Glucosinolate and Amino Acid Biosynthesis in Arabidopsis[1]

Ben Field, Guillermo Cardon, Maria Traka, Johan Botterman, Guy Vancanneyt, and Richard Mithen*

Department of Metabolic Biology, John Innes Centre, Norwich NR4 7UH (B.F.); BayerCrop Science N.V., B–9000 Gent, Belgium (G.C., J.B., G.V.); and Nutrition Division, Institute of Food Research, Norwich NR4 7UA, United Kingdom (M.T., R.M.)

Enzymes that catalyze the condensation of acetyl coenzyme A and 2-oxo acids are likely to be important in two distinct metabolic pathways in Arabidopsis. These are the synthesis of isopropylmalate, an intermediate of Leu biosynthesis in primary metabolism, and the synthesis of methylthioalkylmalates, intermediates of Met elongation in the synthesis of aliphatic glucosinolates (GSLs), in secondary metabolism. Four Arabidopsis genes in the ecotype Columbia potentially encode proteins that could catalyze these reactions. *MAM1* and *MAML* are adjacent genes on chromosome 5 at the *Gsl-elong* locus, while *MAML-3* and *MAML-4* are at opposite ends of chr 1. The isopropylmalate synthase activity of each member of the MAM-like gene family was investigated by heterologous expression in an isopropylmalate synthase-null *Escherichia coli* mutant. Only the expression of *MAML-3* restored the ability of the mutant to grow in the absence of Leu. A *MAML* knockout line (KO) lacked long-chain aliphatic GSLs, which were restored when the KO was transformed with a functional *MAML* gene. Variation in expression of *MAML* did not alter the total levels of Met-derived GSLs, but just the ratio of chain lengths. *MAML* overexpression in Columbia led to an increase in long-chain GSLs, and an increase in 3C GSLs. Moreover, plants overexpressing *MAML* contained at least two novel amino acids. One of these was positively identified via MS/MS as homo-Leu, while the other, with identical mass and fragmentation patterns, was likely to be homo-Ile. A *MAML-4* KO did not exhibit any changes in GSL profile, but had perturbed soluble amino acid content.

Arabidopsis contains a series of glucosinolates (GSLs) derived from elongated forms of Met (Table I; Haughn et al., 1991). Elongation of Met is by a process analogous to the synthesis of Leu from 2-oxo-3-methylbutanoate (Fig. 1, a and b). Met is initially transaminated to an oxo acid. The oxo acid undergoes condensation with acetyl coenzyme A (CoA) to result in 2-methylthiopropylmalate, which then undergoes isomerization and oxidative decarboxylation to result in the net gain of a single methyl unit to the oxo acid. This elongated oxo acid can be transaminated to an amino acid and enter core GSL biosynthesis, or else undergo further rounds of acetyl CoA condensation (Fig. 1, a and b; Chisholm and Wetter, 1964; Graser et al., 2000). In Arabidopsis leaves, the major GSLs are derived from Met that has either undergone a single round of elongation to produce 3C GSLs, such as 3-methylsulphinylpropyl, 3-hydroxypropyl, or 2-propenyl GSLs, or two rounds of elongation to produce 4C GSLs, such as 4-methylsulphinylbutyl and 3-butenyl GSLs. The 3C GSLs predominate in the ecotype Landsberg *erecta* (L*er*), while 4C GSLs predominate in ecotype Columbia (Col-0). In addition, all ecotypes produce similar amounts of long chain GSLs in leaves and seeds, derived from four, five, and six rounds of Met elongation. The variation in 3C to 4C ratio is determined by alleles at the *Gsl-elong* locus on chromosome 5 (de Quiros et al., 2000; Kroymann et al., 2001). In Col-0, this locus contained two genes, *MAM1* (At5g23020) and *MAML* (At5g23010), which have high levels of homology to isopropylmalate synthase (IPMS; EC 2.3.3.13) that catalyzes acetyl CoA-oxo acid condensation in Leu biosynthesis. Of these, the *MAM1* gene has been shown to determine the production of 4C GSLs (Kroymann et al., 2001). Variation in *MAM1* has no effect on the expression of 6C, 7C, and 8C GSLs. Hence it is likely that there are other genes in Arabidopsis responsible both for 3C GSL synthesis, and for 5C, 6C, 7C, and 8C GSL biosynthesis.

In addition to determining the ratio of 3C to 4C chain lengths, the *Gsl-elong* locus functions as a quantitative trait locus determining overall levels of Met-derived GSLs (de Quiros et al., 2000; Kroymann et al., 2001), similar to a quantitative trait locus in *Brassica* that determines both 3C to 4C ratio and total amounts (Mithen et al., 2003). Fine mapping and sequence analysis in several Arabidopsis ecotypes reveal that the *Gsl-elong* locus comprises, in addition to *MAM1* and *MAML*, a further gene designated *MAM2*. Whereas all ecotypes appear to have a functioning *MAML* gene, only some ecotypes have both *MAM1* and *MAM2*, while others have either a functioning *MAM1* or a functioning *MAM2*. The relationship between *MAM1* and *MAM2* is complex; putative reciprocal deletion of *MAM2* in Col-0 and *MAM1* in

[1] This work was supported by the John Innes Foundation (BF), by the BBSRC, and by UK Bayer Crop Science.

* Corresponding author; e-mail richard.mithen@bbsrc.ac.uk; fax 44–1603–507–723.

Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.104.039347.

Table I. *Met-derived GSLs in seeds of Arabidopsis*
The numbers correspond to those in Figure 4.

3_1	3C	3-Methylthiopropyl
3_2	3C	3-Methylsulphinylpropyl
3_3	3C	3-Hydroxypropyl
3_4	3C	3-Benzooxylopropyl
		Σ = Total 3C
4_1	4C	4-Methylthiobutyl
4_2	4C	4-Methylsulphinylbutyl
4_3	4C	4-Hydroxybutyl
4_4	4C	4-Benzoxyolbutyl
		Σ = Total 4C
5_1	5C	5-Methylthiopentyl
5_2	5C	5-Methylsulphinylpentyl
		Σ = Total 5C
6_1	6C	6-Methylthiohexyl
6_2	6C	6-Methylsulphinylhexyl
		Σ = Total 6C
7_1	7C	7-Methylthioheptyl
7_2	7C	7-Methylsulphinylheptyl
		Σ = Total 7C
8_1	8C	8-Methylthiooctyl
8_2	8C	8-methylsulphinyloctyl
		Σ = Total 8C

L*er* has resulted in the remaining *MAM1* and *MAM2* genes segregating as alleles of each other, and, in other ecotypes, there is evidence for genetic interchange between *MAM1* and *MAM2* (Kroymann et al., 2003). Ecotypes that have a functioning *MAM1* gene, such as Col-0, make 4C GSLs, whereas those in which *MAM1* is impaired in function synthesize 3C GSLs. Whether 3C synthesis is due to *MAM2* expression, or to other genes, has not been functionally demonstrated. In addition to these three genes, Arabidopsis contains two more genes in this family, designated *MAML*-3 (At1g74040) and *MAML*-4 (At1g18500), both found on chromosome 1.

The predicted proteins of all five MAM genes suggest that they may catalyze the condensation of acetyl CoA with 2-oxo acids (Fig. 1c); they all have highly conserved domains near the amino terminus which are signatures of the active site for oxo acid condensation reactions (Evans et al., 1991; de Quiros et al., 2000) as well as strong overall similarity to IPMS and homocitrate synthase (EC 2.3.3.14), which catalyze such condensation reactions in yeast and bacteria (Fig. 1c). Within primary metabolism, one or more of these genes must encode an IPMS, which catalyzes the condensation of 2-oxo-3-methylbutanoic acid with acetyl CoA to result in isopropylmalate (3-carboxy-3-hydroxy-4-methylpentanoate), a necessary intermediate of Leu biosynthesis (Fig. 1b). IPMS has been studied extensively in yeast and *Escherichia coli*, and has been shown in both these organisms to have an important regulatory function (Umbarger, 1997). In plants, an IPMS activity has been partially purified from spinach chloroplasts and shown to be strongly feedback inhibited by Leu (Hagelstein and Schultz, 1993). Several IPMS-like genes have also been isolated from higher plants, but without associated functional analysis. Junk and Mourad (2002) report the expression of three of the Arabidopsis IPMS-like genes in an *E. coli* Leu auxotroph, and report complementation with *MAM1* and *MAML*, although no data were provided.

Within secondary metabolism, all or some of these genes are likely to function in the synthesis of GSLs, and one or more must function as an IPMS. Currently, a functional analysis has only been undertaken for *MAM1* (Kroymann et al., 2001). In the current study we firstly show that of the four Col-0 MAM genes, only *MAML*-3 has IPMS activity when heterologously expressed in *E. coli* and is thus likely to function as IPMS in planta. Secondly, we demonstrate that expression of *MAML* results in synthesis of long chain GSLs, but that overexpression of *MAML* in Arabidopsis leads to the synthesis of homoleucine and isohomoleucine, novel amino acids not found in wild-type Arabidopsis. Thirdly, while we cannot establish a precise biochemical role for *MAML*4, we show that a *MAML*-4 knockout mutant has perturbed amino acid profile, not inconsistent with IPMS activity, suggesting an important role in amino acid biosynthesis. The intimate association between GSL biosynthesis and amino acid biosynthesis is discussed.

RESULTS

The Arabidopsis Col-0 MAM-Like Gene Family

Within this paper, we will refer to the four Col-0 members of the gene family as *MAM1* (At5g23020), *MAML* (At5g23010), *MAML-3* (At1g74040), and *MAML-4* (At1g18500). The predicted protein for each MAM synthase contains a LeuA domain and a chloroplast leader-peptide as predicted by TargetP, $P > 0.9$ (Fig. 1c). *MAM1* and *MAML* form a subgroup of the Arabidopsis MAM synthase family sharing 78% amino acid identity, while *MAML-3* and *MAML-4* form another subgroup sharing 90% identity (Fig. 1c).

Heterologous Expression of MAM Synthases in a ΔLeuA (IPMS-Null) *E. coli* Mutant

Expression constructs were designed to heterologously express the predicted mature peptide encoded by each MAM-like Arabidopsis gene, with the addition of an N-terminal 6× His tag. *MAM1, MAML, MAML-3,* and *MAML-4* lacking their predicted leader-peptide sequences were amplified from total Arabidopsis cDNA and cloned into the *E. coli* expression vector pQE-30 to give pQE-*MAM1*, pQE-*MAML*, pQE-*MAML-3*, and pQE-*MAML-4*. The *E. coli* strain CV512 has a nonfunctional IPMS (Δ*Leu*A), rendering it unable to grow on media lacking Leu (Somers et al., 1973). Proteins of the predicted sizes were produced upon induction of CV512 containing either pQE-*MAM1*,

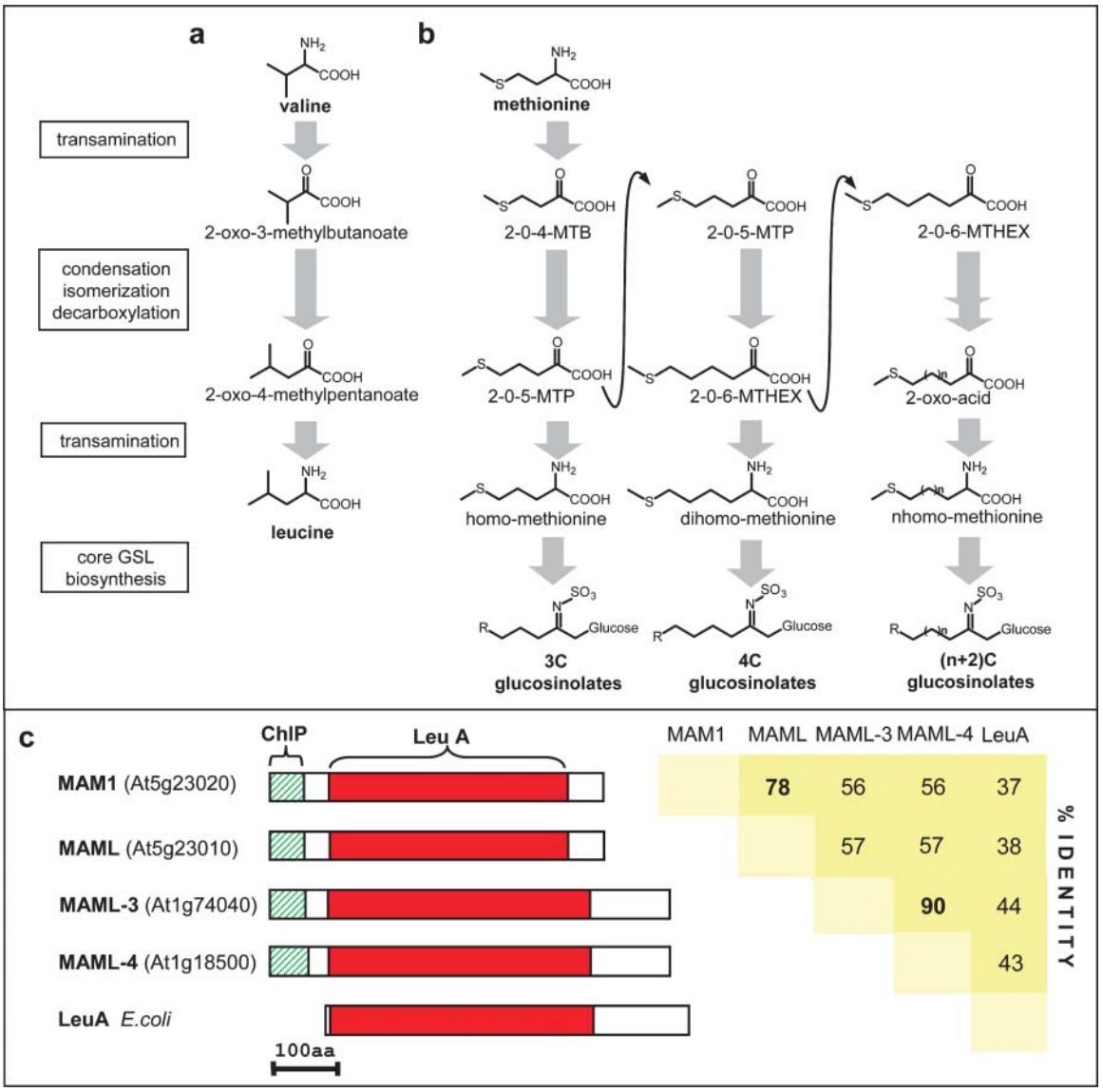

Figure 1. Models of 2-oxo acid elongation and LeuA homologs from Arabidopsis. a, Elongation in primary metabolism for Leu biosynthesis in which IPMS (LeuA) catalyzes the initial condensation with acetyl CoA. b, An analogous elongation mechanism in secondary metabolism for GSL biosynthesis, note that the elongated 2-oxo acid product may undergo further rounds of elongation. c, Aligned domains of the putative Arabidopsis *MAM*-like family and *E. coli* LeuA showing predicted chloroplast target peptides (ChIP) and LeuA catalytic domains (left) as well as pair-wise identities (right). 2-O-4-MTB, 2-oxo-4-methylthiobutanoate; 2-O-5-MTP, 2-oxo-5-methylthiopentanoate; 2-O-6-MTHEX, 2-oxo-6-methylthiohexanoate.

MAML, *MAML-3*, or *MAML-4* (Fig. 2a). Only CV512 producing the mature *MAML*-3 protein was able grow on minimal media lacking Leu (Fig. 2, b and c).

Characterization of MAM-Like Insertion Lines

As neither *MAML* nor *MAML-4* exhibited IPMS activity, they may be involved in GSL biosynthesis. We obtained homozygous knockout (KO) lines for *MAML* and *MAML-4*. MAML::En1 (Col-0 background) contains a single, stable *En-1* insertion in the 1st exon of *MAML* and Garlic1175 (Col-0 background) contains a single T-DNA insertion in the 10th intron of *MAML-4* (Fig. 3a). These two lines are referred to as the MAML KO and MAML-4 KO lines, respectively. Reverse transcription (RT)-PCR analysis of MAML KO cDNA showed that no *MAML* transcript was produced (data not shown), while analysis of MAML-4 KO cDNA showed that chimeric *MAML-4*::*T-DNA* fusion transcripts were produced. However, the chimeric transcript in MAML-4 KO was present at considerably reduced levels compared to the normal transcript in Col-0 (Fig. 3b). No compensatory increase in expression was observed for the other MAM synthase genes.

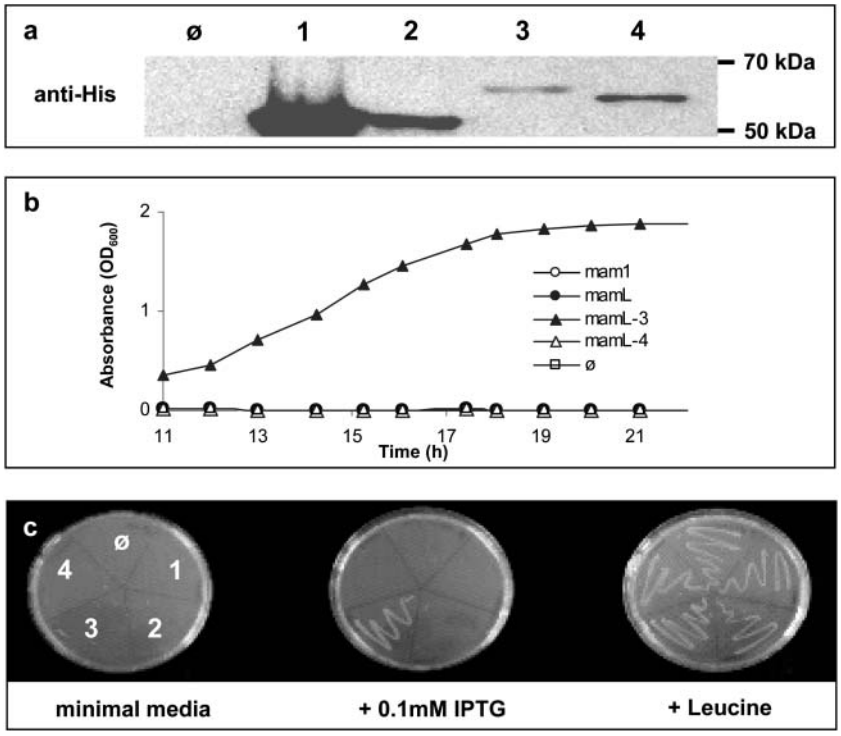

Figure 2. Complementation analysis of the *E. coli* Leu auxotroph CV512 expressing the empty vector control pQE-30 (0), pQE-*MAM1* (1), pQE-*MAML* (2), pQE-*MAML-3* (3), and pQE-*MAML-4* (4). a, Western blot of total protein extracts from CV512 containing each construct following 2 h induction with 0.1 mM IPTG at 37°C. Peptide sizes were determined from a Coomassie gel run in parallel: (1) and (2) approximately 52 kD, (3) approximately 65 kD, and (4) approximately 64 kD. b, Growth curves of CV512 containing each construct in liquid minimal media supplemented with 0.02 mM IPTG and incubated at 37°C. Each point represents the average of three independent growth curve experiments. c, Growth of CV512 containing each construct on solid minimal media (left) + 0.1 mM IPTG (center) or + 30 mM Leu (right).

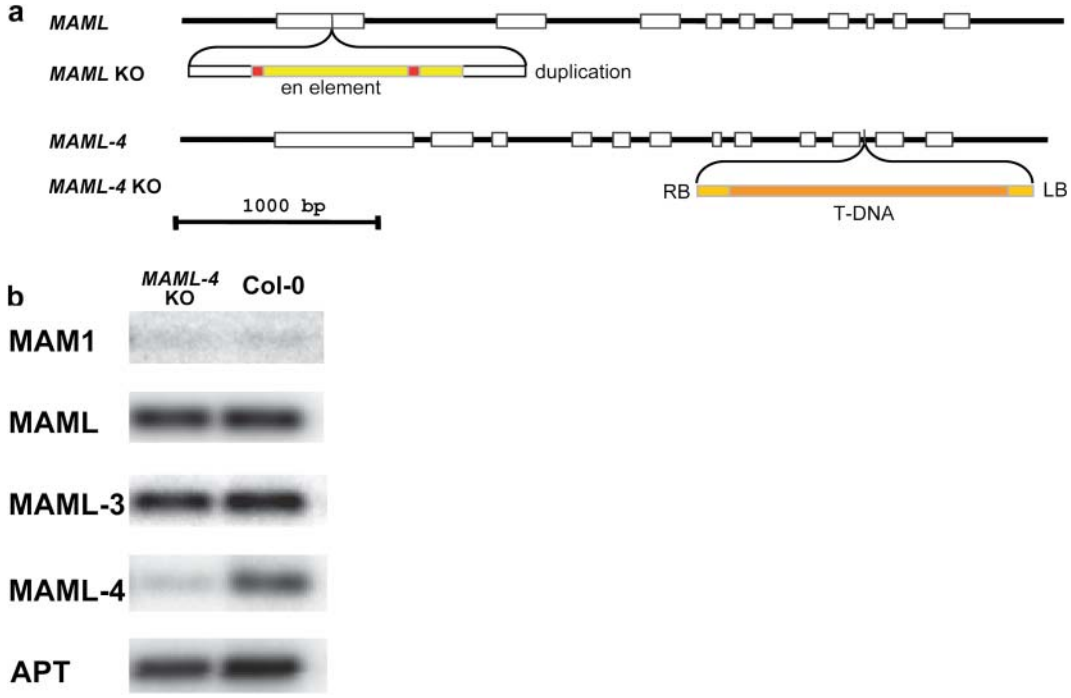

Figure 3. Two MAM KO lines. a, Position of the *En-1* transposon insertion in *MAML* and T-DNA insertion in *MAML-4*. In each case the complete gene structure is shown, white boxes denote exons while black boxes denote untranslated regions. b, Semiquantitative RT-PCR analysis of MAM gene expression in the MAML-4 KO line compared to the wild-type Col-0. The PCR product is a 539 bp section of exon 1. In each case 18 cycles of PCR amplification was used, and analysis was repeated three times to ensure reproducibility. MAM1 expression in Col-0 could be detected after 40 rounds of PCR. The constitutively expressed housekeeping gene APT was used as an mRNA loading control.

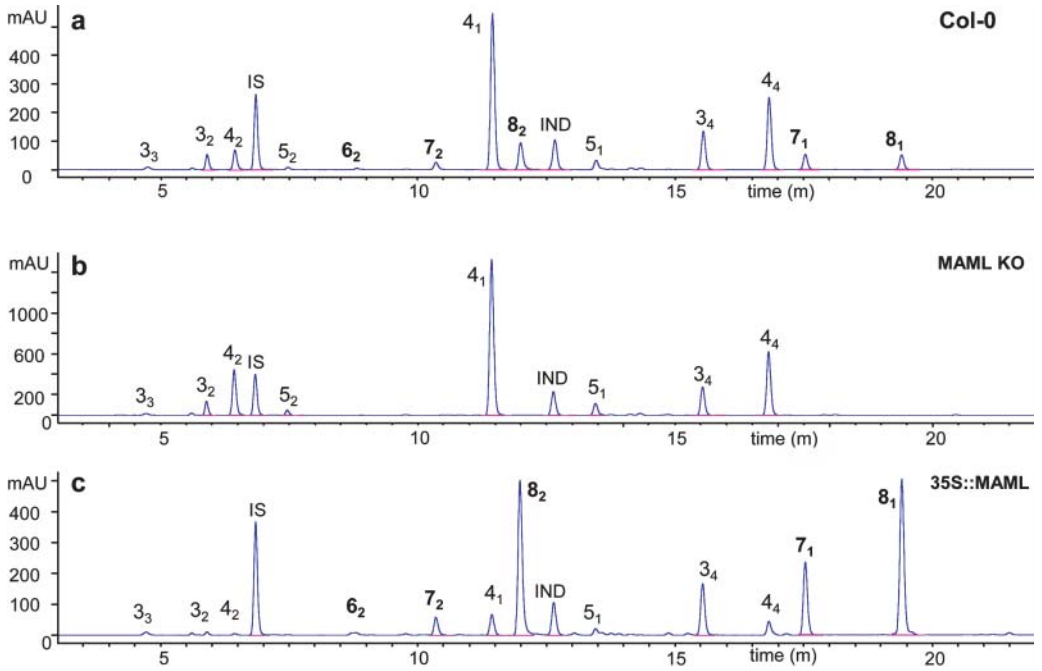

Figure 4. The MAML KO lacks 6C, 7C, and 8C GSLs. Chromatograms of desulphoglucosinolates extracted from the seeds of (a) wild-type Col-0, (b) the MAML KO line, and (c) the MAML KO line transformed with a 35S driven copy of *MAML*. The 6C, 7C, and 8C GSLs are indicated in bold; none were detected in the MAML KO. See Table I for peak number identification. IND, indole GSLs (derived from Trp); IS, internal standard.

The growth and appearance of the *MAML* KO line did not differ significantly from the wild-type controls. However, the *MAML*-4 KO line showed a 20% reduction in germination compared to Col-0. Subsequent growth appeared normal.

MAML Is Required for Long Chain GSLs

GSLs were analyzed in the seeds and leaves of the MAML KO line. There was a complete absence of 6C, 7C, and 8C GSLs (Fig. 4, a and b). Despite the loss of these compounds, the total level of Met-derived GSLs in the knockout and controls was not significantly different (Table II), due to enhanced levels of 4C GSLs. To test whether these alterations in GSL composition were controlled by a nonfunctional *MAML* allele, both the MAML KO line and wild-type Col-0 were transformed with vector pTKC28, containing a wild-type genomic copy of the *MAML* gene with the endogenous promoter replaced by the cauliflower mosaic virus

Table II. *Seed GSL content (μmol g^{-1}) of primary transformants of Col-0 and the MAML KO line transformed with an empty vector (pTKC24) or the same vector containing a genomic copy of MAML (pTKC28)*

Data from independent transformants are shown. Additional transformants gave similar results.

Line	GSLs						
	3C	4C	5C	6C	7C	8C	Total
				μmol g^{-1} seed			
Col-0 + pTKC24 (empty vector control)							
729	4.0	25.7	1.3	0.0	2.1	5.6	38.7
731	5.2	30.6	1.7	0.0	2.5	6.9	46.9
Col-0 + pTKC28 (35S::MAML)							
549	5.0	3.8	0.9	0.4	11.9	25.4	47.4
567	6.3	3.0	0.0	0.5	5.4	27.2	42.4
570	12.3	5.0	1.1	0.4	5.0	20.7	44.5
MAML KO + pTKC24 (empty vector control)							
659	5.9	45.5	2.7	0.0	0.0	0.0	54.1
669	3.0	29.8	2.0	0.0	0.0	0.0	34.8
674	5.4	38.2	2.2	0.0	0.0	0.0	45.8
MAML KO + pTKC28 (35S::MAML)							
599	6.6	4.0	0.0	0.5	10.3	36.3	57.7
628	13.2	16.4	0.3	1.0	19.9	23.9	74.7
597	24.8	8.2	1.2	0.0	4.2	12.0	50.4

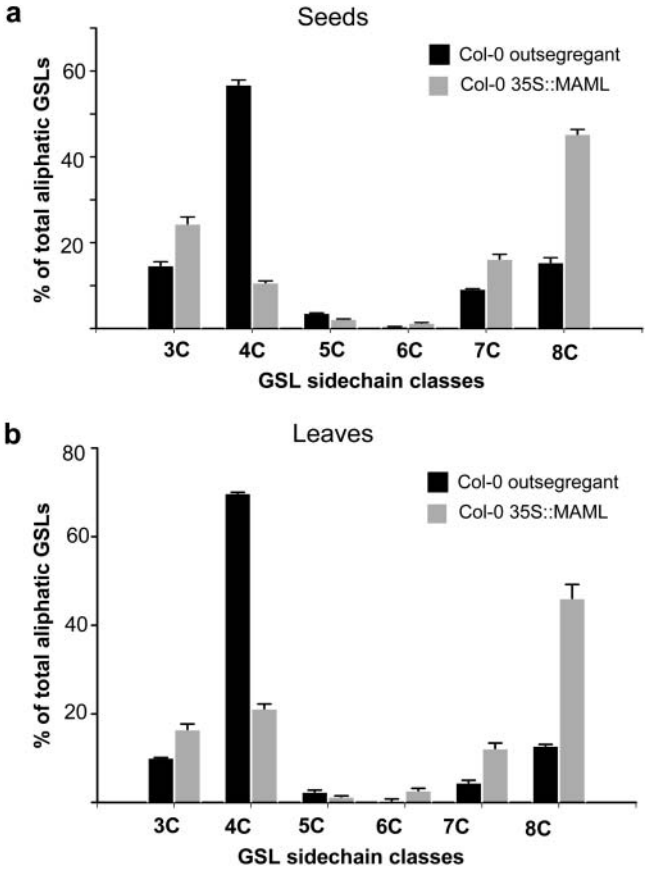

Figure 5. Overexpression of *MAML* enhances the ratio of long chain to short chain aliphatic GSLs. a, GSLs in seeds of Col-0 plants homozygous for 35S::MAML and outsegregants from the same initial primary transformation event that have lost 35S::MAML. The total level of Met-derived GSLs between the two classes are not significantly different (70.2 ± 1.31 and 75.9 ± 4.79 μmol g^{-1}, respectively). b, GSLs in rosette leaves of Col-0 plants homozygous for 35S::MAML and outsegregants from the same initial primary transformation event that have lost 35S::MAML. The total level of Met-derived GSLs between the two classes is not significantly different (20.5 ± 1.95 and 17.4 ± 2.41 μmol g^{-1}, respectively).

(CAMV) 35S promoter. Production of 6C, 7C, and 8C GSLs was restored in the MAML KO line and enhanced in wild-type Col-0 (Fig. 4c; Table II). Segregation analysis of T_2 transformants and subsequent analysis of T_3 families confirmed that the enhanced levels of long chain GSLs in transgenic Col-0 was due to the presence of the transgene (Fig. 4). The total level of aliphatic GSLs was the same in the seeds of T_3 plants whether they possessed the transgene or had lost it through segregation (70.2 ± 1.31 μmol g^{-1}, $n = 15$ and 75.9 ± 4.79 μmol g^{-1}, $n = 5$, respectively), and also in leaves (20.5 ± 1.95 μmol g^{-1} and 17.4 ± 2.41 μmol g^{-1}, respectively). However, the ratio of chain lengths varied in both tissues; those plants that contained the transgene had higher levels of long chain GSLs and 3C GSLs, but lower levels of 4C GSLs compared to wild-type plants and outsegregants in both seeds and leaves (Fig. 4).

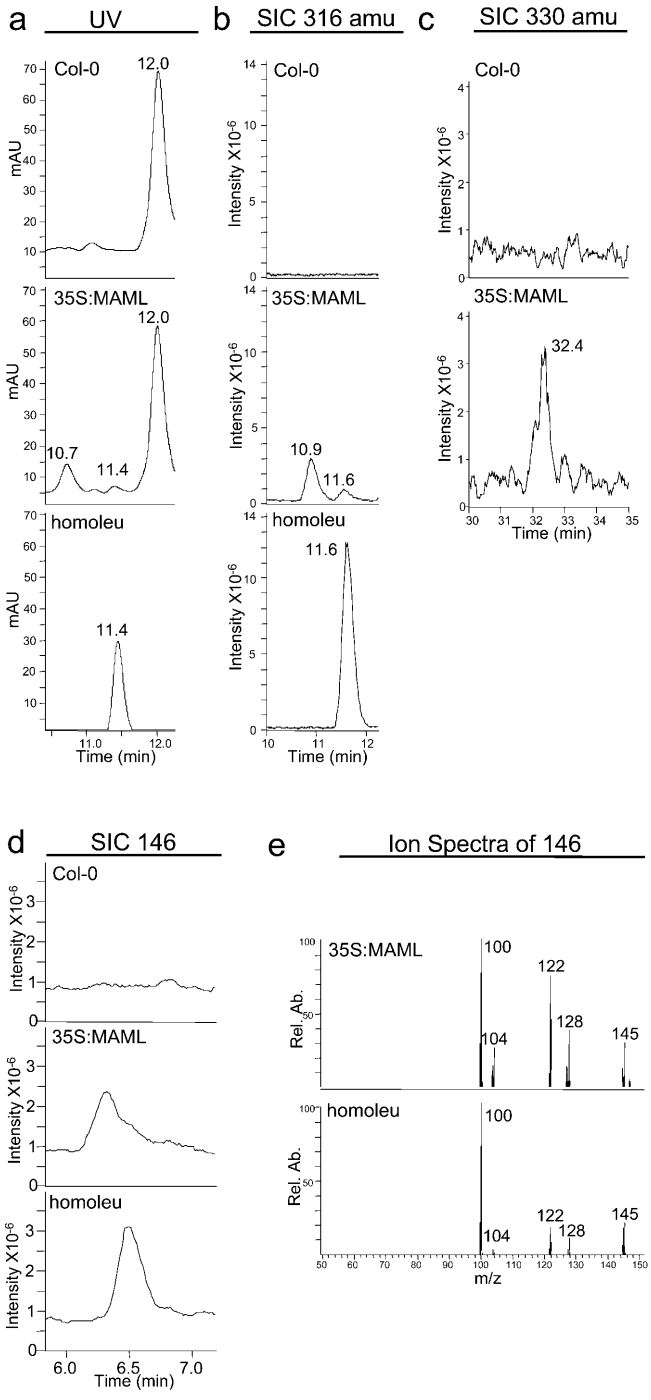

Figure 6. Plants overexpressing *MAML* contain novel amino acids. a, UV-absorbance chromatograms of derivatized amino acid extracts from Col-0 seedlings, 35S::MAML seedlings, and a homoleucine standard, separated using gradient 1. There are two new peaks at 10.7 and 11.4 min in 35S::MAML seedlings. b, Single ion chromatograms at 316 amu in the same region. c, In the same extracts there is a new 330 amu peak at 32.4 min in 35S::MAML seedlings, separated using gradient 2. d, In underivatized extracts the two new compounds from 35S::MAML coelute and (e) have the same ionization spectra as homoleucine.

Overexpression of MAML Results in the Synthesis of Novel Amino Acids

As acetyl CoA-oxo acid condensation is also an important component of amino acid biosynthesis, we investigated the amino acid content of 35S::MAML. Soluble amino acids were extracted from 9-d-old seedlings of 35S::MAML and derivatized with the AccQ Tag reagent. Amino acids were separated and detected by liquid chromatography-fluorescence detection (LC-FLD), and quantified using standards. All amino acids identified in 35S::MAML seedlings were at equivalent levels to wild type (Col-0) with the exception of Tyr. Soluble Tyr decreased more than 2-fold in 35S::MAML seedlings (37 ± 2.5 pmol mg^{-1} in 35S::MAML and 100 ± 7.4 pmol mg^{-1} in Col-0, $n = 4$, $P < 0.005$; Fig. 5). Furthermore, two new peaks, not seen in the wild type, were observed in the LC-FLD chromatograms of 35S::MAML extracts (Fig. 6a). These compounds were investigated by liquid chromatography-mass spectrometry (LC-MS) in derivatized 35S::MAML extracts. The most abundant ion in both the new peaks had a mass of 316, which corresponds to a prederivatization mass of 145 (Fig. 6b). The second of the two peaks has an identical retention time to derivatized homoleucine (11.6 min; Fig. 6, a and b).

To obtain ionization spectra for the novel amine compounds an LC-MS/MS method was developed for underivatized extracts. Two novel peaks of mass 146 (i.e. $M+H^+$) coeluted between 6.2 and 6.6 min in 35S::MAML extracts (Fig. 6d). Homoleucine also eluted at 6.5 min. Ionization spectra were obtained for both novel peaks and the authentic standard (Fig. 6e). All three mass spectra were identical, with a major transition from 146 to 100 amu. This transition is likely to be due to the loss of an HCOOH group. In summary, the underivatized compound eluting at 6.5 min is homoleucine. The earlier peak represents an isomer of homoleucine and is almost certainly homo-Ile. Further analysis suggested the occurrence of a further novel amino acid in 35S::MAML with derivatized M^+H^+ mass of 330, which is consistent with dihomoleucine and/or dihomoisoleucine (Fig. 6c).

MAML-4 KO Plants Have Normal GSLs but Perturbed Amino Acid Content

MAML-4 KO plants had similar GSLs to their wild-type control (data not shown), but had perturbed amino acid content. There were highly significant ($P < 0.005$) increases in His, Val, Asn, and Gln when compared to wild-type Col-0, and less significant decreases in Leu ($P = 0.057$; Fig. 7). The less significant decrease in Leu was due to just one of the four replicates that had similar levels to wild-type Col-0 (MAML-4 KO: 26.6, 27.5, 18.2, and 36.7 pmol mg^{-1}; Col-0: 40.0, 38.0, 38.0, and 32.0 pmol mg^{-1}). If this anomalous result in excluded, the level of probability of a decrease in Leu in MAM4-L KO compared to wild type decreases to $P = 0.01$.

DISCUSSION

Heterologous expression in an *E. coli* IPMS null mutant indicated that *MAML-3* is likely to function as

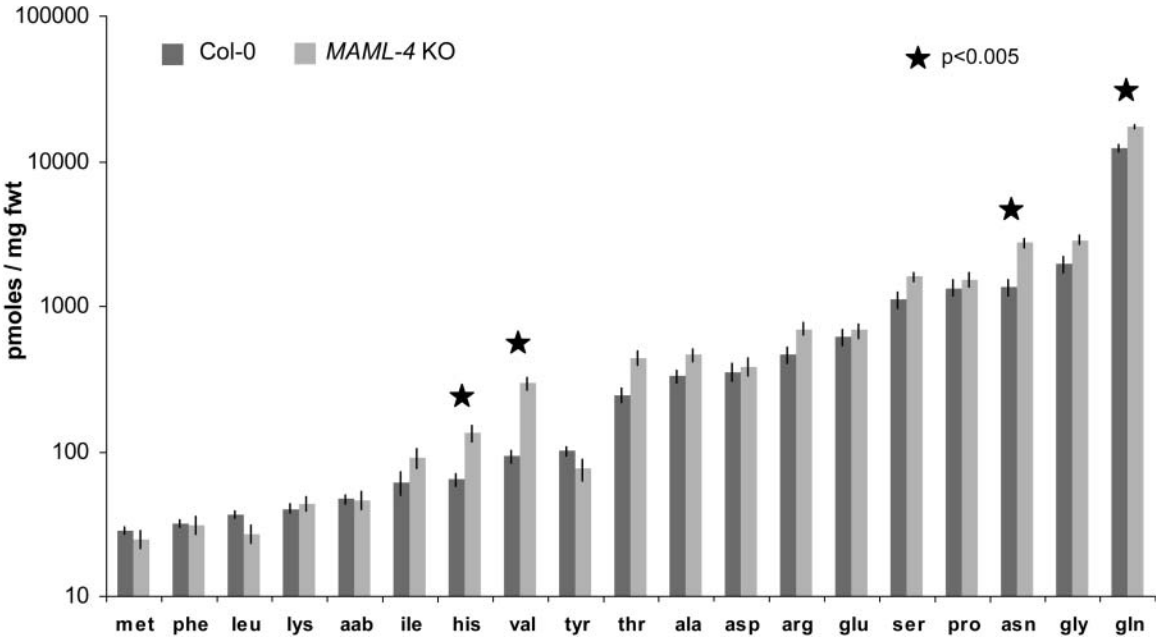

Figure 7. The MAML-4 KO has significantly altered soluble amino acid composition. Mean soluble amino acid content of wild-type Col-0 (dark gray box) and MAML-4 KO (light gray box) 9-d-old seedlings. SE bars are present for each point ($n = 4$). Stars indicate changes that are highly significant ($P < 0.005$). Cys could not be reliably resolved, and Trp was below the detection threshold. Standard amino acid abbreviations are used. aab, α-amino-butyrate.

an IPMS gene in planta (Fig. 2). No IPMS activity was detected for *MAM1*, *MAML-4*, or *MAML*. While this is not definite proof for lack of IPMS activity, as eukaryotic posttranslational processing may be required, it suggests that these genes are less likely to function as IPMS in planta compared to *MAML-3* and may therefore be involved in GSL biosynthesis. This result is contrary to a previous report that claimed that heterologously expressed *MAM1* and *MAML* possessed IPMS activity, although no supporting data were provided (Junk and Mourad, 2002). As the activity of *MAM1* had been previously studied (Kroymann et al., 2001), we obtained knockout lines for *MAML* and *MAML-4*. As part of characterization of these lines, it was interesting to note the relatively poor expression of *MAM1* in Col-0 compared to the other members of this gene family (Fig. 3b).

Through both analysis of MAML KO and overexpression of *MAML* in MAML KO and wild-type Col-0, we have shown that the *MAML* gene is required for 6C, 7C, and 8C GSL synthesis, suggesting that its product can catalyze the condensation of 2-oxo-8-methylthiooctanoate and longer homologs with acetyl CoA (Fig. 8). The ability of *MAML* to catalyze multiple condensation reactions is analogous to an enzyme in *Methanococcous jannaschii* that can catalyze condensation of a series of elongated 2-oxo acids (Howell et al., 1998). Unexpectedly, 3C GSL also were significantly enhanced in 35S::MAML plants, although their levels were not altered in either the original MAML KO plants (compared to Col-0 and MAML KO transformed with empty vector in Table II). This suggests that while the product of *MAML* is unlikely to catalyze the initial 2-oxo-3-methylthiobutanoate/acetyl CoA condensation reaction, overexpression perturbs the condensation of acetyl CoA with 2-oxo-5-methylthiopentanoate, possibly interfering with the activity of the *MAM1* gene product. 5C GSLs are still detectable in the MAML KO (Fig. 4b), suggesting that other MAM-like genes, probably *MAM1*, encode products capable of catalyzing the 2-oxo-6-methylthiohexanoate/acetyl CoA condensation.

Knocking out *MAML* led to the loss of 6C, 7C, and 8C GSLs, and an equivalent increase in 4C and 5C

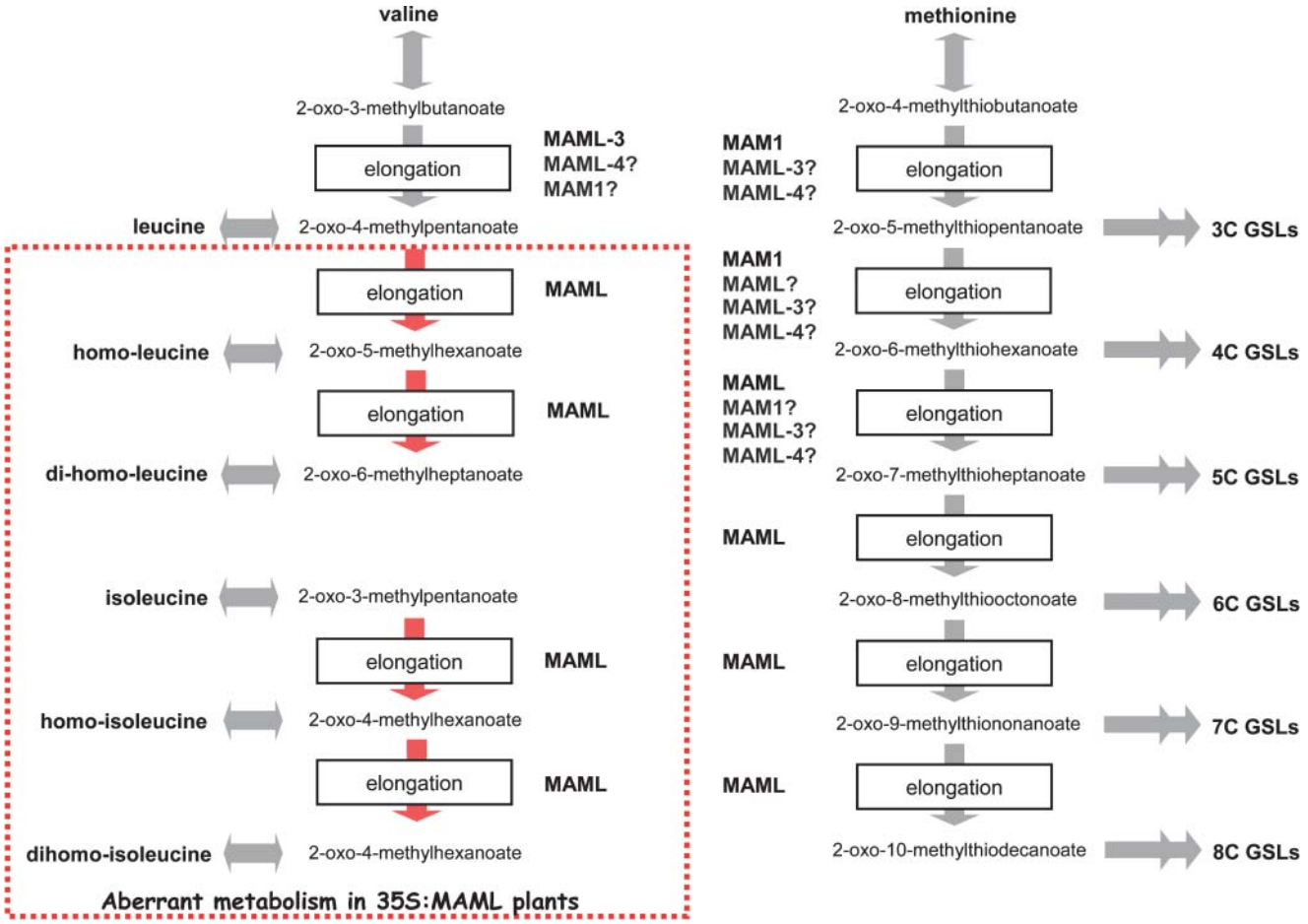

Figure 8. A model for the control of 2-oxo acid elongation by MAM-like genes in Arabidopsis. Elongation consists of the three reactions shown in Figure 1a. MAM-like gene-products only catalyze the initial condensation reaction. Names in bold indicate steps where direct evidence exists for the involvement of that MAM gene-product in elongation, while names in gray indicate steps where only indirect evidence exists. Reactions in the red box occur only in 35S::MAML plants.

GSLs (Fig. 3b; Table II), resulting in no overall change in total Met-derived GSL content. Likewise, overexpression of *MAML* led to an increase in 3C, 6C, 7C, and 8C GSLs and a reduction in 4C and 5C GSLs, again with no change in overall amounts. This indicates that *MAML* does not alter the flux of Met homologs into GSL biosynthesis (possibly by 2-oxo-3-methylthiobutanoate/acetyl CoA condensation) but only the extent of subsequent 2-oxo acid elongation (Fig. 8). The total levels of Met-derived GSLs found in the analyses of the wild types and primary transgenics (Table II), and those found in subsequent analyses of transgenic plants and outsegregants (Fig. 5), are quite different, but are within the range of variation observed previously.

When *MAML* was overexpressed in Arabidopsis Col-0 two novel amino acids were detected in 35S::MAML seedlings with mass 145 (Fig. 6, a and b). On the basis of HPLC and MS we positively identified one of these as homoleucine, and it is highly likely that the other, which has an identical mass and fragmentation pattern but a slightly different retention time, is isohomoleucine (Fig. 6, d and e). The presence of these novel amino acids strongly indicates that *MAML* is capable of initiating elongation of the Leu intermediate, 2-oxo-4-methylpentanoate, and the Ile intermediate, 2-oxo-3-methylpentanoate (Fig. 8). Indeed, both intermediates are 2-oxo acids and possess five-carbon backbones and proximal methyl groups, as does 2-oxo-5-methylthiopentanoate in Met elongation for GSL biosynthesis (Fig. 1, a and b). Moreover we detected low levels of a third amino acid, with a mass consistent with dihomoleucine and dihomoisoleucine indicating that *MAML* can initiate multiple rounds of Leu/Ile elongation, in an analogous manner to Met elongation (Fig. 6c). The absence of homoleucine and homoisoleucine from wild-type plants indicates that under normal conditions *MAML* is confined to GSL biosynthesis, but when spatially, temporally, and/or quantitatively misexpressed *MAML* may promiscuously initiate elongation of oxo acids, leading to aberrant metabolism (Fig. 8). Overexpression or misexpression of *MAM1* may also result in novel amino acid biosynthesis.

Overexpression of *MAML* also caused a significant decrease in soluble Tyr, without a concomitant decrease in the related amino acid Phe. This change cannot be explained by current models of amino acid biosynthesis.

Although we can now ascribe functions in GSL synthesis to *MAM1* (Kroymann et al., 2001) and *MAML* (this work), the function of *MAML*-4 remains unclear. The *MAML-4* KO showed significant alterations in soluble His, Asn, and Gln (Fig. 7), and an increase in Val. There was also a decrease in Leu, but at a lower level of significance ($P = 0.057$). If *MAML-4* functions as an IPMS in planta, despite lack of activity when heterologously expressed in *E. coli*, we may expect an increase in Val and a decrease in Leu. Thus, while the results are inconclusive as to IPMS activity, they do indicate that firstly *MAML-4* plays a role, if undefined, in amino acid biosynthesis, and secondly, perturbing *MAML-4* has effects on several other amino acids. These changes may be the pleiotropic consequences of a nonfunctional IPMS. Indeed, Zhu and Galili (2003) recently showed that modification of Lys biosynthesis in Arabidopsis also caused unexpected increases in levels of soluble His, Gln, and Asns, precisely the same amino acids that are perturbed in our study. Together, these data may be evidence for a plant general amino acid control mechanism. In yeast, general amino acid control is a system that regulates amino acid biosynthesis on a global scale (Hinnebusch, 1992). Thus, starving yeast of one amino acid can result in derepression of the biosynthetic pathways for multiple amino acids and a concomitant 2- to 10-fold increase in the pool sizes of multiple amino acids.

There remains the question of which enzymes in Arabidopsis catalyze the initial 2-oxo-3-methylthiobutanoate condensation in Arabidopsis ecotypes with impaired *MAM1* function. While *MAM2* is a strong candidate, M*AML-3* and *MAML*-4 may also be involved. As we show that the *MAML* gene can function both in long-chain GSL biosynthesis and in the synthesis of long-chain forms of Leu, it is possible that both *MAML-3* and *MAML-4*, in addition to *MAM2*, can also function in short chain GSL biosynthesis and in amino acid biosynthesis (Fig. 8). A degree of redundancy in genes determining this reaction may explain why no null Met-derived GSL mutants have been described in Arabidopsis.

MATERIALS AND METHODS

Bioinformatic Analyses

The CD search at the National Center for Biotechnology Information (Altschul et al., 1997; www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) was used to identify strongly conserved LeuA motifs in the predicted proteins of each MAM synthase. TargetP (Emanuelsson et al., 2000; www.cbs.dtu.dk/services/TargetP/) was used to identify likely chloroplast target peptides ($P > 0.9$) and to predict their cleavage sites: *MAM1* and *MAML* after 49 amino acids, *MAML-3* after 46 amino acids, and *MAML-4* after 57 amino acids. ClustalW at EBI (www.ebi.ac.uk/clustalw/) was used to perform identity comparisons between the four Arabidopsis MAM proteins (lacking their degenerate target peptides) and the *Escherichia coli* LeuA (National Center for Biotechnology Information accession no. AAC73185).

Plant Growth

Plants were routinely grown in Arabidopsis mix (2 parts Levington's M3 potting compost to 1 part grit/sand) under standard glasshouse conditions at approximately 20°C. For aseptic growth, seeds were surface-sterilized and plated on growth medium (1× Murashige and Skoog salts plus vitamins [Duchefa, Haarlem, the Netherlands], 1% Suc, 0.8% agar [BACTOAGAR, Fisher Chemicals, Loughorough, UK], and 2.5 mM MES, pH 5.7). The seeds were stratified for 2 d at 4°C in the dark before germination in a growth room (16 h light/8 h dark, 20°C).

RNA Extraction and cDNA Synthesis

Leaf tissue samples were ground in liquid nitrogen, and total RNA was extracted using the RNeasy plant mini kit (Qiagen, Crawley, UK) and eluted in 40 μL of diethyl pyrocarbonate treated water. Contaminating DNA was

removed by DNase treatment using the Ambion DNA-*free* kit (Huntingdon, UK). Five micrograms of total RNA was used to make first strand cDNA using SuperScript II (Invitrogen, Paisley, UK) in a 20-μL reaction with oligo(dT) primers according to the manufacturer's instructions. The completed reaction was diluted 50-fold, and in subsequent PCR 1 μL of the dilution was used per 10 μL of reaction mix.

cDNA Cloning

MAM1, *MAML*, *MAML-3*, and *MAML-4* lacking their predicted leader-peptide sequence were PCR amplified from Col-0 leaf cDNA in a 50-μL reaction containing 5 μL of cDNA dilution and 2 units of PfuUltra (Stratagene, Amsterdam), 1 $\times$ supplied buffer, 0.3 μM each primer, and 0.2 mM dNTPs. An initial denaturation step of 96°C for 2 min was followed by 30 cycles of 94°C for 10 s, 55°C for 15 s, and 72°C for 2 min. Finally the products were extended by incubation at 72°C for 10 min. *MAM1* was amplified using the primers *MAM1*/49*Sac*I and *MAM1*/TGA*Xho*I; *MAML* with *MAML*/49*Sac*I and *MAML*/TGA*Xho*I; *MAML-3* with *MAML-3*/46*Bgl*II and *MAML-3*/TGA*Pst*I; and finally *MAML-4* with *MAML-4*/57*Sac*I and *MAML-4*/TGA*Xho*I. After amplification the *MAML-3* product was digested with *Bgl*II *and Pst*I, while the *MAM1*, *MAML*, and *MAML-4* products were digested with *Sac*I and *Xho*I and gel purified. The products were then ligated into the inducible expression vector pQE30 (Qiagen) digested with *Bam*HI and *Pst*I in the case of *MAML-3* to give pQE-*MAML-3*, or *Sac*I and *Sal*I in the case of *MAM1*, *MAML*, and *MAML-4* to give pQE-*MAM1*, pQE-*MAML* and pQE-*MAML-4*. The four constructs and the empty vector were used to transform *E. coli* strain m15 (Qiagen) containing the Lac repressor plasmid pREP4. The constructs were isolated and sequenced to ensure no mutations had been introduced.

E. coli CV512 Complementation

Plasmids pQE-*MAM1*, pQE-*MAML*, pQE-*MAML-3*, pQE-*MAML-4*, and an empty vector control were transferred into the Leu auxotrophic *E. coli* strain CV512 obtained from the CGSC *E. coli* Genetic Stock Centre (CGSC no. 5539) and containing pREP4. To test for complementation CV512 containing each of the constructs grown on Luria-Bertani medium was streaked onto plates of solid M9 media (1 $\times$ M9 salts [48 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.5 mM NaCl, and 18.7 mM NH_4Cl]; 2% [w/v] Glc; 1 mM thiamine; 1 mM $MgSO_4$; 0.1 mM $CaCl_2$; kanamycin, 50 μg mL^{-1}; and carbenicillin, 100 μg mL^{-1}), M9 media supplemented with 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG), and M9 media supplemented with 0.3 mM Leu. The plates were incubated overnight at 37°C before being photographed. This was performed at least three times. CV512 containing each construct was also used to inoculate the same liquid M9 medium to an OD_{600} of 0.01 and incubated at 37°C with shaking. Growth curves were obtained by measuring the OD_{600} at the times indicated in Figure 2b. The growth curves were repeated three times and the average OD_{600} used.

Protein Analysis

CV512 containing pQE-*MAM1*, *MAML*, *MAML-3*, *MAML-4*, or an empty vector control was used to inoculate 10 mL of Luria-Bertani medium and induced with 0.1 mM IPTG at mid-log phase. After 2 h total protein was extracted, separated by SDS-PAGE (Laemmli, 1970), immunoblottted with an anti-His horseradish peroxidase conjugated antibody (Invitrogen), and detected with the SuperSignalWest Pico chemiluminsecent system (PerBio, Tattenhall, UK) according to the manufacturer's instructions.

Isolation of MAM Insertion Lines

The MAML insertion line, MAML KO, was identified by PCR screening an En-1 mutagenized population of Arabidopsis, ecotype Col-0 (Baumann et al., 1998; Wisman et al., 1998a, 1998b) using forward and reverse primers specific to the *MAML* gene (*MAML*/F and *MAML*/R) and the *En-1* transposable element (EN/F and EN/R). After back-crossing once to Col-0 a line was isolated that contained a single copy of the En-1 transposon, inserted into *MAML*. The *MAML-4* T-DNA insertion line, Garlic1175_E02, was identified by screening the SAIL/GARLIC T-DNA insertion population of Arabidopsis ecotype Col-0 (Sessions et al., 2002). Homozygous individuals were generated for each line, and compared with the relevant wild type in all subsequent analyses. Col-0 and ecotype Wassilewskija seed was obtained from the Nottingham Arabidopsis Stock Centre (UK).

DNA Extraction and PCR Analysis of Insertion Sites

DNA was extracted from the leaves of each insertion line using a modified version of the cetyl-trimethyl-ammonium bromide method (Lister et al., 2000). Sequencing was performed using the ABI BigDye dye terminator system (Perkin-Elmer Applied Biosytems, Foster City, CA). The T-DNA insertion site for Garlic1175 was confirmed by PCR and sequencing using T-DNA primers (GARLICLB1 and GARLICRB1) and *MAML-4* specific primers (*MAML-4*/R1 and *MAML-4*/L1). The *MAML*::En insertion site was further analyzed by PCR and sequencing using various *MAML* specific primers and *En-1* specific primers (not listed).

Semiquantitative RT-PCR Analysis

For semiquantitative RT-PCR analysis, five microliters of cDNA dilution was used in a 50-μL PCR mixture containing 1 unit Taq polymerase (Invitrogen) and 0.15 μM each primer; the reaction was allowed to proceed for 18 cycles. The primer pairs used are shown below. The constitutively expressed housekeeping gene adenine phosphoribosyltransferase (APT) was used as an mRNA loading control (Moffatt et al., 1994). Each reaction was repeated at least three times independently. The products were then blotted onto a positively charged nylon membrane (Hybond-N+, Amersham Pharmacia, Uppsala) and probed with ^{32}P-random-labeled probes according to established procedures. Probes were derived from plasmids (pGEMTEasy, Promega, Southampton, UK) that contained the sequenced product from each primer pair used in the RT-PCR reaction.

GSL Analysis

GSLs were extracted from 300 mg of seeds, converted to desulphoglucosinolates, and analyzed by LC-MS with atmospheric pressure chemical ionization, as previously described (de Quiros et al., 2000).

MAML Cloning and Arabidopsis Transformation

Transgenic Arabidopsis were generated by *Agrobacterium tumefaciens* mediated transformation with the T-DNA vector pTKC28 through floral dipping (Clough and Bent, 1998) and selection of the seedlings for BASTA resistance. The Agrobacterium host was a rifampicin resistant derivative of C58 containing a nononcogenic Ti plasmid, supplying the virulence functions and an intermediate vector containing the genes of interest between the T-DNA borders. These consisted of the bialophos resistance gene (bar) from *Streptomyces hygroscopius* driven by the Arabidopsis small subunit Rubisco promoter as a marker gene, and the *MAML* gene consisting of a Sfc/*Hin*cII 3,379 bp genomic fragment derived from bacterial artificial chromosome clone T20O7 (Arabidopsis Biological Resource Center, Columbus, OH) containing the complete open reading frame (including introns) plus 339 bp downstream of the stop codon driven by the 35S CaMV promoter.

Soluble Amino Acid Analysis

A total of 100 mg of 9 d-old seedlings was homogenized in 1 mL of 45°C 70% methanol and incubated at 45°C for 10 min. The supernatant was removed and the pellet extracted twice more and the supernatants pooled. The supernatants were dried under a fixed nitrogen line at 45°C and resuspended in 0.02 M HCl. The extract was then filtered through an ultrafree-MC 0.22-μm filter column (Millipore, Bedford, MA). Ten microliters of eluate was derivatized with 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate using the AccQ Tag system (Waters, Milford, MA). The derivatized amino acids were separated on a Waters Alliance 2695 Separation Module through a reverse phase AccQ Tag column (Waters) at 37°C using a 65-min gradient of sodium acetate buffer (0.1 M sodium acetate pH 5.80, 2.7 μM EDTA, and 6.9 mM triethylamine), acetonitrile, and water at a flow rate of 1 mL min^{-1}. Derivatized amino acids were detected by excitation at 250 nm and emission at 395 nm using a Waters 474 scanning fluorescence detector. Millenium^{32} Chromatography Manager software (Waters) was used to analyze the data. Individual amino acids were identified and quantified using a calibration curve generated by the injection of standards of known concentrations. Four extractions were performed for each line analyzed to ensure reproducibility.

For LC-MS/MS analysis of derivatized amino acids, derivatized extracts were separated on a Thermofinnigan Surveyor HPLC (Hemel Hempstead, UK) through a reverse phase Luna C18 column (Phenomonex, Cheshire, UK) using a gradient of solvent A (0.1% formic acid in water) and solvent B (0.1% formic acid in methanol) at a flow rate of 300 μL/min throughout. Two gradients were used. Gradient 1 was as follows: 30% solvent B at start, linear gradient to 70% solvent B over 26.6 min, linear gradient to 97% solvent B over 2.4 min, linear gradient to 30% solvent B over 0.2 min, and 30% solvent B for 6.8 min. Gradient 2 was as follows: 5 min at 7% solvent B, linear gradient to 40% solvent B over 20 min, linear gradient to solvent 97% B over a further 20 min, and 97% solvent B for 5 min. Between runs the column was reequilibrated with 7% solvent B for 10 min. Derivatized amino acids were detected by UVA_{250}.

For LC MS/MS analysis of underivatized amino acids, untreated extracts were separated on a Thermofinnigan Surveyor HPLC through a reverse phase Luna C18 column using a gradient of solvent A (0.1% heptafluorobutyric acid in water; Sigma, St. Louis) and solvent B (acetonitrile) at a flow rate of 300 μL min^{-1} throughout. The gradient was as follows: 10% solvent B at start, linear gradient to 25% solvent B over 20 min, linear gradient to 40% solvent B over 2.5 min, linear gradient to 10% solvent B over 0.5 min. Between runs the column was reequilibrated with 10% solvent B for 10 min. After column separation, amino acids in the derivatized and underivatized extracts were detected by positive mode electrospray ionization in a Thermofinnigan LCQ DecaXP ion-trap mass spectrometer. The source conditions were 5.2 kV source voltage, 350°C capillary temperature, 50 units sheath gas, no auxiliary gas. Ions of interest were selected for fragmentation using an isolation width of 2 *m/z* and collision energy of 35% without wideband activation.

Primers

All primers were synthesized by Sigma Genosys (Cambridge, UK).
For cDNA cloning:

MAM1/49*Sac*I 5′-ACGAGCTCTGCTCCGCTGAGTCCAAA-AAG-3′
MAM1/TGA*Xho*I 5′-GACCTCGAGCCAAACTTATACAACAG-CGGAAA-3′
MAML/49*Sac*I 5′-ACGAGCTCTGCTCTTCTGTGTCCAAAA-ATG-3′
MAML/TGA*Xho*I 5′-GACCTCGAGCGTGTTTACACATTCGAT-GAAA-3′
MAML-3/46*Bgl*II 5′-ACAGATCTCTTACCACCGCCGGAAAA-TTC-3′
MAML-3/TGA*Pst*I 5′-ACCTGCAGTTTCTTCAGGCAGGGACT-TC-3′
MAML-4/57*Sac*I 5′-ACGAGCTCTGCTCAATCTCAGATCCTTC-TC-3′
MAML-4/TGA*Xho*I 5′-GACCTCGAGTTCAGGCAGCGACTCT-GTT-3′

Insertion site analysis:

GARLICLB1 5′-GCCTTTTCAGAAATGGATAAATAGCCTTGC-TTCC-3′
MAML/F 5′-TATGCGAAGAGGCCGAGGGTAATG-3′
MAML/R 5′-CTTTACCATACCCTGCCGACACATACC-3′
EN/F 5′-AGAAGCACGACGGCTGTAGAATAGGA-3′
EN/R 5′-GAGCGTCGGTCCCCACACTTCTATAC-3′

For RT-PCR analysis:

MAM1/L1 5′-ACCACTAGCTGTCGCTCCAT-3′
MAM1/R1 5′-CTTGGCGATGGTTTTAATAGC-3′
MAML/L1 5′-CCGGTCAGTGTTACCCTTTC-3′
MAML/R1 5′-TTGGCGATGGTCTTAATGGT-3′
MAML-3/L1 5′-CTTATCCTCCTCCTCCACCT-3′
MAML-3/R1 5′-CAGGGACATAGCCATTTTCG-3′
MAML-4/L1 5′-TTTCCGTTTCCAACCATCTC-3′
MAML-4/R1 5′-TCCTAGCGATTTCGATGACC-3′
APT/F 5′-TCCCAGAATCGCTAAGATTGCC-3′
APT/R 5′-CCTTTCCCTTAAGCTCTG-3′

Upon request, all novel materials described in this publication will be made available in a timely manner for noncommercial research purposes, subject to the requisite permission from any third-party owners of all or parts of the material. Obtaining any permissions will be the responsibility of the requestor.

ACKNOWLEDGMENTS

We thank Lionel Hill for the LC-MS/MS analyses, and Konstantinos Petritis for help with separating underivatized amino acids. We also thank Alison Smith and Steve Rawsthorne for proof reading preliminary drafts.

Received January 20, 2004; returned for revision March 1, 2004; accepted March 1, 2004.

LITERATURE CITED

Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:** 3389–3402

Baumann E, Lewald J, Saedler H, Schulz B, Wisman E (1998) Successful PCR-based reverse genetic screens using an En-1-mutagenised Arabidopsis thaliana population generated via single-seed descent. Theor Appl Genet **97:** 729–734

Chisholm M, Wetter L (1964) Biosynthesis of mustard oil glucosides. IV. The administration of methionine ^{14}C and related compounds to horseradish. Can J Biochem **42:** 1033–1040

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J **16:** 735–743

de Quiros HC, Magrath R, McCallum D, Kroymann J, Scnabelrauch D, Mitchell-Olds T, Mithen R (2000) Alpha-keto acid elongation and glucosinolate biosynthesis in Arabidopsis thaliana. Theor Appl Genet **101:** 429–437

Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol **300:** 1005–1016

Evans DJ, Jones R, Woodley PR, Wilborn JR, Robson RL (1991) Nucleotide sequence and genetic analysis of the *Azootobacter chlorooccum* nifUSVWZM gene cluster, including a new gene (*nifP*) which encodes a serine acetyltransferase. J Bacteriol **173:** 5457–5469

Graser G, Schneider B, Oldham NJ, Gershenzon J (2000) The methionine chain elongation pathway in the biosynthesis of glucosinolates in Eruca sativa (Brassicaceae). Arch Biochem Biophys **378:** 411–419

Hagelstein P, Schultz G (1993) Leucine synthesis in spinach-chloroplasts: partial characterization of 2-isopropylmalate synthase. Biol Chem Hoppe Seyler **374:** 1105–1108

Haughn GW, Davin L, Giblin M, Underhill EW (1991) Biochemical genetics of plant secondary metabolites in Arabidopsis thaliana: the glucosinolates. Plant Physiol **97:** 217–226

Hinnebusch AG (1992) General and pathway-specific regulatory mechanisms controlling the synthesis of amino acid biosynthetic enzymes in *Saccharomyces cervisiae. In* EW Jones, JR Pringle, JR Broach, eds, The Molecular and Cellular Biology of the Yeast *Saccharomyces*, Vol 2, Gene Expression. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 319–411

Howell DM, Harich K, Xu HM, White RH (1998) Alpha-keto acid chain elongation reactions involved in the biosynthesis of coenzyme B (7-mercaptoheptanoyl threonine phosphate) in methanogenic archaea. Biochemistry **37:** 10108–10117

Junk DJ, Mourad GS (2002) Isolation and expression analysis of the isopropylmalate synthase gene family of Arabidopsis thaliana. J Exp Bot **53:** 2453–2454

Kroymann J, Donnerhacke S, Schnabelrauch D, Mitchell-Olds T (2003) Evolutionary dynamics of an Arabidopsis insect resistance quantitative trait locus. Proc Natl Acad Sci USA **100:** 14587–14592

Kroymann J, Textor S, Tokuhisa JG, Falk KL, Bartram S, Gershenzon J, Mitchell-Olds T (2001) A gene controlling variation in arabidopsis glucosinolate composition is part of the methionine chain elongation pathway. Plant Physiol **127:** 1077–1088

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:** 680–685

Lister C, Anderson M, Dean C (2000) Genetic mapping using recombinant inbred lines. *In* Z Wilson, ed, *Arabidopsis:* A Practical Approach. Oxford University Press, Oxford, pp 51–75

Mithen R, Faulkner K, Magrath R, Rose P, Williamson G, Marquez J (2003) Development of isothiocyanate-enriched broccoli, and its enhanced ability to induce phase 2 detoxification enzymes in mammalian cells. Theor Appl Genet **106:** 727–734

Moffatt BA, Mcwhinnie EA, Agarwal SK, Schaff DA (1994) Adenine phosphoribosyltransferase-encoding gene of Arabidopsis-thaliana. Gene **143:** 211–216

Sessions A, Burke E, Presting G, Aux G, McElver J, Patton D, Dietrich B, Ho P, Bacwaden J, Ko C, et al (2002) A high-throughput Arabidopsis reverse genetics system. Plant Cell **14:** 2985–2994

Somers JM, Amzallag A, Middleton RB (1973) Genetic fine structure of the leucine operon of Escherichia coli K-12. J Bacteriol **113:** 1268–1272

Umbarger HE (1997) Biosynthesis of branched-chain amino acids. *In* FC Niedhardt, ed, *Escherichia coli* and *Salmonella* Cellular and Molecular Biology, Ed 2, Vol 1. ASM Press, Washington DC, pp 422–456

Wisman E, Cardon GH, Fransz P, Saedler H (1998) The behaviour of the autonomous maize transposable element En/Spm in Arabidopsis thaliana allows efficient mutagenesis. Plant Mol Biol **37:** 989–999

Wisman E, Hartmann U, Sagasser M, Baumann E, Palme K, Hahlbrock K, Saedler H, Weisshaar B (1998) Knock-out mutants from an En-1 mutagenized Arabidopsis thaliana population generate phenylpropanoid biosynthesis phenotypes. Proc Natl Acad Sci USA **95:** 12432–12437

Zhu X, Galili G (2003) Increased lysine synthesis coupled with a knockout of its catabolism synergistically boosts lysine content and also transregulates the metabolism of other amino acids in Arabidopsis seeds. Plant Cell **15:** 845–853