transferred between organisms through an endosymbiotic event

3. MOSAIC ORIGIN OF PLANT PRIMARY METABOLISM

Planta (Archaeplastida) genomes are derived from at least three origins. A eukaryotic host (an archaebacterium) acquired mitochondria through endosymbiosis of an α-proteobacteria progenitor (related to *Rickettsia*), followed by another endosymbiosis of a cyanobacterial ancestor (related to *Gloeomargarita*) giving rise to the plastids (106, 191) (**Figure 3*a***). Since all three lineages had the majority of primary metabolic pathways and enzymes (with some exceptions, such as oxygenic photosynthesis, which was only present in the cyanobacterial ancestor), primary metabolic pathways of plants are the results of the mixing and matching of genes and enzymes derived from different origins. These two endosymbiosis events also provided multiple subcompartments, which enabled remarkable expansion of complex plant metabolic networks that includes multiple redundant pathways of core metabolism with alleviated cross-pathway metabolite enzyme inhibition (1, 167). Furthermore, many plant genes and enzymes show the closest homology to those of unexpected lineages, beyond the three major origins. Thus, much of plant primary metabolism represents highly mosaic origins through ancient gene transfers (**Figure 3**).

3.1. Central Carbon Metabolism of Plants

Plastidic and cytosolic pathways of glycolysis are prime examples of redundant primary metabolic pathways of plants. Both pathways are of mosaic origin and are now nuclear encoded (103). Isoforms of the plastidic pathway are largely of cyanobacterial origin, but triose phosphate isomerase (TPI) is derived from α-proteobacterial endosymbiotic gene transfer (EGT) (**Figure 3*b***). Cytosolic isoforms, by contrast, are predominantly of α-proteobacterial origin, with the exception of phosphoglycerate kinase (PGK) derived from cyanobacterial EGT (103). The best-studied enzyme from an evolutionary perspective is plastidic and cytosolic glyceraldehyde 3-phosphate dehydrogenases (GAPDHs), which are derived from cyanobacteria and α-proteobacterial EGT and dependent on nicotinamide adenine dinucleotide phosphate (NADPH) and nicotinamide adenine dinucleotide (NADH), respectively (105, 132) (**Figure 3*b***). The plastidic PGK and GAPDH are also used in the Calvin-Benson cycle (**Figure 3*b***), whose committed enzyme Rubisco is also of cyanobacterial origin and belongs to the form I with hexadecameric structure of large and small subunits (L_8S_8) (169). Notably, however, plants have class I fructose bisphosphate aldolase (FBA, or simply aldolase), which is distantly related to class II FBA (found in cyanobacteria) and was later duplicated into plastidic and cytosolic isoforms (59, 130) (**Figure 3*b***). Although the cyanobacteria likely brought in a full set of the glycolytic and Calvin-Benson cycle enzymes to the plastids, some of them have been replaced by homologous enzymes from other origins, potentially acquired through HGT and α-proteobacterial EGT.

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

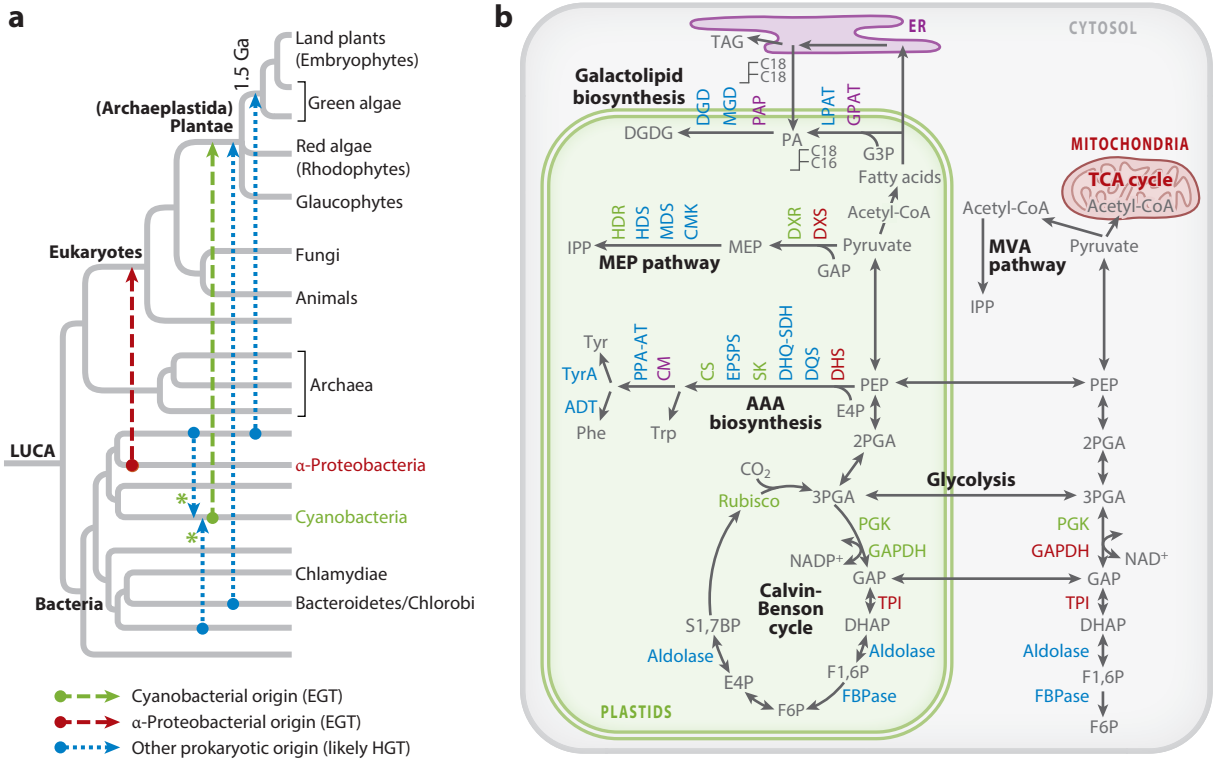


Figure 3

The mix and match evolution of plant primary metabolism. Enzymes and pathways from eukaryotic and various prokaryotic sources were acquired and used to form plant primary metabolic pathways. (*a*) A simplified tree of life, which starts from LUCA, highlights various gene transfer events—endosymbiotic gene transfer (EGT) and horizontal gene transfer (HGT)—to Plantae (Archaeplastida), which diverged around 1.5 Ga. (*b*) Representative primary metabolic enzymes were derived from different evolutionary origins. Green indicates likely acquisitions through cyanobacterial EGT, while dark red indicates likely acquisitions via α-proteobacterial EGT. Blue denotes potential acquisitions via HGT from other prokaryotes, though these genes might have been first transferred to the cyanobacteria progenitor prior to EGT (*green asterisks*). Purple indicates likely eukaryotic origin. Abbreviations: AAA, aromatic amino acid; ADT, arogenate dehydratase; CM, chorismate mutase; CMK, 4-(cytidine 5′-diphospho)-2-*C*-methyl-D-erythritol kinase; CS, chorismate synthase; DGD, DGDG synthase; DGDG, digalactosyldiacylglycerol; DHAP, dihydroxyacetone phosphate; DHQ-SDH, bifunctional dehydroquinate dehydratase-shikimate dehydrogenase enzyme; DHS, 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase; DQS, dehydroquinate synthase; DXR, 1-deoxy-D-xylulose 5-phosphate (DXP) reductoisomerase; DXS, DXP synthase; E4P, erythrose-4-phosphate; EPSPS, 5-enolpyruvylshikimate-3-phosphate synthase; ER, endoplasmic reticulum; F1,6P, fructose-1, 6-bisphosphate; F6P, fructose-6-phosphate; FBPase, fructose 1,6-bisphosphatase; G3P, glycerol-3-phosphate; Ga, billion years ago; GAP, glyceraldehyde 3-phosphate; GAPDH, GAP dehydrogenase; GPAT, glycerol 3-phosphate acyltransferase (known as ATS1 in *Arabidopsis*); HDR, 4-hydroxy-3-methyl-butenyl 1-diphosphate (HMBPP) reductase; HDS, HMBPP synthase; IPP, isopentenyl diphosphate; LPAT, lysophosphatidic acid acyltransferase; LUCA, last universal common ancestor; MDS, 2-*C*-methyl-D-erythritol 2,4-cyclodiphosphate synthase; MEP, 2-*C*-methyl-D-erythritol 4-phosphate; MGD, monogalactosyldiacylglycerol synthase; MVA, mevalonate; PA, phosphatidic acid; PAP, PA phosphatase; PEP, phosphoenolpyruvate; 2PGA, 2-phosphoglycerate; 3PGA, 3-phosphoglycerate; PGK, phosphoglycerate kinase; Phe, phenylalanine; PPA-AT, prephenate aminotransferase; Rubisco, ribulose-1, 5-bisphosphate carboxylase/oxygenase; S1,7BP, sedoheptulose 1,7-bisphosphate; SK, shikimate kinase; TAG, triacylglycerol; TPI, triose phosphate isomerase; Trp, tryptophan; Tyr, tyrosine; TyrA, arogenate dehydrogenase.

The oxidative pentose phosphate pathway is also broadly regarded as operating in the cytosol and plastid; the oxidative steps are duplicated in both locations, whereas some of the reductive steps are confined to the plastid (90). However, the reactions catalyzed by glucose-6-phosphate dehydrogenase, 6-phosphogluconolactonase, and 6-phosphogluconate dehydrogenase can also occur

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

in the peroxisome, most likely operating to provide NADPH for use in peroxisomal metabolism (112). The conditional targeting of these enzymes to the peroxisome is triggered by a cytosolic redox switch (112) and emphasizes the importance of cellular circumstance in the evolution of plant metabolism.

The TCA cycle and associated electron transport pathways, by contrast, are principally confined to mitochondria; however, several enzymes are additionally localized outside of mitochondria and encoded by a mosaic of organelle and nuclear genomes, with both aspects potentially mitigating against adverse redox conditions (167). Prior to the α-proteobacterial endosymbiotic event, the TCA cycle seemed to operate only as isolated steps in both the host and α-proteobacteria cells (18). The large number of TCA cycle–associated genes in plants is likely due to both polyploidization and HGT from eubacteria, with alternate localizations resulting from unfaithful targeting of the proteins (18, 150). Such genes were subsequently retained in the plant lineage since their redundancy likely provides robustness in the face of environmental adversity.

Glycolysis and the TCA cycle are remarkably similar between plants and other organisms but with a few exceptions (44). Glycolysis in plants has lost regulatory control by adenosine triphosphate (ATP) and acquired a number of bypass reactions relying on pyrophosphate as an alternative energy donor (133). Conversely, succinyl CoA ligase of the plant TCA cycle has lost the ability to utilize GTP and specifically uses ATP as a cosubstrate, while other organisms can use both GTP and ATP for this step (164), suggesting that a unique functional specialization occurred for this TCA cycle reaction in the plant lineage. Some cyanobacteria were thought to have incomplete TCA cycles but had instead acquired different enzymes for the conversion of 2-oxoglutarate to succinate (195). The greatest innovation with regard to respiratory pathways is that of the alternative respiratory pathways via NADH dehydrogenases and the alternative oxidase (AOX), which conveys robustness toward the prevailing environmental conditions (33, 44). It is important to note that, contrary to what was previously thought, AOX is by no means confined to plants, is found across the kingdoms of life, and is likely derived from α-proteobacterial EGT (109).

3.2. Galactolipid Biosynthesis

Galactolipids are present in the photosynthetic membranes of cyanobacteria, algae, and land plants and are synthesized in plants via both the endoplasmic reticulum and plastidic pathways (**Figure 3*b***), though their relative contribution is species dependent (159). The plastidic pathway is commonly referred to as the prokaryotic pathway, mainly because it produces galactolipids having C18 and C16 acyl chains at the *sn*-1 and -2 positions, respectively, like in cyanobacteria, whereas the endoplasmic reticulum–derived galactolipids have C18 in both positions. Despite such conventional naming, however, comparative genomics and detailed phylogenetic analyses from a wide range of photosynthetic and nonphotosynthetic organisms revealed that none of the genes and enzymes involved in the plastidic galactolipid biosynthetic pathways originated from the cyanobacterial EGT (143). For example, phosphatidic acid (PA) phosphatase (PAP) is likely of eukaryotic origin, while the initial acyltransferase reaction is mediated by nonhomologous enzymes in plants and cyanobacteria—glycerol 3-phosphate acyltransferase (GPAT) versus PlsX-PlsY enzymes, respectively (**Figure 3*b***). The final step, digalactosyldiacylglycerol (DGDG) synthesis, is catalyzed by different enzymes: DgdA in cyanobacteria and DGD1 in plants and most algae; however, the cyanobacterial-type DgdA is found in some red algae (64, 142), suggesting that the primary endosymbiont of Archaeplastida likely had both types of DGDG synthases. Therefore, even though photosynthetic membranes of both plastids and cyanobacteria have very similar galactolipid compositions, the underlying biosynthetic pathway likely underwent complete replacement after the endosymbiosis (**Figure 3*b***). Interestingly, many of these galactolipid biosynthetic

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

enzymes in cyanobacteria can be functionally replaced by plant-type enzymes (107, 192), suggesting that galactolipid biosynthetic pathways are flexible and may be relatively easy to engineer. Furthermore, the acquisitions of dual lipid biosynthetic pathways likely provided a robust foundation to support both the biogenesis of photosynthetic membranes and extraplastidic triacylglycerol (TAG) synthesis in algae and land plants.

3.3. Plastidic and Nonplastidic Isoprenoid Biosynthesis

Isoprenoid (terpenoid) biosynthesis supports tremendous chemical diversity of plant natural products and operates in two alternative pathways, the plastidic 2-*C*-methyl-D-erythritol 4-phosphate (MEP) and extraplastidic mevalonate (MVA) pathways (174). Isopentenyl diphosphate (IPP) is the precursor of diverse isoprenoid compounds, such as quinones, sterols, hormones, photosynthetic pigments, and other diverse terpenoid compounds. The MVA pathway is commonly found in animals, fungi, archaea, and gram-positive bacteria, while other bacteria, including cyanobacteria, typically have the MEP pathway; however, land plants and many algae have both pathways (92, 97, 108) and so do a few bacteria (7, 153). The MEP and MVA pathways evolved independently, as they use different enzymes—except for the final isomerization step—and start from different precursors: pyruvate and glyceraldehyde-3-phosphate for the MEP pathway and acetyl-CoA for the MVA pathway (174). The MVA pathway is likely an ancient pathway that potentially dates back to LUCA (66, 97). The MEP pathway is absent in all non-plastid-bearing eukaryotes and hence was likely introduced to the plant lineage through cyanobacterial EGT. Indeed, two reductases catalyzing the second and last steps are closely related to cyanobacterial enzymes (108) (**Figure 3*b***). Notably, however, other MEP pathway enzymes of plants and algae are more closely related to other eubacteria, such as α-proteobacteria and Chlamidia (108). Therefore, both EGT and HGT contributed to the mosaic evolutionary origin of the plastidic MEP pathway.

3.4. Plant Amino Acid Biosynthetic Pathways

Most amino acid biosynthetic pathways are also localized in the plastids, but many of these enzymes are derived from noncyanobacterial origins (139). While potential phylogenetic artifacts must be carefully evaluated (71), plants likely acquired some of them through HGT, either directly to the common ancestor of the Archaeplastida or to the cyanobacterial progenitor before EGT to the plant lineage (but they were lost or have not been found in extant cyanobacteria) (**Figure 3*a***). One striking example is seen in the aromatic amino acid (AAA) biosynthetic pathways, which produces L-phenylalanine (Phe), L-tyrosine, and L-tryptophan through a highly branched pathway of over twenty enzyme-catalyzed steps (102). These enzymes might be derived from at least six different sources (**Figure 3*b***). Plant shikimate kinases (SKs) are most closely related to cyanobacterial counterparts, while chorismate mutases (CMs) appear to be of eukaryotic origin (36, 140). Plant dehydroquinate dehydratases (DHQs) belong to type I, which are distinct from cyanobacterial type II DHQs, and are fused with the subsequent enzyme, shikimate dehydrogenase (SDH), into a single polypeptide (140) (**Figure 3*b***). The last two steps of plant Phe biosynthesis catalyzed by prephenate aminotransferase (PPA-AT) and arogenate dehydratase (ADT) are both most closely related to the Chlorobi/Bacteroidetes counterparts (36). Cyanobacterial PPA-ATs belong to class IV BCAA aminotransferases (57), rather than plant-type class II PPA-ATs (36). Therefore, plants use enzymes from diverse sources to operate AAA biosynthesis, which now supports the production of diverse and often abundant AAA-derived compounds (e.g., phenylpropanoids).

A prior phylogenomic study also revealed that one third of nuclear-encoded proteins specifically found in the genomes of plants and green algae, but not of nonphotosynthetic eukaryotes,

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

have no cyanobacterial homologs (80). Some biosynthetic enzymes of photosynthetic quinones (e.g., phylloquinone and tocopherols) were also replaced by some with noncyanobacterial origins (22, 180). Thus, plastidic metabolic pathways uniquely found in the plant kingdom were not simply acquired from cyanobacterial counterparts (**Figure 3*b***).

4. STEPWISE EMERGENCE OF PLANT-SPECIFIC METABOLISM

Besides the acquisition of primary metabolic pathways and enzymes from other kingdoms, additional metabolic pathways evolved uniquely within the plant kingdom (Plantae), which we refer to as plant-specific metabolism, as compared to more recently emerged specialized metabolism (see Section 6). The colonization of land by plants approximately 600 Ma was a pivotal event of the biological and ecological history of the Earth (**Figure 4*a***), which led to tremendous diversification of land plants (embryophytes) as well as other life such as animals (117). The adaptation to the terrestrial environment, however, came with enormous challenges (e.g., stresses imposed by UV radiation, dehydration, gravity, new pathogens, and herbivores) (**Figure 4*a***), which were overcome by a number of metabolic innovations. Recent genome sequencing and comparative genomic analyses of bryophytes and algae, especially charophyte green algae whose ancestor gave rise to land plants (10, 63, 76, 121), are now allowing us to examine the evolutionary history of the emergence of these plant-specific metabolic pathways.

Plant hormones orchestrate complex plant developmental processes by integrating environmental stimuli. For example, abscisic acid (ABA) is a critical phytohormone in desiccation responses. Some ABA biosynthetic and signaling pathway genes, homologous to land plant counterparts [e.g., abscisic aldehyde oxidase 3 (AAO3)], are present in charophyte algae, and so is the ABA compound itself (63) (**Figure 4*b***). Interestingly, soluble ABA receptors were likely acquired about 600 Ma through HGT from soil bacteria to an ancestor of Zygnematophyceae, the subaerial charophytes that are sister to land plants (21). However, the committed enzyme of ABA biosynthesis, 9-*cis*-epoxycarotenoid dioxygenase (NCED), is absent in charophytes but present in land plants (63, 121). Two enzymes involved in the major auxin biosynthetic pathway derived from tryptophan (via TAA1 and YUCCA enzymes) are present in all land plants, but their clear orthologs are absent in most green algae (63, 76, 121) (**Figure 4*b***). However, some auxin signaling and transport components (e.g., PIN and ARF) are found in charophytes (63, 76, 121), which also accumulate the primary auxin, indole-3-acetic acid (IAA) (63). Therefore, ABA and auxin might have been already produced by promiscuous enzymes or alternative noncanonical pathway(s) (e.g., tyramine-derived IAA pathway) prior to land plant evolution (**Figure 4*b***).

Charophyte algae appear to have some early enzymatic steps for the synthesis of a defense hormone, jasmonic acid (JA), such as allene oxide synthase (AOS), and hence can synthesize the key intermediate 12-oxophytodienoic acid (OPDA) (63) (**Figure 4*b***). Downstream enzymes such as OPDA reductase 3 (OPR3), however, are missing in green algae (63, 76) and even in the liverwort *Marchantia polymorpha* (10). JA response components [e.g., CORONATINE-INSENSITIVE PROTEIN 1 (COI1) and JASMONATE-ZIM DOMAIN (JAZ)] are present in the liverwort and other land plant genomes (10) but absent in charophyte algae (63, 76). Given that JA along with other oxylipins can be produced via nonenzymatic oxidation of polyunsaturated fatty acids (176), stepwise recruitments of JA biosynthetic enzymes likely resulted in the production of JA in a stereo-specific manner (**Figure 4*b***).

A recent comparative analysis of metabolic pathway genes across 72 genomes from green algae to angiosperms (17) revealed metabolic gains and losses during land plant evolution. A complete set of genes encoding enzymes of brassinosteroid biosynthesis and inactivation is present in seed plants (spermatophytes), while ferns and lycophytes have only partial sets and bryophytes

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

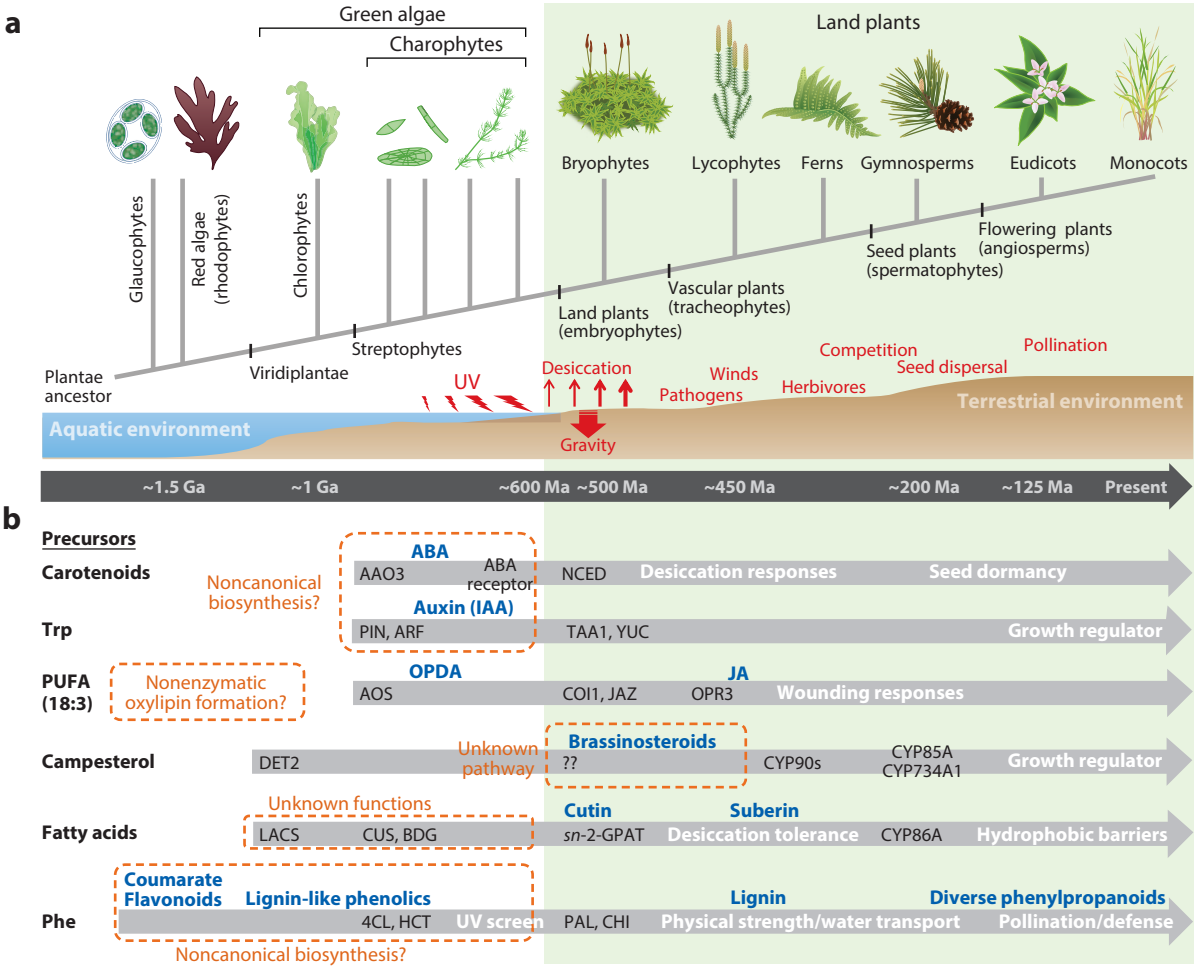


Figure 4

Stepwise evolution of plant-specific metabolism critical for land plant evolution. (*a*) Land plants (*green box*) evolved from an ancestor of charophyte algae around 600 Ma. A number of plant-specific metabolic pathways evolved during the gradual transition from aquatic to terrestrial environments and likely played critical roles in overcoming various challenges on land (*red*). Timeline is not drawn to scale. (*b*) Various metabolic innovations within Plantae are depicted with particular emphasis on time lags observed between chemical (*blue*) and biochemical (*black*) evidence. Such contradictions suggest that the earlier pathways were possibly catalyzed by analogous but nonhomologous enzymes or mediated by different pathways (*orange*). Major chemical functions are indicated in white letters on gray arrows. Abbreviations: 4CL, 4-coumarate CoA-ligase; AAO3, abscisic aldehyde oxidase 3; ABA, abscisic acid; AOS, allene oxide synthase; ARF, auxin response factor; BDG, BODYGUARD; CHI, chalcone isomerase; COI1, CORONATINE-INSENSITIVE PROTEIN 1; CUS, cutin synthase; DET2, DEETIOLATED2 steroid 5-α-reductase; Ga, billion years ago; HCT, hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyltransferase; IAA, indole-3-acetic acid; JA, jasmonic acid; JAZ, JASMONATE-ZIM DOMAIN; LACS, long chain acyl-CoA synthase; Ma, million years ago; NCED, 9-*cis*-epoxycarotenoid dioxygenase; OPDA, 12-oxophytodienoic acid; OPR3, OPDA reductase 3; PAL, phenylalanine ammonia lyase; Phe, phenylalanine; PIN, PIN-FORMED auxin efflux carriers; PUFA, polyunsaturated fatty acid; *sn*-2-GPAT, *sn*-2-specific glycerol 3-phosphate acyltransferase; TAA1, tryptophan aminotransferase of *Arabidopsis* 1; Trp, tryptophan; UV, ultraviolet; YUC, YUCCA flavin monooxygenase.

and green algae have only the initial pathway gene [i.e., DEETIOLATED2 steroid 5-α-reductase (*DET2*)] (**Figure 4*b***). Since brassinosteroids also occur in nonseed plants, albeit at low concentrations (190), these plants make brassinosteroids via an unknown pathway or pathways, and the canonical pathway found in seed plants today likely developed in a stepwise manner. Similarly, cytochrome P450 oxidases (e.g., CYP86As) involved in cutin and suberin production are absent in genomes of nonseed plants that nevertheless contain these biopolymers. Many green algae contain genes encoding long chain acyl-CoA synthetase (LACS) involved in fatty acid activation (17). Also, some charophytes have cutin synthase (CUS) and BODYGUARD (BDG) required for cuticle assembly (76), though their functions remain to be investigated. Algae, however, mostly lack other genes (17, 76), consistent with their lack of true cutin and suberin polyesters. Therefore, stepwise acquisition of additional enzymes, such as *sn-2*-specific GPAT and CYP86A enzymes, likely contributed to the sequential emergence of cutin in land plants and then suberin in vascular plants, which created robust hydrophobic barriers to withstand drier terrestrial habitats (187) (**Figure 4*b***).

Complex polysaccharide cell walls confer both biomechanical support and protection against environmental stress, and their evolution has often been cited as vital to the plant colonization of terrestrial habitats (76, 113). Land plant primary walls are composed of cellulose that is embedded within a matrix of pectin and hemicellulose together with glycoproteins. However, many of the polysaccharides extant in embryophyte cell walls evolved during the divergence of the charophyte algae (10, 162). The recent release of the *Penium margaritaceum* genome revealed considerable expansion of carbohydrate active enzyme gene families, including glucosyl hydrolases, carbohydrate esterases, and polysaccharide lyase class enzymes (76). Also, detailed phylogenetic analyses based on charophyte algae transcriptomes revealed that most of the core cell wall polysaccharide enzymes of land plants likely evolved before the emergence of land plants (113), consistent with the detection of cell walls similar to those of land plants in charophyte algae (162).

Phenylpropanoid compounds also played critical roles during land plant evolution: Soluble phenolic compounds (e.g., hydroxycinnamate esters and flavonoids) can absorb harmful UV radiation and have antimicrobial activities. Phenolic polymers, such as sporopollenin and lignin, provide physical barriers and support to withstand desiccation and other physical stresses imposed on land plants (**Figure 4*a***). Phenylalanine ammonia-lyase (PAL) catalyzes the committed enzyme of phenylpropanoid biosynthesis. Plant PAL orthologs are found in all land plants but absent in algae and are most similar to fungal PALs (32, 38). One of the core flavonoid biosynthetic enzymes, chalcone isomerase (CHI), also arose in land plants (20), intriguingly from a noncatalytic ancestor, a fatty acid binding protein (79). These rare recruitment events represent key metabolic innovations of land plants (**Figure 4*b***). Interestingly, however, some algae appear to have several genes encoding downstream phenylpropanoid [e.g., 4-coumarate CoA-ligase (4CL)], flavonoid, and monolignol [e.g., hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyltransferase (HCT)] pathway enzymes (32, 91). Also, many algae accumulate phloroglucinol, phenylpropanoid intermediates (e.g., *p*-coumarate), and trace levels of flavonoids (54). Wall-bound lignin-like phenolics are detectable in some red and green algae (34, 162). Therefore, algae likely have phenylpropanoid biosynthetic capacity via rudimentary and uncharacterized pathways, some of which were further modified, such as by the addition of the plant/fungi-type PALs, during the land plant evolution to efficiently produce diverse phenylpropanoid compounds (**Figure 4*b***).

A number of plant-specific metabolic pathways evolved during early land plant evolution, which played critical roles in terrestrialization (**Figure 4*a***) and also provided stepping-stones for the diversification of specialized metabolism (as discussed in Section 6). Accumulating evidence suggests that the ancestor of charophyte algae and land plants had already developed fundamental machineries of thought-to-be land plant–specific metabolic pathways (e.g., phytohormone and

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

phenylpropanoid biosynthesis) and hence preadapted to life on land (**Figure 4*b***). The lack of obvious orthologs of corresponding plant enzymes in charophyte algae suggests that these lineages use different enzymes and/or pathways from those of land plants (**Figure 4*b***). These earlier rudimentary pathways were later modified or replaced by the recruitment of additional genes and enzymes, resulting in the formation of more efficient metabolic pathways. Experimental characterizations of candidate enzymes from charophytes and diverse land plants are still largely pending and may further reveal earlier functions of these prototype enzymes, pathways, and compounds, which were later co-opted to overcome challenges in various terrestrial environments.

5. DIVERSIFICATION OF PRIMARY METABOLISM

Primary metabolism is generally assumed to be conserved, especially within the plant kingdom, which mainly relies on photosynthetic carbon fixation as the primary source of organic carbon and energy. This notion is true in relative comparison to highly diversified secondary or specialized metabolism. However, there are some relatively rare alterations in lineage-specific primary metabolism, which likely have significant impacts on the overall metabolic networks and physiology in specific plants (100).

5.1. C_3, C_4, and Crassulacean Acid Metabolism Photosynthesis

Arguably the best-studied pathway variations seen in plant primary metabolism are those of carbon assimilation. In fact, three different main modes of photosynthesis currently exist, namely C_3, C_4, and crassulacean acid metabolism (CAM) (149). Because anatomical evolution of CAM and C_4 from C_3 metabolism has been extensively covered elsewhere (62, 149), we largely focus on more metabolic aspects. Unlike 3-phosphoglycerate, the three-carbon molecule produced by Rubisco in C_3 photosynthesis, C_4 photosynthesis initially generates a four-carbon molecule, oxaloacetate, via the action of phospho*enol*pyruvate carboxylase (PEPC). Oxaloacetate is subsequently converted to malate or aspartate and shuttled to the chloroplasts of the bundle sheath cells, where the CO_2 is released for refixation by Rubisco. While at first glance this appears to be an inferior strategy compared to that of ancestral C_3 photosynthesis due to the high metabolic costs associated with its function, it provides an adaptive advantage under arid, warm, and high light conditions via the attenuation of the oxygenation side reaction of Rubisco, thereby substantially reducing photorespiration. The C_4 photosynthetic pathway has evolved more than 60 times independently yet unevenly across the phylogeny, being particularly prevalent in the Poaceae and Caryophyllales (141, 175). A battery of recent comparative genomics studies on C_3, C_4, and C_3–C_4 intermediates suggests that this repeated evolution was likely facilitated by preconditions or enabling traits that were either present or emerged within given plant lineages (5, 61, 149). One such precondition was the shuttling of photorespiratory glycine from the mesophyll to bundle sheath cells—the so-called C_2 cycle, which is present in many sister species to C_4 lineages and is accompanied by the shuttling of other metabolites such as alanine/pyruvate and aspartate/malate (56). Once these were established, alongside alterations in gene expression and anatomy (99), C_4 could evolve relatively easily and be maintained in habitats in which it proved advantageous.

CAM photosynthesis, which temporally separates CO_2 uptake and fixation as a means to conserve water, has also arisen independently at least 35 times (62) and may have allowed species to diversify in deserts. Also, like C_4, CAM appears to have arisen in a stepping-stone fashion, while genome sequencing suggests signatures of convergence in the protein sequence and diel transcript expression of genes involved in nocturnal CO_2 fixation and stomatal movement (188). Unlike the evolutionary trajectory of C_4, it appears that the final step (i.e., the transition from weak CAM to strong CAM) did not evolve easily (37). Researchers have argued, however, that the C_3-to-CAM

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

transition represents a true continuum (12), which, alongside recent modeling analysis (154), bodes well for attempts to improve plant productivity by converting C_3 species into CAM species.

Despite having quite distinctive trajectories, the evolution of C_4 and of CAM appear to have many commonalities. Both a focused study on the key enzyme PEPC (23) and a more extensive survey of 19 gene families (55) within the Caryophyllales indicate shared amino acid substitution patterns between the two modes, in addition to mode-specific substitutions. While most of the attention on evolution in photosynthesis has recently been paid to C_4 and CAM modes, relatively little work has focused on C_3 metabolism, which represents more than 90% of extant terrestrial plant species. However, a recent metabolomics analysis of Calvin-Benson cycle intermediates in panels of C_4 and C_3 species (4) revealed considerable differences between the metabolite contents of the species sets. The study additionally demonstrated striking differences among different C_3 species, likely reflecting their different evolutionary trajectories from one another. Thus, a deeper study of the evolution of primary metabolism within C_3 species is also warranted.

5.2. Lineage-Specific Amino Acid Biosynthetic Pathways

Some amino acid biosynthetic pathways have also diversified in specific plant lineages. Amino acids are essential for protein synthesis in all organisms but are also used as precursors for the biosynthesis of numerous plant specialized metabolites, such as alkaloids, cyanogenic glycosides, glucosinolates, and phenylpropanoids. Thus, lineage-specific alterations of amino acid biosynthesis are likely linked to different demands of certain amino acid precursor(s) for the synthesis of downstream specialized metabolism (100).

An interesting alteration in BCAA biosynthesis was found in the trichomes of wild and cultivated tomatoes, which produce various acylsugar defense compounds (41). IPMS catalyzes the committed step of leucine biosynthesis (**Figure 2**) and is typically feedback-inhibited by leucine at its C-terminal regulatory domain, which controls relative flux toward leucine and valine biosynthesis. The IPMS3 isoform specifically expressed in trichome cells was found to be truncated at its C-terminal and hence insensitive to the leucine feedback inhibition in cultivated tomatoes (*Solanum lycopersicum*) (120). In contrast, wild tomato (*Solanum pennellii*) had inactive IPMS3. As a result, *S. lycopersicum* and *S. pennellii* produce acylsugars derived from the leucine and valine pathways (with 2-methylpropanoic and 3-methylbutanoic acid acyl chains), respectively (120). Therefore, altered supply of the primary metabolic precursors contributed to the chemical diversity of specialized metabolism.

The final step of tyrosine biosynthesis, catalyzed by arogenate TyrA dehydrogenase, is also strictly feedback-inhibited by the product, tyrosine. The *TyrA* genes were tandemly duplicated in core Caryophyllales species, such as beets, which uniquely produce tyrosine-derived betalain pigments rather than the ubiquitous phenylalanine-derived anthocyanin pigments (13). One of the duplicated TyrA isoforms, whose expression correlates with betalain pigmentation, exhibits relaxed sensitivity to the tyrosine feedback inhibition (98). Interestingly, the deregulated TyrA enzymes emerged before the evolution of betalain pigmentation. Thus, the increased availability of the tyrosine precursor might have facilitated later evolution of the novel pigment pathway—betalain biosynthesis—in this specific plant lineage (98). Another example of deregulated TyrA enzymes was also found in legumes (146, 147), which are also associated with elevated production of specialized defense metabolites, such as tyrosine and gallate conjugates accumulated in some *Inga* species at up to 20% of dry weight (26).

These examples of primary metabolic diversification are mediated by (*a*) altering gene expression, (*b*) repurposing preexisting enzymes to form a different pathway (e.g., C_4 and CAM carbon fixation), or (*c*) altering the regulation of key metabolic enzymes. Although rare, these changes have major impacts on plant physiology (e.g., stress tolerance) and metabolism (e.g., evolution of

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

specialized metabolism). The rapidly increasing capacity of genome sequencing, gene synthesis, and high-throughput enzyme assays will likely facilitate discovery of novel variants in primary metabolic enzymes in the coming years, which could potentially provide useful tools to redesign often difficult-to-manipulate primary metabolic pathways (101).

6. EVOLUTION OF SPECIALIZED METABOLISM

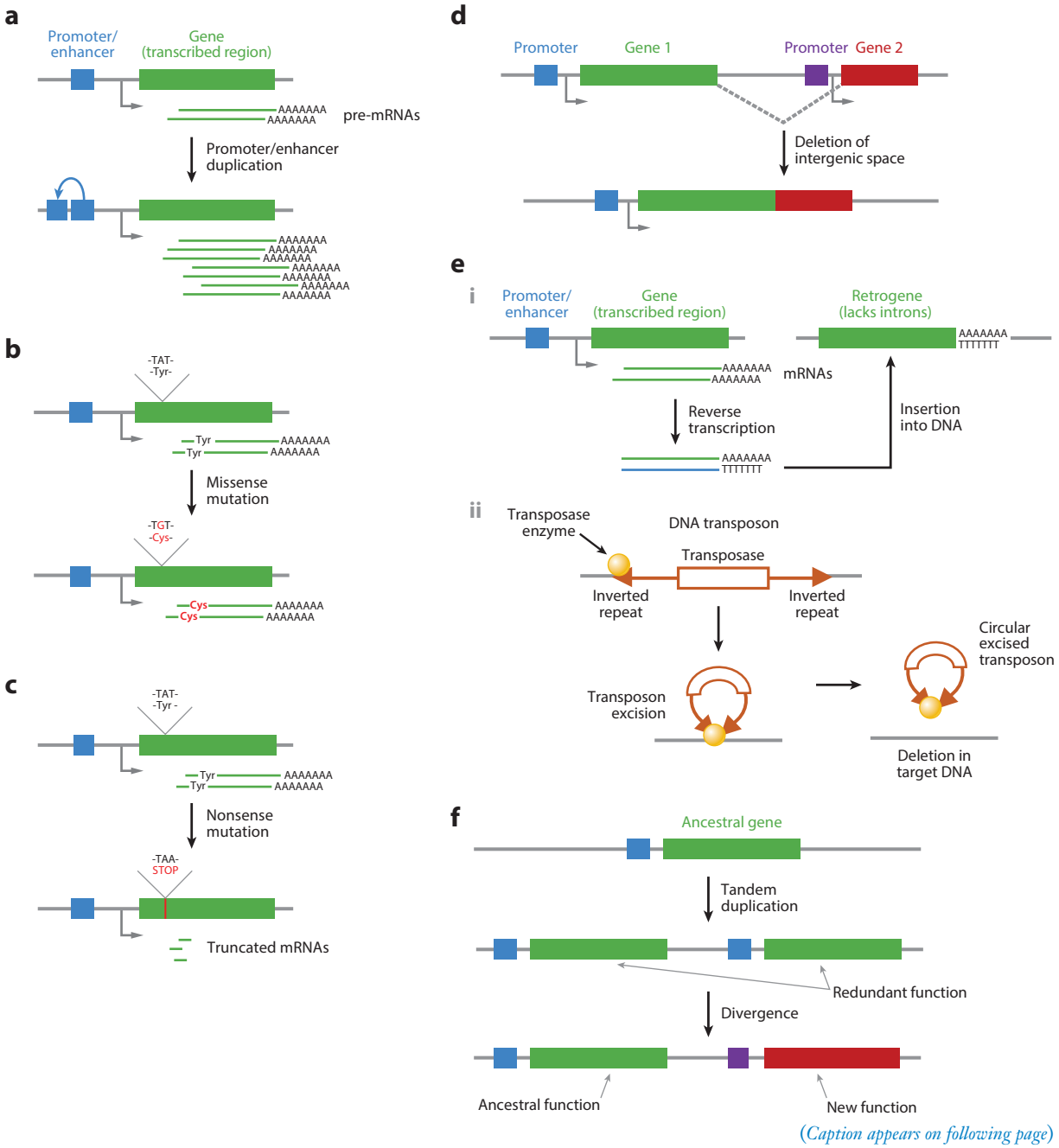
Having established a core metabolism, several factors have been identified that produce the vast metabolic diversity in various plant lineages, i.e., specialized (or secondary) metabolism. These include (*a*) differences in promoter strength resulting from differences in methylation or copy number variation in the promoter region; (*b*) single-nucleotide polymorphisms in the coding region corresponding to enzymatic activity, substrate preference, or both; (*c*) polymorphisms resulting in a premature stop codon; (*d*) gene fusions; (*e*) large gene deletions or insertions caused by transposons; and (*f*) tandem gene duplications (46) (**Figure 5**). Importantly, comparisons across the domains of life have revealed that gene duplication is particularly prevalent in plants (193), which then acts as the initial step for the co-opting or hijacking of core metabolic enzymes into specialized metabolism, leading to tremendous expansion of the plant chemical repertoire (89, 116, 160). To place this in context, primary metabolism is widely assumed to comprise around 1,000 metabolites, whereas the secondary metabolites of the plant kingdom have been estimated to number 200,000 to 1 million (42).

An early example of the importance of tandem duplication was the glucosinolate polymorphism in *Arabidopsis thaliana* being encoded by the AOP2/3 and MAM1/3 tandem gene duplication region (88). The key enzyme behind *Arabidopsis* accession-specific phenylacylated flavonoids was similarly found to be encoded within a serine carboxypeptidase-like (SCPL) tandem duplication (170). Such genes are thought to be the result of recent neofunctionalization following tandem gene duplication because they exhibit genetic polymorphism among natural accessions and are not conserved in any but the most closely related species (46). Plant secondary metabolism is more tolerant of mutations than its more evolutionarily constrained counterparts in primary metabolism (179), which partially explains the massive chemodiversity of plant specialized metabolites (114). Substrate specificity is functionally relevant only when alternative substrates are available at appropriate concentrations, meaning that evolutionary changes in what Schenk & Last (148) refer to as the "cellular context" are greatly important in shaping the metabolic diversity of a species. Such changes can include (*a*) alterations in cell-type specificity of expression, (*b*) subcellular relocalization, (*c*) pathway sequestration, and (*d*) the mixing of cellular contents following tissue damage. Indeed, enzymes and metabolic diversity coevolved step-by-step (122); therefore, tremendous expansion and diversification of some enzyme families (e.g., P450 oxygenase, acyltransferases, and *O*-methyltransferases) were facilitated by the availability of new metabolites, which promiscuous enzymes acted on and further specialized to use these new substrates.

Tandem gene duplication and neofunctionalization is a leitmotif recurring across the diversity of the phenylpropanoids (46), in both the establishment of major subclasses of flavonoids, such as isoflavone diversification in legumes (24), and decorative reactions, such as the flavone wogonin production in the medicinal plant *Scutellaria baicalensis* (198). More recently, phylogenetic and structure-function analyses revealed that the evolution of the phenolic rosmarinic acid occurred independently in Lamiaceae and Boraginaceae, highlighting that chemotypic convergence arose via disparate evolutionary trajectories (94). In alkaloid biosynthesis, an atypical polyketide synthesis and P450-mediated cyclization led to the innovation of the tropane alkaloids (6), while the molecular basis of the chlorination of alkaloids in Menispermaceae revealed an example of cross-kingdom parallel evolution (86). Moreover, new sources of the evolution of specialized

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

metabolic enzymes were reported recently. Plant triterpenoid metabolism co-opted a component of the cell wall biosynthetic machinery via recruitment of cellulose synthase-like enzymes alongside other enzymes of saponin biosynthesis to the endoplasmic reticulum (78). Also, the aromatization of secondary metabolites evolved via specialization of the detoxification enzymes glyoxalases (67).



(*Caption appears on following page*)

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

Figure 5 (*Figure appears on preceding page*)

Some of the major genetic mechanisms underlying the diversity of secondary metabolism and plant chemodiversity. (*a*) Enhanced gene expression is shown as a consequence of promoter duplication. A mutational event in regulatory sequences may underlie tissue-specific expression of metabolism-related genes and thus lead to the diversity of metabolic profiles between different tissues or organs. (*b*) A missense mutation (here, a nonsynonymous SNP) causes a change of gene function. (*c*) The introduction of a premature stop codon (a nonsense mutation) results in truncated gene product(s). (*d*) Deletion of the space between two different genes (making them contiguous) results in gene fusion. (*e*) Transposable elements mediate (*i*) insertion or (*ii*) deletion events. (*f*) Tandem (local) gene duplication and emergence of a new function in one of the paralogs occurs after divergence (neofunctionalization). TAT, TGT, and TAA represent nucleotide codons. Abbreviations: Cys, cysteine; mRNA, messenger RNA; pre-mRNA, precursor mRNA; SNP, single-nucleotide polymorphism; Tyr, tyrosine.

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

A review by Jacobowitz & Weng (73) highlights recent technologies that have been used in the study of the evolution of secondary metabolism, and several reviews cover the organization, and to a lesser extent evolution, of gene clusters in plant secondary metabolism (46, 126). Therefore, with three exceptions—ancestral protein resurrection, the 1,000 Plant Transcriptomes Project, and phylogenomics of genes derived from transposable elements—we will not retread this ground. Two articles from the Barkman lab (68, 69) highlight the power of the first exception. The first study investigates the role of ancestral functional variation in determining modern-day protein specificity by looking at protein functional changes in the salicylic acid/benzoic acid/theobromine (SABATH) lineage of plant secondary metabolite–producing enzymes. In each case, they demonstrated that ancestrally nonpreferred activities were improved upon in a daughter enzyme following gene duplication, suggesting that these functional shifts were likely coincident with positive selection (68). In the second study, they revealed that the convergent evolution of caffeine biosynthesis was the result of the co-option of exapted ancestral enzymes (69). Also of note here is an excellent recent review on the effect of epistasis and dominance on the evolution of the terpene gene synthase family (19). The 1,000 Plant Transcriptomes Project, is, as its name suggests, an impressive resource housing exome data on a thousand species, which provide good coverage of the green lineage (127). As the examples for primary metabolism described above attest, there will likely be a massive boon in understanding the evolution of the diversity of plant metabolism, especially in combination with advanced metabolomics. The final area that deserves mention is the role that transposable elements play in metabolic innovation. While several studies have already highlighted this fact (77, 119), the increasing interest and tractability of studying structural variations at the genome-wide level (see, for example, 3) are likely to allow a far deeper assessment of such events in the near future.

7. IMPACT OF DOMESTICATION ON PLANT METABOLISM

A specific variant of evolution that is of high relevance to humans is artificial selection—more specifically, the processes of domestication and crop improvement. Both processes are well documented to cause genetic bottlenecks and massively reduce allelic diversity (45). The consequence of this reduced diversity has been studied for decades for certain metabolic traits linked to agronomic traits; however, such studies have until recently been carried out on single metabolic traits that have profound effects on our foodstuffs. Examples include the century-long breeding for attractive color; reduced bitterness; and altered acidity, sweetness, and starchiness, as well as fragrances (15, 25, 53, 129, 197, 199). While these studies proved important in identifying the genetic loci or even the genes underlying metabolic changes occurring during domestication, it is arguably the combination of next-generation sequencing and metabolomics that has advanced our understanding of these metabolic changes the most.

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

Earlier studies that focused on the levels of phenolic compounds across a panel, which was set up to establish the domestication of the eggplant, revealed that agronomic features (e.g., fruit size and texture) aside from nutrition were prominently targeted via selection (111). Beleggia et al. (8) performed a relatively simple, yet revolutionary, evaluation of metabolic changes occurring during wheat domestication by investigating the composition of 51 central kernel metabolites in three *Triticum turgidum* L. subspecies (wild emmer, emmer, and durum wheat). The study found that the primary domestication (that of emmer) was marked by a reduction in unsaturated fatty acids, while a decrease in amino acid levels characterized the secondary domestication (that of durum wheat) (**Figure 6*a***). Importantly, these metabolic effects were partially independent of the associations that any of these metabolites have with other domestication-related traits (e.g., kernel weight). Moreover, the changes in metabolite contents were coupled to alterations in metabolite correlation networks, suggesting that a deep metabolic restructuring took place on domestication.

While the details differ, the patterns of change with domestication in tomato were similar in a large-scale multiomics study of 610 accessions (199). Selection of alleles of genes associated with larger fruits altered metabolite profiles as a consequence of linkage drag, while selection at five major loci reduced the accumulation of antinutritional steroidal glycoalkaloids in ripe fruits (**Figure 6*b***). In addition, breeding for pink fruits, an Asian preference, also modified the content of over 100 metabolites, and the introgression of resistance genes from wild relatives also caused unexpected disturbance to the metabolome.

In another study, the metabolic divergence between maize and its wild ancestor, teosinte, was assessed (186) (**Figure 6*c***). This study revealed that, as for other species, certain metabolite classes varied in specific evolutionary transitions. In this case, alkaloids, terpenoids, and lipids were targeted at the divergence between teosinte and tropical maize, whereas benzoxazinoids were targeted on the divergence between tropical and temperate maize. Some of the genetic loci responsible were determined by studies of either independent maize and teosinte populations or populations resulting from a maize-teosinte cross (95, 186).

The changes with domestication in rice as compared to maize were recently addressed in a comparative metabolomic study (35), which revealed that these species displayed differing metabolomic shifts during their evolution (**Figure 6*d***). Moreover, both shifts were different from those revealed in the early study in wheat. Domestication has additionally been extensively studied at the transcriptomic level and has been found to cause a decrease in nucleotide and expression diversity as well as to modify coexpression patterns in common bean (9) and, less prominently, in maize (165). Conversely, differences in the levels of metabolites in common bean (131), like those discussed in maize above, and the types of metabolites present in the cultivated crop are more divergent than in the wild progenitors. RNA sequencing was also used to provide insight into the evolution of lettuce and the regulation of flavonoid biosynthesis (194), though metabolic profiling has not yet been carried out at scale in these species.

While by no means comprehensive in terms of species yet studied, the massive changes that have been documented to occur during domestication at both qualitative and quantitative levels render the use of de novo domestication of wild species relatives of our major crops attractive as both a research tool and a means of engineering metabolically valuable crops (47). Intriguingly, analysis of the Gephebase database (**http://www.gephebase.org**), which compiles published data about the genes responsible for evolutionary and domesticated changes across eukaryotes, suggests that breeders have selected large-effect mutations underlying adaptive traits in specific settings but that these mutations and associated phenotypes would not survive the vagaries of changing environments (29). This observation aside, the database also highlights several other domestication genes of interest, including some associated with metabolism such as the *betaine aldehyde*

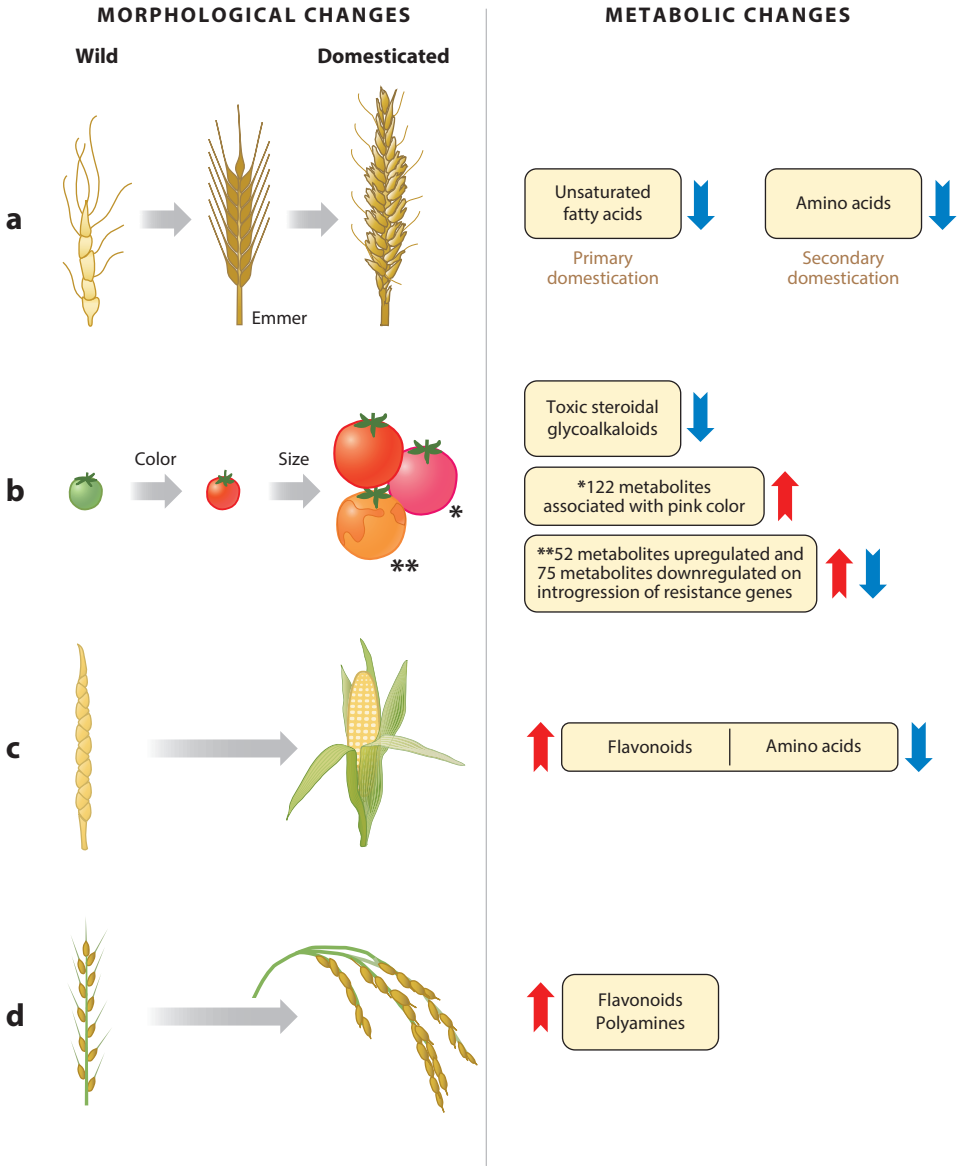


Figure 6

Metabolic changes concurrent with domestication and improvement of crops. Images on the left show morphological changes in the harvested organs of four crops: (*a*) wheat, (*b*) tomato, (*c*) maize, and (*d*) rice. Images on the right show the metabolic changes associated with (*a*) primary and secondary domestication in wheat, (*b*) increase in size, selection for color, and introgression of resistance genes in tomato, (*c*) the domestication of maize, and (*d*) the domestication of rice. Red and blue arrows indicate increase or decrease, respectively, in the metabolites shown. Data from References 8, 35, 95, and 199.

dehydrogenase 2 (*BADH2*) gene in rice—inactivation of which has been demonstrated to enrich aromatic properties (155).

Domestication (or, strictly speaking, the combination of domestication and crop improvement) is a special case of evolution, and often the changes that occurred during this process are quite

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

different from those during natural selection, which in plants was characterized by a massive expansion of metabolic diversity. Domestication is instead mainly characterized by quantitative changes in content, although the elimination of some bitter and toxic compounds has occurred in some instances. It is important to note that the domestication process has largely been driven by selection for yield, and, as such, quality is often compromised, as seen, for example, in the deterioration of tomato fruit taste (87). That said, widespread studies of the effects of domestication on the metabolome are in their infancy, and our understanding of these will dramatically increase in the near future.

8. CONCLUSIONS AND PERSPECTIVES

The evolutionary history of plant metabolism broadly reviewed here highlights a number of key metabolic innovations that led to the tremendous chemodiversity of plants (**Figure 1**). Metabolites and underlying biochemical pathways initially evolved under environments having certain primordial substrates and cofactors, which appear to still impose constraints on current metabolic networks. Although various CO_2 fixation, anabolic, and catabolic pathways exist in different organisms, the core metabolism is largely conserved across kingdoms, likely stemming from nonenzymatic reactions and common reaction modules that contributed to the formation of these core pathways (**Figure 2**). Mixing and matching of various enzymes and pathways derived from different evolutionary origins, along with multiple subcellular compartments in the plant cells, further led to the formation of the unique and robust framework of the plant primary metabolic network (**Figure 3**). Besides oxygenic photosynthesis acquired through cyanobacterial endosymbiosis, the duplication of some key pathways, such as glycolysis and isoprenoid biosynthesis, likely reinforced robust connections between photosynthetic carbon fixation, energy metabolism, and various biosynthetic pathways in plants. Such metabolic foundations of plants further enabled evolution and subsequent diversification of numerous defense compounds (e.g., phenylpropanoids), biopolymers (e.g., cell walls and cutin), and phytohormones, which collectively contributed to overcome the enormous challenges that these sessile organisms faced during colonization of land (**Figure 4**). These plant-specific pathways were further used as stepping-stones for evolutionary expansions of downstream specialized metabolism (e.g., production of diverse terpenoids and phenolics). The rapid expansion of the metabolic repertoire was driven by the relatively high rates of gene (and even whole-genome) duplication that characterize the plant kingdom (**Figure 5**). Such tremendous chemodiversity plays critical roles in plants to resist, defend, and flexibly adjust their growth and development in response to changing environments. Most recently, domestication has introduced quantitative changes in various nutritional compounds, though with limited effect on the metabolic repertoire by itself (**Figure 6**).

A deeper and holistic understanding of plant metabolic evolution will surely help inform metabolic engineering strategies to harness the immense potential of plant metabolic and chemical diversity. A rapidly growing number of plant genomes and transcriptome sequences (127) is accelerating the identification of genes and enzymes involved in specialized metabolism and expanding the repertoire of plant metabolic tool kits for synthetic biology in both microbial and plant hosts. However, some of these plant enzymes do not work efficiently in heterologous systems (52), suggesting that we still have to understand the biochemical and cellular contexts in which these specialized metabolic enzymes function and evolve in coordination (e.g., unknown interactors and certain microenvironments). As compared to recently evolved specialized metabolic enzymes and pathways, core metabolism that evolved much earlier may be more difficult to engineer. However, synthetic biology on plant hosts will require the modification of core pathways to improve the supply of primary metabolite precursors for the efficient production of downstream

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

plant natural products (156). Such major efforts will likely require repeated cycles of trial and error but should be facilitated through deeper understanding of their evolutionary path. This will be particularly critical in plants that have a much longer generation time and a more limited capacity of high-throughput screening than microbes. Some rare examples of evolutionary alterations of primary metabolism, for instance, can provide powerful tools and strategies to redesign primary metabolism (101). We may also have to restart from ancestral states and introduce modifications to negate the buildup of negative genetic epistasis, though we have to understand both intragenic and intergenic epistasis, and the latter is currently not easily tractable. To date, such studies have been carried out at the level of individual enzymes or short metabolic pathways; however, ultimately, we should be able to integrate the knowledge of evolutionary history with metabolic pathway modeling to redesign plant metabolism at the level of the complex network.

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

SUMMARY POINTS

1. Metabolism evolved under certain (ancient) environments, initially mediated via nonenzymatic reactions and limited numbers of ambiguous enzymes.
2. Plant primary metabolic pathways were formed through mixing and matching of enzymes from diverse origins, leading to the remarkable expansion of complex metabolic networks localized in different subcellular compartments.
3. A number of plant-specific metabolisms emerged in a stepwise manner, which played critical roles in plants' colonization of land and also served as stepping-stones for the diversification of specialized metabolism.
4. Although rare, some primary metabolic pathways have been recently altered in specific plant lineages, which has had significant impacts on both physiology and metabolism of plants.
5. The expansion of the specialized metabolic network occurred mostly in a stepwise manner, by massive gene duplication and neofunctionalization of the existing enzyme repository and the evolving cellular context.
6. Domestication has resulted in complex and unpredictable changes in metabolite contents, but has had largely quantitative changes and relatively little effect on the metabolic repertoire of our crops.

FUTURE ISSUES

1. To what extent is the promiscuity of core metabolic enzymes still maintained, and how much does it contribute to the functionality of the present metabolic network (e.g., underground metabolism)?
2. How much diversity exists in primary metabolic enzymes across the tree of life, and how flexibly can these enzymes from different origins be mixed and matched to modify the current metabolic networks?
3. What other cellular processes changed concomitantly or subsequently when a new pathway evolved?

4. What were the underlying selection pressures that acted on the evolution of certain metabolic pathways, if any (or are they the result of a simple genetic drift)?
5. What were the ancestral states under which a given metabolic pathway emerged? Were there any enabling events or preconditions that later facilitated the evolution of a new metabolic pathway?
6. Ancestral protein resurrection is an underused tool in plants but is useful to contextualize the evolutionary framework and to overcome the possibility of negative epistasis.
7. Can we uncouple agronomic and metabolic traits selected together during domestication, and improve metabolic traits (e.g., enhanced nutrition and defense compounds) without compromising agronomic traits (e.g., fruit sizes)?

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Due to space constraints we were unable to cite many relevant studies and apologize to authors whose work was not mentioned. Discussions with Saleh Alseekh, Federico Scossa, Koichiro Awai, and Kaan Koper were highly important for the writing of this article. The support of the Max Planck Society for the work on evolution in the Fernie group is gratefully acknowledged. The work of H.A.M. was supported by grants from the US National Science Foundation (IOS-1836824, MCB-1818040, and DEB-1938597), the US Department of Agriculture (NIFA-AFRI 2020-67013-30898), the US Department of Energy (BER DE-SC0020390), and the Alexander von Humboldt Research Fellowship.

LITERATURE CITED

1. Alam MT, Olin-Sandoval V, Stincone A, Keller MA, Zelezniak A, et al. 2017. The self-inhibitory nature of metabolic networks and its alleviation through compartmentalization. *Nat. Commun.* 8:16018
2. Alifano P, Fani R, Liò P, Lazcano A, Bazzicalupo M, et al. 1996. Histidine biosynthetic pathway and genes: structure, regulation, and evolution. *Microbiol. Rev.* 60(1):44–69
3. Alonge M, Wang X, Benoit M, Soyk S, Pereira L, et al. 2020. Major impacts of widespread structural variation on gene expression and crop improvement in tomato. *Cell* 182(1):145–161.e23
4. Arrivault S, Alexandre Moraes T, Obata T, Medeiros DB, Fernie AR, et al. 2019. Metabolite profiles reveal interspecific variation in operation of the Calvin–Benson cycle in both C_4 and C_3 plants. *J. Exp. Bot.* 70(6):1843–58
5. Aubry S, Kelly S, Kümpers BMC, Smith-Unna RD, Hibberd JM. 2014. Deep evolutionary comparison of gene expression identifies parallel recruitment of *trans*-factors in two independent origins of C_4 photosynthesis. *PLOS Genet.* 10(6):e1004365
6. Bedewitz MA, Jones AD, D'Auria JC, Barry CS. 2018. Tropinone synthesis via an atypical polyketide synthase and P450-mediated cyclization. *Nat. Commun.* 9(1):5281
7. Begley M, Gahan CGM, Kollas A-K, Hintz M, Hill C, et al. 2004. The interplay between classical and alternative isoprenoid biosynthesis controls γδ T cell bioactivity of *Listeria monocytogenes*. *FEBS Lett.* 561(1–3):99–104

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

8. Beleggia R, Rau D, Laidò G, Platani C, Nigro F, et al. 2016. Evolutionary metabolomics reveals domestication-associated changes in tetraploid wheat kernels. *Mol. Biol. Evol.* 33(7):1740–53
9. Bellucci E, Bitocchi E, Ferrarini A, Benazzo A, Biagetti E, et al. 2014. Decreased nucleotide and expression diversity and modified coexpression patterns characterize domestication in the common bean. *Plant Cell* 26(5):1901–12
10. Bowman JL, Kohchi T, Yamato KT, Jenkins J, Shu S, et al. 2017. Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell* 171(2):287–304.e15
11. Braakman R. 2013. Mapping metabolism onto the prebiotic organic chemistry of hydrothermal vents. *PNAS* 110(33):13236–37
12. Bräutigam A, Schlüter U, Eisenhut M, Gowik U. 2017. On the evolutionary origin of CAM photosynthesis. *Plant Physiol.* 174(2):473–77
13. Brockington SF, Walker RH, Glover BJ, Soltis PS, Soltis DE. 2011. Complex pigment evolution in the Caryophyllales. *New Phytol.* 190(4):854–64
14. Buchanan BB, Arnon DI. 1990. A reverse KREBS cycle in photosynthesis: consensus at last. *Photosyn. Res.* 24:47–53
15. Butelli E, Licciardello C, Ramadugu C, Durand-Hulak M, Celant A, et al. 2019. *Noemi* controls production of flavonoid pigments and fruit acidity and illustrates the domestication routes of modern citrus varieties. *Curr. Biol.* 29(1):158–164.e2
16. Caetano-Anollés G, Yafremava LS, Gee H, Caetano-Anollés D, Kim HS, Mittenthal JE. 2009. The origin and evolution of modern metabolism. *Int. J. Biochem. Cell Biol.* 41(2):285–97
17. **Cannell N, Emms DM, Hetherington AJ, MacKay J, Kelly S, et al. 2020. Multiple metabolic innovations and losses are associated with major transitions in land plant evolution. *Curr. Biol.* 30(10):1783–1800.e11**
18. Cavalcanti JHF, Esteves-Ferreira AA, Quinhones CGS, Pereira-Lima IA, Nunes-Nesi A, et al. 2014. Evolution and functional implications of the tricarboxylic acid cycle as revealed by phylogenetic analysis. *Genome Biol. Evol.* 6(10):2830–48
19. Cheema J, Faraldos JA, O'Maille PE. 2017. REVIEW: Epistasis and dominance in the emergence of catalytic function as exemplified by the evolution of plant terpene synthases. *Plant Sci.* 255:29–38
20. Cheng A-X, Zhang X, Han X-J, Zhang Y-Y, Gao S, et al. 2018. Identification of chalcone isomerase in the basal land plants reveals an ancient evolution of enzymatic cyclization activity for synthesis of flavonoids. *New Phytol.* 217(2):909–24
21. Cheng S, Xian W, Fu Y, Marin B, Keller J, et al. 2019. Genomes of subaerial Zygnematophyceae provide insights into land plant evolution. *Cell* 179(5):1057–1067.e14
22. Cheng ZG, Sattler S, Maeda H, Sakuragi Y, Bryant DA, DellaPenna D. 2003. Highly divergent methyltransferases catalyze a conserved reaction in tocopherol and plastoquinone synthesis in cyanobacteria and photosynthetic eukaryotes. *Plant Cell* 15(10):2343–56
23. Christin P-A, Arakaki M, Osborne CP, Bräutigam A, Sage RF, et al. 2014. Shared origins of a key enzyme during the evolution of C_4 and CAM metabolism. *J. Exp. Bot.* 65(13):3609–21
24. Chu S, Wang J, Cheng H, Yang Q, Yu D. 2014. Evolutionary study of the isoflavonoid pathway based on multiple copies analysis in soybean. *BMC Genet.* 15:76
25. Cohen S, Itkin M, Yeselson Y, Tzuri G, Portnoy V, et al. 2014. The *PH* gene determines fruit acidity and contributes to the evolution of sweet melons. *Nat. Commun.* 5:4026
26. Coley PD, Endara M-J, Ghabash G, Kidner CA, Nicholls JA, et al. 2019. Macroevolutionary patterns in overexpression of tyrosine: an anti-herbivore defence in a speciose tropical tree genus, *Inga* (Fabaceae). *J. Ecol.* 107(4):1620–32
27. Cooper G, Reed C, Nguyen D, Carter M, Wang Y. 2011. Detection and formation scenario of citric acid, pyruvic acid, and other possible metabolism precursors in carbonaceous meteorites. *PNAS* 108(34):14015–20
28. Court SJ, Waclaw B, Allen RJ. 2015. Lower glycolysis carries a higher flux than any biochemically possible alternative. *Nat. Commun.* 6:8427
29. Courtier-Orgogozo V, Martin A. 2020. The coding loci of evolution and domestication: current knowledge and implications for bio-inspired genome editing. *J. Exp. Biol.* 223:jeb208934

17. This innovative study used a comparative computational approach to identify metabolic gains and losses during land plant evolution.

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

30. Cronin JR, Moore CB. 1971. Amino acid analyses of the murchison, murray, and allende carbonaceous chondrites. *Science* 172(3990):1327–29
31. D'Ari R, Casadesús J. 1998. Underground metabolism. *Bioessays* 20(2):181–86
32. de Vries J, de Vries S, Slamovits CH, Rose LE, Archibald JM. 2017. How embryophytic is the biosynthesis of phenylpropanoids and their derivatives in streptophyte algae? *Plant Cell Physiol.* 58(5):934–45
33. Del-Saz NF, Ribas-Carbo M, McDonald AE, Lambers H, Fernie AR, Florez-Sarasa I. 2018. An *in vivo* perspective of the role(s) of the alternative oxidase pathway. *Trends Plant Sci.* 23(3):206–19
34. Delwiche CF, Graham LE, Thomson N. 1989. Lignin-like compounds and sporopollenin *Coleochaete*, an algal model for land plant ancestry. *Science* 245(4916):399–401
35. Deng M, Zhang X, Luo J, Liu H, Wen W, et al. 2020. Metabolomics analysis reveals differences in evolution between maize and rice. *Plant J.* 103(5):1710–22
36. Dornfeld C, Weisberg AJ, Ritesh KC, Dudareva N, Jelesko JG, Maeda HA. 2014. Phylobiochemical characterization of class-Ib aspartate/prephenate aminotransferases reveals evolution of the plant arogenate phenylalanine pathway. *Plant Cell* 26(7):3101–14
37. Edwards EJ. 2019. Evolutionary trajectories, accessibility and other metaphors: the case of C_4 and CAM photosynthesis. *New Phytol.* 223(4):1742–55
38. Emiliani G, Fondi M, Fani R, Gribaldo S. 2009. A horizontal gene transfer at the origin of phenylpropanoid metabolism: a key adaptation of plants to land. *Biol. Direct.* 4:7
39. Erb TJ, Zarzycki J. 2018. A short history of RubisCO: the rise and fall (?) of Nature's predominant CO_2 fixing enzyme. *Curr. Opin. Biotechnol.* 49:100–7
40. Evran S, Telefoncu A, Sterner R. 2012. Directed evolution of $(\beta\alpha)_8$-barrel enzymes: establishing phosphoribosylanthranilate isomerisation activity on the scaffold of the tryptophan synthase α-subunit. *Protein Eng. Des. Sel.* 25(6):285–93
41. Fan P, Leong BJ, Last RL. 2019. Tip of the trichome: evolution of acylsugar metabolic diversity in Solanaceae. *Curr. Opin. Plant Biol.* 49:8–16
42. Fang C, Fernie AR, Luo J. 2019. Exploring the diversity of plant metabolism. *Trends Plant Sci.* 24(1):83–98
43. Fani R, Liò P, Lazcano A. 1995. Molecular evolution of the histidine biosynthetic pathway. *J. Mol. Evol.* 41(6):760–74
44. Fernie AR, Carrari F, Sweetlove LJ. 2004. Respiratory metabolism: glycolysis, the TCA cycle and mitochondrial electron transport. *Curr. Opin. Plant Biol.* 7(3):254–61
45. Fernie AR, Tadmor Y, Zamir D. 2006. Natural genetic variation for improving crop quality. *Curr. Opin. Plant Biol.* 9(2):196–202
46. Fernie AR, Tohge T. 2017. The genetics of plant metabolism. *Annu. Rev. Genet.* 51:287–310
47. Fernie AR, Yan J. 2019. De novo domestication: an alternative route toward new crops for the future. *Mol. Plant* 12(5):615–31
48. Fischer WW, Hemp J, Johnson JE. 2016. Evolution of oxygenic photosynthesis. *Annu. Rev. Earth Planet. Sci.* 44:647–83
49. Flamholz A, Noor E, Bar-Even A, Liebermeister W, Milo R. 2013. Glycolytic strategy as a tradeoff between energy yield and protein cost. *PNAS* 110(24):10039–44
50. Fondi M, Emiliani G, Fani R. 2009. Origin and evolution of operons and metabolic pathways. *Res. Microbiol.* 160(7):502–12
51. Fuchs G. 2011. Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? *Annu. Rev. Microbiol.* 65:631–58
52. Galanie S, Thodey K, Trenchard IJ, Filsinger Interrante M, Smolke CD. 2015. Complete biosynthesis of opioids in yeast. *Science* 349(6252):1095–100
53. Gao L, Gonda I, Sun H, Ma Q, Bao K, et al. 2019. The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nat. Genet.* 51(6):1044–51
54. Goiris K, Muylaert K, Voorspoels S, Noten B, De Paepe D, et al. 2014. Detection of flavonoids in microalgae from different evolutionary lineages. *J. Phycol.* 50(3):483–92
55. Goolsby EW, Moore AJ, Hancock LP, De Vos JM, Edwards EJ. 2018. Molecular evolution of key metabolic genes during transitions to C_4 and CAM photosynthesis. *Am. J. Bot.* 105(3):602–13

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

56. Gowik U, Bräutigam A, Weber KL, Weber APM, Westhoff P. 2011. Evolution of C_4 photosynthesis in the genus *Flaveria*: How many and which genes does it take to make C_4? *Plant Cell* 23(6):2087–105
57. Graindorge M, Giustini C, Kraut A, Moyet L, Curien G, Matringe M. 2014. Three different classes of aminotransferases evolved prephenate aminotransferase functionality in arogenate-competent microorganisms. *J. Biol. Chem.* 289(6):3198–208
58. Granick S. 1957. Speculations on the origins and evolution of photosynthesis. *Ann. N. Y. Acad. Sci.* 69(2):292–308
59. Gross W, Lenze D, Nowitzki U, Weiske J, Schnarrenberger C. 1999. Characterization, cloning, and evolutionary history of the chloroplast and cytosolic class I aldolases of the red alga *Galdieria sulphuraria*. *Gene* 230(1):7–14
60. He Y, Mawhinney TP, Preuss ML, Schroeder AC, Chen B, et al. 2009. A redox-active isopropylmalate dehydrogenase functions in the biosynthesis of glucosinolates and leucine in Arabidopsis. *Plant J.* 60(4):679–90
61. Heckmann D. 2016. C_4 photosynthesis evolution: the conditional Mt. Fuji. *Curr. Opin. Plant Biol.* 31:149–54
62. Heyduk K, McKain MR, Lalani F, Leebens-Mack J. 2016. Evolution of a CAM anatomy predates the origins of Crassulacean acid metabolism in the Agavoideae (Asparagaceae). *Mol. Phylogenet. Evol.* 105:102–13
63. Hori K, Maruyama F, Fujisawa T, Togashi T, Yamamoto N, et al. 2014. *Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation. *Nat. Commun.* 5:3978
64. Hori K, Nobusawa T, Watanabe T, Madoka Y, Suzuki H, et al. 2016. Tangled evolutionary processes with commonality and diversity in plastidial glycolipid synthesis in photosynthetic organisms. *Biochim. Biophys. Acta. Mol. Cell Bio. Lipids* 1861(9 Part B):1294–308
65. Horowitz NH. 1945. On the evolution of biochemical syntheses. *PNAS* 31(6):153–57
66. Hoshino Y, Gaucher EA. 2018. On the origin of isoprenoid biosynthesis. *Mol. Biol. Evol.* 35(9):2185–97
67. Huang J-Q, Fang X, Tian X, Chen P, Lin J-L, et al. 2020. Aromatization of natural products by a specialized detoxification enzyme. *Nat. Chem. Biol.* 16(3):250–56
68. Huang R, Hippauf F, Rohrbeck D, Haustein M, Wenke K, et al. 2012. Enzyme functional evolution through improved catalysis of ancestrally nonpreferred substrates. *PNAS* 109(8):2966–71
69. Huang R, O'Donnell AJ, Barboline JJ, Barkman TJ. 2016. Convergent evolution of caffeine in plants by co-option of exapted ancestral enzymes. *PNAS* 113(38):10613–18
70. Hudson AO, Singh BK, Leustek T, Gilvarg C. 2006. An LL-diaminopimelate aminotransferase defines a novel variant of the lysine biosynthesis pathway in plants. *Plant Physiol.* 140(1):292–301
71. Husnik F, McCutcheon JP. 2018. Functional horizontal gene transfer from bacteria to eukaryotes. *Nat. Rev. Microbiol.* 16(2):67–79
72. Ingle RA. 2011. Histidine biosynthesis. *Arabidopsis Book* 9:e0141
73. Jacobowitz JR, Weng J-K. 2020. Exploring uncharted territories of plant specialized metabolism in the postgenomic era. *Annu. Rev. Plant Biol.* 71:631–58
74. Jensen RA. 1976. Enzyme recruitment in evolution of new function. *Annu. Rev. Microbiol.* 30:409–25
75. Jensen RA, Gu W. 1996. Evolutionary recruitment of biochemically specialized subdivisions of Family I within the protein superfamily of aminotransferases. *J. Bacteriol.* 178(8):2161–71
76. Jiao C, Sørensen I, Sun X, Sun H, Behar H, et al. 2020. The *Penium margaritaceum* genome: hallmarks of the origins of land plants. *Cell* 181(5):1097–111.e12
77. Joly-Lopez Z, Bureau TE. 2018. Exaptation of transposable element coding sequences. *Curr. Opin. Genet. Dev.* 49:34–42
78. Jozwiak A, Sonawane PD, Panda S, Garagounis C, Papadopoulou KK, et al. 2020. Plant terpenoid metabolism co-opts a component of the cell wall biosynthesis machinery. *Nat. Chem. Biol.* 16(7):740–48
79. Kaltenbach M, Burke JR, Dindo M, Pabis A, Munsberg FS, et al. 2018. Evolution of chalcone isomerase from a noncatalytic ancestor. *Nat. Chem. Biol.* 14(6):548–55
80. Karpowicz SJ, Prochnik SE, Grossman AR, Merchant SS. 2011. The GreenCut2 resource, a phylogenomically derived inventory of proteins specific to the plant lineage. *J. Biol. Chem.* 286(24):21427–39

63. The first report on the charophyte algae genome revealed that many of the thought-to-be land-plant-specific metabolic pathways had emerged already from these algae in a stepwise manner.

74. This early review article provides a fundamental framework for understanding how the metabolic pathway evolved through gene duplication and functionalization.

78. This study discovered that the committed glycosyltransferase enzyme of saponin specialized metabolism was recruited from an enzyme related to classical cellulose synthase (CesAs).

79. This study used ancestral protein reconstruction and revealed that the core phenylpropanoid enzyme chalcone isomerase arose from a noncatalytic ancestor, suggesting that enzymatic innovation played a key role in the evolutionary expansion of flavonoid metabolism in plants.

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

81. Keefe AD, Lazcano A, Miller SL. 1995. Evolution of the biosynthesis of the branched-chain amino acids. *Orig. Life Evol. Biosph.* 25(1–3):99–110
82. Keller MA, Kampjut D, Harrison SA, Ralser M. 2017. Sulfate radicals enable a non-enzymatic Krebs cycle precursor. *Nat Ecol Evol.* 1(4):83
83. Keller MA, Zylstra A, Castro C, Turchyn AV, Griffin JL, Ralser M. 2016. Conditional iron and pH-dependent activity of a non-enzymatic glycolysis and pentose phosphate pathway. *Sci. Adv.* 2(1):e1501235
84. Kelley DS, Karson JA, Blackman DK, Früh-Green GL, Butterfield DA, et al. 2001. An off-axis hydrothermal vent field near the Mid-Atlantic Ridge at 30° N. *Nature* 412(6843):145–49
85. Khersonsky O, Tawfik DS. 2010. Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu. Rev. Biochem.* 79:471–505
86. Kim CY, Mitchell AJ, Glinkerman CM, Li F-S, Pluskal T, Weng J-K. 2020. The chloroalkaloid (−)-acutumine is biosynthesized via a Fe(II)- and 2-oxoglutarate-dependent halogenase in Menispermaceae plants. *Nat. Commun.* 11(1):1867
87. Klee HJ, Tieman DM. 2018. The genetics of fruit flavour preferences. *Nat. Rev. Genet.* 19(6):347–56
88. Kliebenstein DJ, Lambrix VM, Reichelt M, Gershenzon J, Mitchell-Olds T. 2001. Gene duplication in the diversification of secondary metabolism: tandem 2-oxoglutarate–dependent dioxygenases control glucosinolate biosynthesis in Arabidopsis. *Plant Cell* 13(3):681–93
89. Kliebenstein DJ, Osbourn A. 2012. Making new molecules—evolution of pathways for novel metabolites in plants. *Curr. Opin. Plant Biol.* 15(4):415–23
90. Kruger NJ, von Schaewen A. 2003. The oxidative pentose phosphate pathway: structure and organisation. *Curr. Opin. Plant Biol.* 6(3):236–46
91. Labeeuw L, Martone PT, Boucher Y, Case RJ. 2015. Ancient origin of the biosynthesis of lignin precursors. *Biol. Direct.* 10:23
92. Lange BM, Rujan T, Martin W, Croteau R. 2000. Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. *PNAS* 97(24):13172–77
93. Lazcano A, Miller SL. 1999. On the origin of metabolic pathways. *J. Mol. Evol.* 49(4):424–31
94. Levsh O, Pluskal T, Carballo V, Mitchell AJ, Weng J-K. 2019. Independent evolution of rosmarinic acid biosynthesis in two sister families under the Lamiids clade of flowering plants. *J. Biol. Chem.* 294(42):15193–205
95. Li K, Wen W, Alseekh S, Yang X, Guo H, et al. 2019. Large-scale metabolite quantitative trait locus analysis provides new insights for high-quality maize improvement. *Plant J.* 99(2):216–30
96. Ljungdahl L, Irion E, Wood HG. 1965. Total synthesis of acetate from CO_2. I. Co-methylcobyric acid and CO-(methyl)-5-methoxybenzimidazolylcobamide as intermediates with *Clostridium thermoaceticum*. *Biochemistry* 4(12):2771–80
97. Lombard J, Moreira D. 2011. Origins and early evolution of the mevalonate pathway of isoprenoid biosynthesis in the three domains of life. *Mol. Biol. Evol.* 28(1):87–99
98. Lopez-Nieves S, Yang Y, Timoneda A, Wang M, Feng T, et al. 2018. Relaxation of tyrosine pathway regulation underlies the evolution of betalain pigmentation in Caryophyllales. *New Phytol.* 217(2):896–908
99. Lundgren MR, Osborne CP, Christin P-A. 2014. Deconstructing Kranz anatomy to understand C_4 evolution. *J. Exp. Bot.* 65(13):3357–69
100. Maeda HA. 2019. Evolutionary diversification of primary metabolism and its contribution to plant chemical diversity. *Front. Plant Sci.* 10:881
101. Maeda HA. 2019. Harnessing evolutionary diversification of primary metabolism for plant synthetic biology. *J. Biol. Chem.* 294(45):16549–66
102. Maeda HA, Dudareva N. 2012. The shikimate pathway and aromatic amino acid biosynthesis in plants. *Annu. Rev. Plant Biol.* 63:73–105
103. Martin W, Herrmann RG. 1998. Gene transfer from organelles to the nucleus: how much, what happens, and why? *Plant Physiol.* 118(1):9–17
104. Martin WF, Bryant DA, Beatty JT. 2018. A physiological perspective on the origin and evolution of photosynthesis. *FEMS Microbiol. Rev.* 42(2):205–31

83. This study demonstrated strong pH- and iron-dependency of nonenzymatic glycolysis and pentose phosphate pathway–like reactions, highlighting possible nonenzymatic precursor activities of cellular carbon metabolic networks.

95. Analysis of the primary metabolome of a cross between teosinte and maize that provides insight into metabolic changes underlying arguably the best-studied plant domestication process.

98. Discovered lineage-specific alteration of tyrosine biosynthetic TyrA enzymes, which likely increased the availability of the tyrosine precursor and later facilitated the evolution of betalain biosynthesis.

105. Martin WF, Cerff R. 2017. Physiology, phylogeny, early evolution, and GAPDH. *Protoplasma* 254(5):1823–34
106. Martin WF, Garg S, Zimorski V. 2015. Endosymbiotic theories for eukaryote origin. *Philos. Trans. R. Soc. B* 370(1678):20140330
107. Matsumoto T, Awai K. 2020. Adaptations in chloroplast membrane lipid synthesis from synthesis in ancestral cyanobacterial endosymbionts. *Biochem. Biophys. Res. Commun.* 528(3):473–77
108. Matsuzaki M, Kuroiwa H, Kuroiwa T, Kita K, Nozaki H. 2008. A cryptic algal group unveiled: a plastid biosynthesis pathway in the oyster parasite *Perkinsus marinus*. *Mol. Biol. Evol.* 25(6):1167–79
109. McDonald AE, Vanlerberghe GC. 2006. Origins, evolutionary history, and taxonomic distribution of alternative oxidase and plastoquinol terminal oxidase. *Comp. Biochem. Physiol. Part D Genom. Proteom.* 1(3):357–64
110. Messner CB, Driscoll PC, Piedrafita G, De Volder MFL, Ralser M. 2017. Nonenzymatic gluconeogenesis-like formation of fructose 1,6-bisphosphate in ice. *PNAS* 114(28):7403–7
111. Meyer RS, Whitaker BD, Little DP, Wu S-B, Kennelly EJ, et al. 2015. Parallel reductions in phenolic constituents resulting from the domestication of eggplant. *Phytochemistry* 115:194–206
112. Meyer T, Hölscher C, Schwöppe C, von Schaewen A. 2011. Alternative targeting of Arabidopsis plastidic glucose-6-phosphate dehydrogenase G6PD1 involves cysteine-dependent interaction with G6PD4 in the cytosol. *Plant J.* 66(5):745–58
113. Mikkelsen MD, Harholt J, Ulvskov P, Johansen IE, Fangel JU, et al. 2014. Evidence for land plant cell wall biosynthetic mechanisms in charophyte green algae. *Ann. Bot.* 114(6):1217–36
114. Milo R, Last RL. 2012. Achieving diversity in the face of constraints: lessons from metabolism. *Science* 336(6089):1663–67
115. Moghe GD, Last RL. 2015. Something old, something new: conserved enzymes and the evolution of novelty in plant specialized metabolism. *Plant Physiol.* 169(3):1512–23
116. Moghe GD, Leong BJ, Hurney SM, Jones AD, Last RL. 2017. Evolutionary routes to biochemical innovation revealed by integrative analysis of a plant-defense related specialized metabolic pathway. *eLife* 6:e28468
117. Morris JL, Puttick MN, Clark JW, Edwards D, Kenrick P, et al. 2018. The timescale of early land plant evolution. *PNAS* 115(10):E2274–83
118. Muchowska KB, Varma SJ, Chevallot-Beroux E, Lethuillier-Karl L, Li G, Moran J. 2017. Metals promote sequences of the reverse Krebs cycle. *Nat. Ecol. Evol.* 1(11):1716–21
119. Murray AW. 2020. Can gene-inactivating mutations lead to evolutionary novelty? *Curr. Biol.* 30(10):R465–71
120. Ning J, Moghe GD, Leong B, Kim J, Ofner I, et al. 2015. A feedback-insensitive isopropylmalate synthase affects acylsugar composition in cultivated and wild tomato. *Plant. Physiol.* 169(3):1821–35
121. Nishiyama T, Sakayama H, de Vries J, Buschmann H, Saint-Marcoux D, et al. 2018. The *Chara* genome: secondary complexity and implications for plant terrestrialization. *Cell* 174(2):448–64.e24
122. Noda-Garcia L, Liebermeister W, Tawfik DS. 2018. Metabolite-enzyme coevolution: from single enzymes to metabolic pathways and networks. *Annu. Rev. Biochem.* 87:187–216
123. Notebaart RA, Szappanos B, Kintses B, Pál F, Györkei Á, et al. 2014. Network-level architecture and the evolutionary potential of underground metabolism. *PNAS* 111(32):11762–67
124. Novikov Y, Copley SD. 2013. Reactivity landscape of pyruvate under simulated hydrothermal vent conditions. *PNAS* 110(33):13283–88
125. Nunoura T, Chikaraishi Y, Izaki R, Suwa T, Sato T, et al. 2018. A primordial and reversible TCA cycle in a facultatively chemolithoautotrophic thermophile. *Science* 359(6375):559–63
126. Nützmann H-W, Scazzocchio C, Osbourn A. 2018. Metabolic gene clusters in eukaryotes. *Annu. Rev. Genet.* 52:159–83
127. One Thousand Plant Transcriptomes Initiative. 2019. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* 574(7780):679–85
128. Orgel LE. 2004. Prebiotic chemistry and the origin of the RNA world. *Crit. Rev. Biochem. Mol. Biol.* 39(2):99–123

120. Discovered alterations of the committed enzyme of leucine biosynthesis in cultivated and wild tomatoes, which likely underlie distinct compositions of acylsugar specialized metabolites derived from leucine and valine.

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

129. Parween S, Anonuevo JJ, Butardo VM, Misra G, Anacleto R, et al. 2020. Balancing the double-edged sword effect of increased resistant starch content and its impact on rice texture: its genetics and molecular physiological mechanisms. *Plant Biotechnol. J.* 18(8):1763–77
130. Patron NJ, Rogers MB, Keeling PJ. 2004. Gene replacement of fructose-1,6-bisphosphate aldolase supports the hypothesis of a single photosynthetic ancestor of chromalveolates. *Eukaryot. Cell* 3(5):1169–75
131. Perez de Souza L, Scossa F, Proost S, Bitocchi E, Papa R, et al. 2019. Multi-tissue integration of transcriptomic and specialized metabolite profiling provides tools for assessing the common bean (*Phaseolus vulgaris*) metabolome. *Plant J.* 97(6):1132–53
132. Petersen J, Brinkmann H, Cerff R. 2003. Origin, evolution, and metabolic role of a novel glycolytic GAPDH enzyme recruited by land plant plastids. *J. Mol. Evol.* 57(1):16–26
133. Plaxton WC. 1996. The organization and regulation of plant glycolysis. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47:185–214
134. Radwanski ER, Last RL. 1995. Tryptophan biosynthesis and metabolism: biochemical and molecular genetics. *Plant Cell* 7(7):921–34
135. Raff RA. 2000. Evo-devo: the evolution of a new discipline. *Nat. Rev. Genet.* 1(1):74–79
136. Ralser M. 2014. The RNA world and the origin of metabolic enzymes. *Biochem. Soc. Trans.* 42(4):985–88
137. Ralser M. 2018. An appeal to magic? The discovery of a non-enzymatic metabolism and its role in the origins of life. *Biochem. J.* 475(16):2577–92
138. Raymond J, Segrè D. 2006. The effect of oxygen on biochemical networks and the evolution of complex life. *Science* 311(5768):1764–67
139. Reyes-Prieto A, Moustafa A. 2012. Plastid-localized amino acid biosynthetic pathways of Plantae are predominantly composed of non-cyanobacterial enzymes. *Sci. Rep.* 2:955
140. Richards TA, Dacks JB, Campbell SA, Blanchard JL, Foster PG, et al. 2006. Evolutionary origins of the eukaryotic shikimate pathway: gene fusions, horizontal gene transfer, and endosymbiotic replacements. *Eukaryot. Cell* 5(9):1517–31
141. Sage RF, Christin P-A, Edwards EJ. 2011. The C_4 plant lineages of planet Earth. *J. Exp. Bot.* 62(9):3155–69
142. Sato N, Awai K. 2016. Diversity in biosynthetic pathways of galactolipids in the light of endosymbiotic origin of chloroplasts. *Front. Plant Sci.* 7:117
143. Sato N, Awai K. 2017. "Prokaryotic Pathway" is not prokaryotic: Noncyanobacterial origin of the chloroplast lipid biosynthetic pathway revealed by comprehensive phylogenomic analysis. *Genome Biol. Evol.* 9(11):3162–78
144. Sato T, Atomi H, Imanaka T. 2007. Archaeal type III RuBisCOs function in a pathway for AMP metabolism. *Science* 315(5814):1003–6
145. Sauer FD, Erfle JD, Mahadevan S. 1975. Amino acid biosynthesis in mixed rumen cultures. *Biochem. J.* 150(3):357–72
146. Schenck CA, Chen S, Siehl DL, Maeda HA. 2015. Non-plastidic, tyrosine-insensitive prephenate dehydrogenases from legumes. *Nat. Chem. Biol.* 11(1):52–57
147. Schenck CA, Holland CK, Schneider MR, Men Y, Lee SG, et al. 2017. Molecular basis of the evolution of alternative tyrosine biosynthetic routes in plants. *Nat. Chem. Biol.* 13:1029–35
148. Schenck CA, Last RL. 2020. Location, location! Cellular relocalization primes specialized metabolic diversification. *FEBS J.* 287(7):1359–68
149. Schlüter U, Weber APM. 2020. Regulation and evolution of C_4 photosynthesis. *Annu. Rev. Plant Biol.* 71:183–215
150. Schnarrenberger C, Martin W. 2002. Evolution of the enzymes of the citric acid cycle and the glyoxylate cycle of higher plants. A case study of endosymbiotic gene transfer. *Eur. J. Biochem.* 269(3):868–83
151. Schopf JW, Kitajima K, Spicuzza MJ, Kudryavtsev AB, Valley JW. 2018. SIMS analyses of the oldest known assemblage of microfossils document their taxon-correlated carbon isotope compositions. *PNAS* 115(1):53–58
152. Scossa F, Fernie AR. 2020. The evolution of metabolism: How to test evolutionary hypotheses at the genomic level. *Comput. Struct. Biotechnol. J.* 18:482–500

143. Detailed phylogenomic and phylogenetic analyses revealed that all enzymes of plastidic glycolipid biosynthesis have been replaced by noncyanobacterial enzymes in Plantae.

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

153. Seto H, Watanabe H, Furihata K. 1996. Simultaneous operation of the mevalonate and non-mevalonate pathways in the biosynthesis of isopentenly diphosphate in *Streptomyces aeriouvifer*. *Tetrahedron Lett.* 37(44):7979–82
154. Shameer S, Baghalian K, Cheung CYM, Ratcliffe RG, Sweetlove LJ. 2018. Computational analysis of the productivity potential of CAM. *Nat. Plants* 4(3):165–71
155. Shan Q, Zhang Y, Chen K, Zhang K, Gao C. 2015. Creation of fragrant rice by targeted knockout of the *OsBADH2* gene using TALEN technology. *Plant Biotechnol. J.* 13(6):791–800
156. Shih PM. 2018. Towards a sustainable bio-based economy: redirecting primary metabolism to new products with plant synthetic biology. *Plant Sci.* 273:84–91
157. Singh BK, Shaner DL. 1995. Biosynthesis of branched chain amino acids: from test tube to field. *Plant Cell* 7(7):935–44
158. Sojo V, Herschy B, Whicher A, Camprubí E, Lane N. 2016. The origin of life in alkaline hydrothermal vents. *Astrobiology* 16(2):181–97
159. Somerville C, Browse J. 1991. Plant lipids: metabolism, mutants, and membranes. *Science* 252(5002):80–87
160. Sonawane PD, Jozwiak A, Panda S, Aharoni A. 2020. "Hijacking" core metabolism: a new panache for the evolution of steroidal glycoalkaloids structural diversity. *Curr. Opin. Plant Biol.* 55:118–28
161. Soo RM, Hemp J, Parks DH, Fischer WW, Hugenholtz P. 2017. On the origins of oxygenic photosynthesis and aerobic respiration in Cyanobacteria. *Science* 355(6332):1436–40
162. Sørensen I, Pettolino FA, Bacic A, Ralph J, Lu F, et al. 2011. The charophycean green algae provide insights into the early origins of plant cell walls. *Plant J.* 68(2):201–11
163. Springsteen G, Yerabolu JR, Nelson J, Rhea CJ, Krishnamurthy R. 2018. Linked cycles of oxidative decarboxylation of glyoxylate as protometabolic analogs of the citric acid cycle. *Nat. Commun.* 9(1):91
164. Studart-Guimarães C, Gibon Y, Frankel N, Wood CC, Zanor MI, et al. 2005. Identification and characterisation of the α and β subunits of succinyl CoA ligase of tomato. *Plant Mol. Biol.* 59(5):781–91
165. Swanson-Wagner R, Briskine R, Schaefer R, Hufford MB, Ross-Ibarra J, et al. 2012. Reshaping of the maize transcriptome by domestication. *PNAS* 109(29):11878–83
166. Sweetlove LJ, Fernie AR. 2005. Regulation of metabolic networks: understanding metabolic complexity in the systems biology era. *New Phytol.* 168(1):9–24
167. Sweetlove LJ, Fernie AR. 2013. The spatial organization of metabolism within the plant cell. *Annu. Rev. Plant Biol.* 64:723–46
168. Tabita FR, Hanson TE, Li H, Satagopan S, Singh J, Chan S. 2007. Function, structure, and evolution of the RubisCO-like proteins and their RubisCO homologs. *Microbiol. Mol. Biol. Rev.* 71(4):576–99
169. Tabita FR, Hanson TE, Satagopan S, Witte BH, Kreel NE. 2008. Phylogenetic and evolutionary relationships of RubisCO and the RubisCO-like proteins and the functional lessons provided by diverse molecular forms. *Philos. Trans. R. Soc. B* 363(1504):2629–40
170. Tohge T, Wendenburg R, Ishihara H, Nakabayashi R, Watanabe M, et al. 2016. Characterization of a recently evolved flavonol-phenylacyltransferase gene provides signatures of natural light selection in Brassicaceae. *Nat. Commun.* 7:12399
171. Varma SJ, Muchowska KB, Chatelain P, Moran J. 2018. Native iron reduces CO_2 to intermediates and end-products of the acetyl-CoA pathway. *Nat. Ecol. Evol.* 2(6):1019–24
172. Velasco AM, Leguina JI, Lazcano A. 2002. Molecular evolution of the lysine biosynthetic pathways. *J. Mol. Evol.* 55(4):445–59
173. Verschueren KHG, Blanchet C, Felix J, Dansercoer A, De Vos D, et al. 2019. Structure of ATP citrate lyase and the origin of citrate synthase in the Krebs cycle. *Nature* 568(7753):571–75
174. Vranová E, Coman D, Gruissem W. 2013. Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annu. Rev. Plant Biol.* 64:665–700
175. Washburn JD, Bird KA, Conant GC, Pires JC. 2016. Convergent evolution and the origin of complex phenotypes in the age of systems biology. *Int. J. Plant Sci.* 177(4):305–18
176. Wasternack C, Feussner I. 2018. The oxylipin pathways: biochemistry and function. *Annu. Rev. Plant Biol.* 69:363–86
177. Weiss MC, Sousa FL, Mrnjavac N, Neukirchen S, Roettger M, et al. 2016. The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* 1(9):16116

171. Discovered that several native metals can facilitate pyruvate and acetate from CO_2, which may provide nonenzymatic precursor reactions of the reductive acetyl-CoA pathway.

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

178. Weng J-K. 2014. The evolutionary paths towards complexity: a metabolic perspective. *New Phytol.* 201(4):1141–49
179. Weng J-K, Philippe RN, Noel JP. 2012. The rise of chemodiversity in plants. *Science* 336(6089):1667–70
180. Widhalm JR, Ducluzeau A-L, Buller NE, Elowsky CG, Olsen LJ, Basset GJC. 2012. Phylloquinone (vitamin K_1) biosynthesis in plants: two peroxisomal thioesterases of Lactobacillales origin hydrolyze 1,4-dihydroxy-2-naphthoyl-CoA. *Plant J.* 71(2):205–15
181. Wilmanns M, Hyde CC, Davies DR, Kirschner K, Jansonius JN. 1991. Structural conservation in parallel β/α-barrel enzymes that catalyze three sequential reactions in the pathway of tryptophan biosynthesis. *Biochemistry* 30(38):9161–69
182. Wong JT. 1975. A co-evolution theory of the genetic code. *PNAS* 72(5):1909–12
183. Xie G, Keyhani NO, Bonner CA, Jensen RA. 2003. Ancient origin of the tryptophan operon and the dynamics of evolutionary change. *Microbiol. Mol. Biol. Rev.* 67(3):303–42
184. Xing A, Last RL. 2017. A regulatory hierarchy of the Arabidopsis branched-chain amino acid metabolic network. *Plant Cell* 29(6):1480–99
185. Xiong J, Bauer CE. 2002. Complex evolution of photosynthesis. *Annu. Rev. Plant Biol.* 53:503–21
186. Xu G, Cao J, Wang X, Chen Q, Jin W, et al. 2019. Evolutionary metabolomics identifies substantial metabolic divergence between maize and its wild ancestor, teosinte. *Plant Cell* 31(9):1990–2009
187. Yang W, Simpson JP, Li-Beisson Y, Beisson F, Pollard M, Ohlrogge JB. 2012. A land-plant-specific glycerol-3-phosphate acyltransferase family in Arabidopsis: substrate specificity, *sn*-2 preference, and evolution. *Plant Physiol.* 160(2):638–52
188. Yang X, Hu R, Yin H, Jenkins J, Shu S, et al. 2017. The *Kalanchoë* genome provides insights into convergent evolution and building blocks of crassulacean acid metabolism. *Nat. Commun.* 8(1):1899
189. Yasutake Y, Yao M, Sakai N, Kirita T, Tanaka I. 2004. Crystal structure of the *Pyrococcus horikoshii* isopropylmalate isomerase small subunit provides insight into the dual substrate specificity of the enzyme. *J. Mol. Biol.* 344(2):325–33
190. Yokota T, Ohnishi T, Shibata K, Asahina M, Nomura T, et al. 2017. Occurrence of brassinosteroids in non-flowering land plants, liverwort, moss, lycophyte and fern. *Phytochemistry* 136:46–55
191. Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D. 2004. A molecular timeline for the origin of photosynthetic eukaryotes. *Mol. Biol. Evol.* 21(5):809–18
192. Yuzawa Y, Shimojima M, Sato R, Mizusawa N, Ikeda K, et al. 2014. Cyanobacterial monogalactosyldiacylglycerol-synthesis pathway is involved in normal unsaturation of galactolipids and low-temperature adaptation of *Synechocystis* sp. PCC 6803. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* 1841(4):475–83
193. Zhang J. 2003. Evolution by gene duplication: an update. *Trends Ecol. Evol.* 18(6):292–98
194. Zhang L, Su W, Tao R, Zhang W, Chen J, et al. 2017. RNA sequencing provides insights into the evolution of lettuce and the regulation of flavonoid biosynthesis. *Nat. Commun.* 8(1):2264
195. Zhang S, Bryant DA. 2011. The tricarboxylic acid cycle in cyanobacteria. *Science* 334(6062):1551–53
196. Zhang XV, Martin ST. 2006. Driving parts of Krebs cycle in reverse through mineral photochemistry. *J. Am. Chem. Soc.* 128(50):16032–33
197. Zhao G, Lian Q, Zhang Z, Fu Q, He Y, et al. 2019. A comprehensive genome variation map of melon identifies multiple domestication events and loci influencing agronomic traits. *Nat. Genet.* 51(11):1607–15
198. Zhao Q, Yang J, Cui M-Y, Liu J, Fang Y, et al. 2019. The reference genome sequence of *Scutellaria baicalensis* provides insights into the evolution of wogonin biosynthesis. *Mol. Plant.* 12(7):935–50
199. Zhu G, Wang S, Huang Z, Zhang S, Liao Q, et al. 2018. Rewiring of the fruit metabolome in tomato breeding. *Cell* 172(1–2):249–61.e12

199. A comprehensive analysis of the effect of domestication and improvement processes on the tomato metabolome revealing the sites of selection for removal of toxic steroidal glycoalkaloids.

Annual Review of Plant Biology

Volume 72, 2021

Contents

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.

Errata

An online log of corrections to *Annual Review of Plant Biology* articles may be found at http://www.annualreviews.org/errata/arplant

Annu. Rev. Plant Biol. 2021.72:185-216. Downloaded from www.annualreviews.org
Access provided by Chinese Academy of Sciences - Institute of Botany on 07/18/22. For personal use only.