

## Research Article

# Changing substrate specificity and iteration of amino acid chain elongation in glucosinolate biosynthesis through targeted mutagenesis of *Arabidopsis* methylthioalkylmalate synthase 1

- ♠ Annette Petersen<sup>1,2,\*</sup>, Lea Gram Hansen<sup>3,\*</sup>, Nadia Mirza<sup>4,\*</sup>, Christoph Crocoll<sup>1,2</sup>, Osman Mirza<sup>5</sup> and
- **Darbara Ann Halkier**

Correspondence: Barbara Ann Halkier (bah@plen.ku.dk)



Methylthioalkylmalate synthases catalyse the committing step of amino acid chain elongation in glucosinolate biosynthesis. As such, this group of enzymes plays an important role in determining the glucosinolate composition of Brassicaceae species, including Arabidopsis thaliana. Based on protein structure modelling of MAM1 from A. thaliana and analysis of 57 MAM sequences from Brassicaceae species, we identified four polymorphic residues likely to interact with the 2-oxo acid substrate. Through site-directed mutagenesis, the natural variation in these residues and the effect on product composition were investigated. Fifteen MAM1 variants as well as the native MAM1 and MAM3 from A. thaliana were characterised by heterologous expression of the glucosinolate chain elongation pathway in Escherichia coli. Detected products derived from leucine, methionine or phenylalanine were elongated with up to six methylene groups. Product profile and accumulation were changed in 14 of the variants, demonstrating the relevance of the identified residues. The majority of the single amino acid substitutions decreased the length of methionine-derived products, while approximately half of the substitutions increased the phenylalanine-derived products. Combining two substitutions enabled the MAM1 variant to increase the number of elongation rounds of methionine from three to four. Notably, characterisation of the native MAMs indicated that MAM1 and not MAM3 is responsible for homophenylalanine production. This hypothesis was confirmed by glucosinolate analysis in mam1 and mam3 mutants of A. thaliana.

\*These authors contributed equally to this work.

Received: 21 December 2018 Revised: 01 May 2019 Accepted: 15 May 2019

Accepted Manuscript Online: 07 June 2019 Version of Record published: 02 July 2019

## Introduction

Glucosinolates are amino acid-derived chemical defence compounds characteristic of the Brassicales order [1]. In biosynthesis of methionine-derived aliphatic glucosinolates, the chain elongation process generates side chain-elongated methionine derivatives as precursor for the corresponding glucosinolates [2]. In the iterative process, one methylene group is added in each cycle (Figure 1). This process evolved from the non-iterative leucine biosynthesis in which one methylene group is added to the side chain of valine [3–6]. The methionine elongation process has uniquely evolved the ability to add up to nine methylene groups [7]. Interestingly, the glucosinolate structural diversity provides differences in biological function with impact on disease resistance and plant fitness [8].

<sup>&</sup>lt;sup>1</sup> DynaMo Center, Department of Plant and Environmental Sciences, Faculty of Science, University of Copenhagen, Frederiksberg, Denmark; <sup>2</sup>Copenhagen Plant Science Centre, Department of Plant and Environmental Sciences, Faculty of Science, University of Copenhagen, Frederiksberg, Denmark; <sup>3</sup>The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark; <sup>4</sup>NanoTemper Technologies ApS, Taastrup, Denmark; <sup>5</sup>Department of Drug Design and Pharmacology, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark



Figure 1. Schematic view of the chain elongation pathway in glucosinolate biosynthesis

The first step is deamination by a branched-chain aminotransferase (BCAT), followed by a three-step cycle: condensation by a MAM, isomerisation by an isopropylmalate isomerase (IPMI) and oxidative decarboxylation by an isopropylmalate dehydrogenase (IPMDH). At this point the intermediate will either go through another cycle or exit by transamination by a BCAT. The pathway intermediates are (1) amino acid, (2) 2-oxo acid, (3) 2-alkylmalic acid, (4) 3-alkylmalic acid and (5) chain-elongated amino acid.

All enzymes of methionine chain elongation are known in *Arabidopsis thaliana* and the methylthioalkylmalate synthases (MAMs) determine the number of elongation cycles [5,9–11]. Previous studies investigated the evolutionary process from the isopropylmalate synthase (IPMS) to MAM through domain-scale modifications and site-directed mutagenesis. They successfully changed substrate preference of IPMS2 to resemble that of MAM1, thereby identifying a set of residues important for substrate specificity [12].

IPMSs and MAMs both condensate 2-oxo acids with acetyl-CoA, which places them in the Claisen-like condensation (CC-like) subgroup of the DRE-TIM metallolyase superfamily [13]. Two IPMS enzymes have been crystallised from this subgroup: *Mycobacterium tuberculosis* (MtIPMS, [14] and *Neisseria meningitidis* (NmIPMS, [15]. The crystal structures indicate that the IPMS enzymes form homo-dimers. Each monomer has two domains with a catalytic site and the LeuA regulatory domain that inhibits activity in the presence of leucine [16–18]. This regulatory domain is lost in MAM enzymes [12]. Previous studies have shown that the LeuA domain is critical for the quaternary structure of IPMS enzymes. Removal of this domain from IPMS2 resulted in a complete loss of quaternary structure and adding the domain to MAM1 and MAM3 caused the otherwise monomeric enzymes to form both dimers and tetramers [12]. The authors hypothesised that losing the quaternary structure increased the active site of MAM synthases and thereby allowed for bulkier substrates. Individual residues were identified as important for evolution of the MAM synthases. Eleven residues that have been under positive selection during the iterative adaptation of the MAM synthases are Glu96, Ser98, Ile138, Cys165, Thr173, Leu177, Val187, Thr257, Ile258, Gly259 and Val289 according to the numbering in MAM1 of *A. thaliana* [5]. Based on MAM homology models, five residues are proposed to bind the substrate: Arg93, Asp94, Thr261, His292 and His292. Additional six residues are predicted to hydrophobically interact with the substrate: Ile162, Leu186, Gly229, Glu231, Thr257 and Gly259 [12].

Sequencing of the MAM gene cluster of several *Arabidopsis* species reveal three distinct clades of MAMs: MAMa, MAMb and MAMc [5]. In *A. thaliana*, MAM1 and MAM2 are homologues of MAMa, while MAM3 is a homologue of MAMb [10]. MAMc was lost in *A. thaliana* but is still found in e.g. *A. lyrata*. MAM3 has a broader substrate specificity than MAM1, which in turn is broader than MAM2 [5,10,11]. MAM2 catalyses only the first round of elongation and is found in accessions rich in glucosinolates derived from homomethionine (e.g. Ler-0) [5], while MAM1 catalyses C3–C5 and is found in accessions accumulating glucosinolates derived from dihomomethionine



(e.g. Col-0) [9]. MAM3 catalyses all lengths from C3–C8 and thus is present in – but not restricted to – all accessions with glucosinolates derived from tetra-, penta- and hexahomomethionine [10].

Quantitative trait loci mapping has implicated the GS-ELONG locus in not only controlling chain elongation of methionine but also phenylalanine, from which 2-phenylethyl glucosinolate is derived [19]. When characterising MAM3 in *in vitro* assays, chain elongation of methionine as well as leucine, isoleucine and phenylalanine were detected [10]. Based on these findings, MAM3 was proposed to be responsible for elongation of phenylalanine *in planta*. When the chain elongation pathway of *A. thaliana*, i.e. *MAM1*, *BCAT4*, *IPMI-LSU1*, *IPMI-SSU3* and *IPMDH1*, was expressed in *Escherichia coli*, methionine- and leucine-derived products elongated by either one or two methylene groups, were produced [20]. Similar results were obtained when the chain elongation pathway was expressed in *Nicotiana benthamiana* [21,22]. As leucine-derived glucosinolates are not found in the native *A. thaliana* plant, the results indicate that MAM synthases lose specificity when expressed in heterologous hosts.

The Brassicaceae is a large family with over 3700 species divided into 338 genera and 19 tribes [23]. Based on a phylogenetic analysis of 114 sequences within the family, a division into three lineages was proposed [24]. Several whole genome and local tandem duplications combined with subsequent deletions have resulted in a poor resolution of the MAM family in Brassicaceae. Furthermore, predicting activity from sequence data alone can be difficult and imprecise [25], as evident from IPMS enzymes, which show nearly identical substrate specificity despite having a relatively moderate sequence identity [18]. Understanding how the substrate specificity of MAMs controls the iterative chain elongation process is important from a basic science as well as a bioengineering perspective.

The present study investigates the impact of polymorphic residues specific to the MAM enzymes on substrate specificity and number of iterations. A well-established method for this purpose is site-directed mutagenesis on evolutionarily conserved residues [17]. We used this approach to create a library of MAM1 variants designed to have changed substrate specificity. Based on a multiple alignment of MAM synthase sequences and IPMS crystal structures, evolutionary important residues were predicted and targeted for mutation. Since the chain elongation pathway has been successfully expressed in *E. coli* previously [20], we chose to characterise the MAM1 variants *in vivo* using this host. Four residues were mutated individually or in combinations, and a total of 15 variants were characterised. Substrate specificity and number of elongations were compared with the native MAM1, MAM3 and IPMS2 from *A. thaliana*. Targeted proteomics was used to monitor expression of the proteins in *E. coli* and to eliminate changes in product or titres caused by unequal expression of proteins between the different *E. coli* strains.

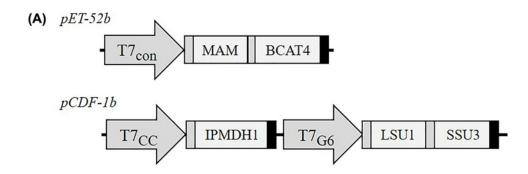
# Materials and methods Bioinformatics and structural modelling

Protein sequences of MAM1, MAM2, MAM3, IPMS1 and IPMS2 from *A. thaliana* were collected from TAIR (Arabidopsis.org). Additional MAM sequences were identified through BLASTp and tBLASTn searches on NCBI using the MAM sequences. MAMs from *Camelina sativa*, *Brassica rapa*, *Brassica oleracea*, *Brassica napus*, *Eutrema salsugineum*, *Capsella rubella*, *Arabidopsis lyrata*, *Arabidopsis cebennensis*, *Boechera divaricarpa*, *Raphanus sativus* and *Eruca vesicaria* were collected and used in an alignment (Supplementary Figures S1 and S2). Sequences with large deletions were removed. Alignments were made using the slow algorithm in CLC Main Workbench 8.0.1 (www.qiagenbioinformatics.com). A homology model of the MAM1 protein structure was constructed based on the *N. meningitis* IPMS crystal structure (PDB-ID 3RMJ) [15] using BLAST [26] to align the two sequences (40.7% identity) and Modeller to create the model [27]. Structural analysis were done in PyMOL (the PyMOL Molecular Graphics System, version 1.5.0.4 Schrödinger, LLC).

# Generation of expression constructs for MAM1 variants

All constructs were cloned by USER cloning [28]. As described in [21], the chain elongation pathway genes were expressed from two plasmids; *BCAT4* and *MAM1* on pET-52b (Novagen®, Merck, #71554) and *IPMI-LSU1*, *IPMI-SSU3* and *IPMDH1* on pCDF-1b (Novagen®, Merck, #71330). Two changes were made to the T7 promoters of *IPMI* and *IPMDH1* compared with previous construct [20]. An alternative start codon was removed between the RBS and *IPMDH1* sequence, and the *IPMI* subunits were expressed from a separate promoter. Point mutations were introduced into *MAM1* by amplification with primers containing the desired mutation in the USER overhang (Supplementary Table S1). The CDS was amplified in two halves and fused by USER fusion [29]. The native *MAM1* sequence was replaced for the mutated MAM versions in the expression construct (Figure 2).





(B)		
	T7 promoter	Sequence
	G6	TAATACGACTCACTA <b>TTTCG</b> GAA
	CC	TAATACGACTCACTA <b>TTGGG</b> GAA
	Consensus	TAATACGACTCACTA <b>TAGGG</b> GAA

Figure 2. Design of expression constructs containing the chain elongation pathway of A. thaliana

(A) The two constructs used in combination to express the biosynthetic genes. The MAM gene is replaced with native or mutated variants of MAM1 and MAM3 as well as IPMS2. Different T7 promoters were used as indicated in the arrows. Ribosomal binding sites, genes and terminators are indicated as dark grey, light grey and black boxes, respectively. (B) Promoters used in this work. Sequence differences are highlighted.

## **Bacterial strains and cultivation**

The *E. coli* NEB10B strain (NEB, #C3019H) was used for plasmid amplification and cloning purposes. The BL21 (DE3) strain (NEB, #C2527I) was used for expression. All transformations were done by heat-shock followed by one hour recovery in non-selective media before plating on LB agar plates (Luria–Bertani broth, Lennox, Duchefa, #L1703) containing 10 g tryptone, 5 g yeast extract, 5 g sodium chloride and 15g bactoagar per litre with 50 µg/ml spectinomycin and 100 µg/ml ampicillin. Expression experiments were performed in 24-well plates (ThermoFisher, #CS15124) containing 4 ml culture. Single colonies containing both chain elongation constructs were inoculated into LB media and grown overnight at 37°C, 220 rpm. Expression cultures were inoculated from the overnight culture into LB media with 50 µg/ml spectinomycin and 50 µg/ml carbenicillin. The cultures were started at a  $D_{600}$  (attenuance) at 0.1 and grown at 37°C, 220 rpm until  $D_{600}$  reached 0.6–0.8. Gene expression was then induced by addition of 1 mM IPTG (isopropyl D- $\beta$ -1 thiogalactopyranoside) and the cultures were moved to 28°C, 220 rpm and harvested after 24 h.

## Metabolite extraction and LC-MS analysis

In *E. coli* cultures expressing the chain elongation pathway, only the medium was analysed for chain-elongated products, as the majority of the product were reported to be exported from the cells [20]. The samples were centrifuged at  $13000 \times g$  and supernatant diluted 25-fold in water, followed by 1:10 mixing with  $^{15}$ N $^{13}$ C-labelled amino acids (Algal amino acids 13C, 15N, Isotec, Miamisburg, U.S.A.) at a concentration of 10 µg/ml. Immediately before analysis, the samples were filtered through Durapore<sup>®</sup> 0.22 µm PVDF filter plates (Merck Millipore, Tullagreen, Ireland).

The LC–MS analysis was performed as previously described [20] with changes as detailed below. Chromatography was performed using an Advance UHPLC system (Bruker, Bremen, Germany) and a Zorbax Eclipse XDB-C18 column ( $100 \times 3.0$  mm, 1.8  $\mu$ m, Agilent Technologies, Germany). The mobile phases A and B were formic acid (0.05% (v/v)) in water and acetonitrile (supplied with 0.05% (v/v) formic acid), respectively. The elution profile was: 0-1.2 min 3% B; 1.2-4.3 min 3-65% B; 4.3-4.4 min 65-100% B; 4.4-4.9 min 100% B, 4.9-5.0 min 100-3% B and 5.0-6.0 min 3% B. Mobile phase flow rate was 500  $\mu$ l/min and column temperature was maintained at 40°C. An EVOQ Elite



Table 1 Peptides of chain elongation enzymes and the *E. coli* housekeeping enzyme (ICD) monitored through targeted proteomics analysis

Enzyme	AGI code	Peptide sequence	Reference	
BCAT4	At3g19710	TGEETLAAK	This study	
		LYETLSDIQTGR*	This study	
		SITNYZPVWIPLAEAK	This study	
		GNVVSTPTIAGTILPGVTR	This study	
MAM1	At5g23010	SGNASLEEVVMALK	This study	
		STYEILSPEDIGIVK	This study	
		DGEQSPGGSLTPPQK	This study	
		SLGFNDIQFGZEDGGR*	This study	
MAM3	At5g23020	GESLMDGVYTK	This study	
		ALVVNGAEISSEK	This study	
		SGNAPLEEVVMALK	This study	
		STYEILSPEDVGIVK*	This study	
		DGEQSPGAALTPPQK	This study	
IPMS2	At1g74040	GTYEIMSPEEIGLER*	This study	
IPMI-LSU1	At4g13430	FILDGEMPSYLQAK*	This study	
		VWMDVYALPVPGAGGK	This study	
IPMI-SSU3	At3g58990	LGSFALNGLPK	This study	
		EDGSSLLINHTTR	This study	
		NCVATGEIFPLESEVR*	This study	
IPMDH1	At5g14200	LSDAILLGAIGGYK*	This study	
		AGSLEGLEFDFK	This study	
		IEDAVVDALNK	This study	
ICD	JW1122	GPLTTPVGGGIR*	[40]	

Peptides marked with asterisk were used for relative quantification.

TripleQuad mass spectrometer with an electrospray ionisation source (ESI) (Bruker, Bremen, Germany) was coupled to the liquid chromatography. Pure standards were used to optimise the instrument parameters. The ion spray voltage was maintained at 3000 V in positive ionisation mode. Cone temperature was set to  $350^{\circ}$ C and cone gas flow to 20 psi. Heated probe temperature was set to  $400^{\circ}$ C and probe gas flow set to 50 psi. Nebulising gas was set to 60 psi and collision gas to 1.6 mTorr. Nitrogen was used as both cone gas and nebulising gas and argon as collision gas.

Multiple reaction monitoring (MRM) was used to monitor analyte parent ion to product ion transitions: MRM transitions for amino acids were chosen as previously described [30]. MRM transitions for chain-elongated amino acids were chosen as described in [20], and comparison to pure standards were used to verify the identification (Supplementary Figures S3 and S4). Both Q1 and Q3 quadrupoles were maintained at unit resolution. Bruker MS Workstation software (Version 8.2.1, Bruker, Bremen, Germany) was used for data acquisition and processing. A mixture of <sup>15</sup>N<sup>13</sup>C-labelled amino acids was used as internal standard for quantification of protein amino acids. Details on transitions and collision energies can be found in Supplementary Table S2. Visualisation and statistical analyses were performed in R studio v1.0.153 [31] using [32–35].

# Protein extraction and targeted proteomics

## Design of reference peptides

Proteotypic peptides were designed for each of the chain elongation enzymes to use as internal reference and a relative quantification of protein expression. Gene sequences were digested *in silico* using the SkyLine software [36]. The resulting list of peptides was then subjected to the following sorting criteria; (1) m/z < 1250, (2) cleavage site with only one arginine or lysine residue and (3) avoiding peptides with methionine or cysteine. All peptides fulfilling these requirements were ordered as synthetic isotopically labelled peptides (JPT, SpikeTides<sup>TM</sup>). Peptides were resuspended by adding 100  $\mu$ l 100 mM ammonium bicarbonate solution (pH 8.5) with 20% (v/v) acetonitrile and incubated gently shaking for 30 min at room temperature. Preferentially, each enzyme was covered by a minimum of two non-neighbouring peptides (Table 1).



## Protein extraction and tryptic digest

Pellet from 1.5 ml culture was harvested by centrifugation and kept at  $-20^{\circ}$ C until extraction. Total protein extraction was performed as previously described [37] and the tryptic digest as described in [38] with the modifications detailed in [39]. Briefly, total protein was obtained by methanol:chloroform extraction, dried in Speed-Vac and dissolved in 100 mM ammonium bicarbonate buffer. Protein concentration was determined by Pierce<sup>TM</sup> BCA Protein Assay Kit (ThermoFisher, #23225). For tryptic digest, 25–50 µg protein was incubated with 1 µg trypsin/Lys-C mix (Promega, #V5073) over night at 37°C. Digest stopped by acidification with trifluoroacetic acid (TFA) and samples were diluted up to 1.5 ml in buffer containing 2% (v/v) acetonitrile and 0.1% (v/v) formic acid. Peptides were purified over Sep-Pak C-18 columns (Waters, Sep-Pak® Vac 1cc 100 mg, #WAT023590) and eluted in 1 ml buffer containing 65% (v/v) acetonitrile and 0.1% (v/v) formic acid. Purified peptides were dried in Speed-Vac (2–4 h, 1000 rpm, max 35°C). Dried peptides were stored at  $-20^{\circ}$ C until analysis. The dried peptides were resuspended just prior to LC–MS/MS analysis in 25 µl buffer containing 2% (v/v) acetonitrile, 0.5% (v/v) formic acid and 0.1% (v/v) TFA spiked with 20 nM isotopically labelled peptide standards (JPT, SpikeTides<sup>TM</sup>) and filtered through 0.22 µm centrifugal filter (#UFC30GV00, Merck, Darmstadt, Germany).

## LC-MS analysis

The gradient was adopted from [40,41] with modifications detailed in [39]. Formic acid (0.1% (v/v)) in water and acetonitrile (supplied with 0.1% (v/v) formic acid) were employed as mobile phases A and B, respectively. The elution profile was: 0–1.0 min 5–10% B; 1.0–3.0 min 10–11% B; 3.0–13.0 min 11–19% B; 13.0–21.0 min 19–27.5% B, 21.0–21.7 min 27.5–24% B; 21.7–22.5 min 34–42% B; 22.5–23.5 min 42–90% B, 23.5–26.9 min 90% B, 26.9–30.0 min 90–5% B and 30.0–34.0 min 5% B. Mobile phase flow rate was 500  $\mu$ l/min and column temperature was maintained at 55°C. Peptide separation was achieved on an Aeris PEPTIDE, XB-C18 column (1.7  $\mu$ m, 2.1 × 150 mm, Phenomenex, Palo Alto, U.S.A.) on an Advance UHPLC-OLE (Bruker Daltonics, Bremen, Germany). The injection volume was 10  $\mu$ l.

The following source settings were used for heated electrospray ionisation: spray voltage 3200 V (positive ionisation mode); cone temperature  $300^{\circ}$ C; cone gas flow 20 psi; heated probe temperature  $350^{\circ}$ C; probe gas flow 40 and nebuliser gas flow 50. Nitrogen was used as cone and probe gas and argon as collision gas. An EVOQ Elite TripleQuad mass spectrometer (Bruker Daltonics, Bremen, Germany) was scanning for parent ion to fragment ion transitions for individual peptides within scheduled 3 min windows. Resolution of the first and third quadrupole was set to  $\pm 1$  Da. Detailed information on peptides including retention times, transitions selected for detection and quantification and collision energies can be found in Supplementary Table S3. Skyline 4.2 were used to manually inspect the acquired chromatograms and calculate ratios between endogenous light and synthetic heavy peptides [36]. Relative quantification was obtained by normalising the ratios to the housekeeping protein isocitrate dehydrogenase (ICD). Results were plotted relative to the MAM1 containing strain. Undetected peptides (N/A) were treated as zero values for analysis. Statistical analyses and visualisation were performed in R studio v1.0.153 (R version 3.4.1) [31] using [32,34,35,42].

## Glucosinolate analysis of A. thaliana mutants

## Plant growth and glucosinolate extraction

Seeds of *A. thaliana* ecotype Col-0 (NASC ID N1092), *mam1* [43] and *mam3* (SALK\_007222) mutants were cultivated in growth champers with long day light conditions (16 h light, 22°C, 19°C at night, 100–120  $\mu$ E, 75% relative humidity). Leaf samples were harvested from 3-week-old plants. Material from three plants (approximately 100 mg) were pooled in each sample and three to seven biological replicates were harvested from each genotype. Glucosinolates were extracted and analysed as desulfo-glucosinolates as previously described in [44] with the 96-well optimisation described in [45] and modifications as outlined in [46]. Briefly, one 3 mm chrome ball was added to a microelution tube (Deltalab, cat. no. 408002) with 300  $\mu$ 1 85% methanol containing 10  $\mu$ M p-hydroxybenzyl glucosinolate as internal standard. Immediately after harvest, the leaf tissue was submerged in the extraction solution and the sample was grinded using a mixer mill (3× 30 s at 30 Hz). Samples were spun 10 min at 2500 rpm and the supernatant was loaded on to a 96-well filter plate (MultiScreenHTS, pore size 0.45  $\mu$ m, EMD Millipore, cat. no. MSHVN 4550) containing 45  $\mu$ 1 DEAE-Sephadex A-25 (GE Healthcare, cat. no. 17-0170-02)that had been equilibrated with 300  $\mu$ 1 milliQ water. The DEAE-Sephadex columns were washed twice with 70% methanol followed by two washes with milliQ water. The desulfo-glucosinolates were released from the column by over-night incubation with a sulfatase solution. On the following day, the desulfo-glucosinolates were eluted in 100  $\mu$ 1 milliQ water and diluted 1:10 in milliQ water prior to LC-MS analysis.



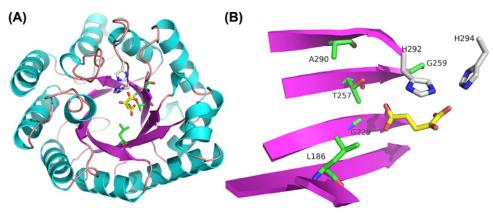


Figure 3. Homology model of MAM1 from A. thaliana

(A) The model of MAM1 was based on the crystal structure of *N. meningitis* IPMS (PDB-ID 3RMJ).  $\beta$  sheets (purple) and  $\alpha$  helixes (turquoise) comprise the  $(\beta/\alpha)_8$  catalytic barrel characteristic of IPMS/MAM family. The binding of IPMS substrate is shown in yellow. Green residues represent the five residues in MAM1 predicted to be within 8 Å of the substrate. Grey residues represent two His residues conserved in the IPMS/MAM family. (B) Simplified presentation of the residues in MAM1, which is predicted to be in close proximity of the substrate. Colours as in (A).

## LC-MS analysis of glucosinolates

The LC–MS analysis was performed using an Advance UHPLC system (Bruker, Bremen, Germany) with a Kinetex XB-C18 column ( $100 \times 2.1$  mm, 1.7 µm, Phenomenex, Palo Alto, U.S.A.) coupled to an EVOQ Elite Triple-Quad mass spectrometer equipped with an electrospray ionisation source (ESI) (Bruker, Bremen, Germany). The method was performed as previously described [46,47] using formic acid (0.05%) in water and formic acid (0.05%) in acetonitrile as mobile phase A and B, respectively. Ionisation was obtained by electrospray ionisation in positive ionisation mode with spray voltage of 3500 V and probe and cone temperatures at 400 and 350°C, respectively. Parent ion to product ion transitions were monitored for one aromatic (2-phenylethyl), three indolic (indol-3-ylmethyl, *N*-methoxy-indol-3-ylmethyl, 4-methoxy-indol-3-ylmethyl) and nine aliphatic (3-methylthiopropyl, 3-methylsulfinylpropyl, 4-methylthiobutyl, 4-methylsulfinylbutyl, 5-methylsulfinylpentyl, 7-methylthioheptyl, 7-methylsulfinylheptyl, 8-methylthiooctyl, 8-methylsulfinyloctyl) glucosinolates. Quantification of individual glucosinolates was done using *p*-hydroxybenzyl glucosinolate as internal reference. Further details and transitions can be found in [46,47].

## Results

## Protein structure modelling and phylogeny analysis of MAM sequences

From a protein sequence alignment of MAM1, MAM3, IPMS1 and IPMS2 from *A. thaliana* ecotype Col-0 and MAM2 from ecotype Ler-0 polymorphic residues specific to MAMs were identified. To date, no MAM enzyme has been crystallised. Since IPMS and MAM enzymes essentially perform the same reaction, we chose to homology model the structure of MAM1 based on a crystallised IPMS (PDB-ID 3RMJ) (Figure 3A). Homocitrate synthases (HCS) are also of the DRE-TIM metallolyase superfamily. Assuming the MAM substrate binds in a similar fashion as in HCS, we retrieved a crystallised HCS (PDB-ID 2ZTJ [48] from *Thermus thermophilus* (TtHCS) with  $\alpha$ -ketoglutarate bound in the catalytic pocket. We superimposed the MAM1 homology model on to 2ZTJ (RMSD of 1.3 Å over 202 residues) and selected MAM1 residues within 8 Å distance of  $\alpha$ -ketoglutarate (Supplementary Figure S5). This subset of residues was used to predict four polymorphic residues (Figure 3B). Three of these sites were previously described as interacting with the substrate (Leu186, Thr257 and Gly259). The remaining residue (Ala290) was included as it fulfilled the criteria of distance to substrate and polymorphism. A fifth residue (Gly229) was within 8 Å of the substrate, but remained constant for all MAM synthases and was thus not targeted for mutation.

A phylogenetic analysis including 57 MAM synthase protein sequences representing 12 species from lineages I and II in the Brassicaceae family was performed. Lineage III is not represented, as to the best of our knowledge no species has been sequenced. This analysis was used to design the substitution mutations so that the mutants represented the natural variation observed at these positions (Table 2 and Supplementary Figure S1). Fifteen substitutions were chosen and residues were mutated in the *A. thaliana* MAM1. In addition, three mutants were constructed, which



Table 2 Polymorphic residues in selected MAM synthases (numbering according to A. thaliana MAM1)

Natural MAM variant	186	257	259	290	
AtiPMS	Н	N	Р	S	
AtMAM1	L	Т	G	Α	
AtMAM2	L	N	А	S	
AtMAM3	M	G	А	Α	
AtbMAM1	L	S	G	Α	
Al/AcMAMb	I/L	G	А	Т	
Cs/CrMAM3	Т	G	А	Α	
Al/AcMAM3/c	С	С	А	S	
Cs/CrMAM3	G/I	G	А	G	
NA	_	V	-	-	
NA	_	D	-	-	

Selected examples of natural variance in the residues targeted for mutation is shown. Species are abbreviated as follows: Ac, A. cebennensis; Al, A. lyrata; At, A. thaliana; Atb: A. thaliana Br-o; Cr, C. rubella; Cs, C. sativa; NA, not found in our analysis but chosen for structural characteristics.

carried two or three mutations in combination. These were inspired from the MAMb clade, which includes MAM3 variants. In this clade, the T257G and G259A mutations are common, and these were therefore combined into the double mutant, GA. Two triple mutants were constructed to mimic specifically MAM3s of *A. thaliana* and *C. sativa*, which are both able to elongate methionine with more than three methylene groups. The mutants were named GAM (T257G/G259A/L186M) and GAG (T257G/G259A/A290G), respectively.

# Metabolite analysis of *E. coli* expressing the chain elongation pathway with MAM variants

## Protein amino acid pools in E. coli expressing the chain elongation pathway

Amino acids used as precursors for the chain elongation pathway based on previous data from mutant plants [49] and heterologous host organisms [10,20] are leucine, methionine, phenylalanine and valine. In the *E. coli* strains expressing the chain elongation pathway with different variants of the MAM synthases, protein amino acid levels were monitored and compared with the levels of *E. coli* strain transformed with empty vector (Figure 4 and Supplementary Figure S6). Generally, the cultures contained approximately three times more phenylalanine than methionine and seven times more leucine than methionine. The levels of leucine, phenylalanine and valine did not statistically differ from the overall mean of all cultures. Methionine levels showed small, but statistically significant differences between cultures (Supplementary Table S4). Slightly reduced levels were seen in the *E. coli* strains producing high levels of methionine-derived products (Figures 4 and 5). However, this did not affect cell density, and thus is likely an effect of the methionine pool being smaller than those of phenylalanine and leucine.

## Levels of chain-elongated amino acids produced by E. coli

By LC-MS analysis of the medium in cultures of E. coli expressing the chain elongation pathway with different MAM variants, we detected nine elongated amino acids derived from one to six elongation cycles: L-homomethionine (HM), L-dihomomethionine (DHM), L-trihomomethionine (TriHM), L-tetrahomomethionine (TetraHM), L-pentahomomethionine (PentaHM), L-hexahomomethionine (HexaHM), L-homoleucine (HL), L-dihomoleucine (DHL), and L-homophenylalanine (HPhe). Additionally two were monitored but not detected: L-trihomoleucine (TriHL) and L-heptahomomethionine (HeptaHM). Strains harbouring the constructs containing the native MAM1 enzyme produced nearly 5-fold more elongated products than those containing MAM3. In strains expressing the G259P variant, we observed an increase in leucine-derived products to 80% of the total products. This was accompanied with a decrease in methionine-derived products and an elimination of HPhe production. In 7 out of 13 strains expressing functional MAM1 variants, the percentage of leucine-derived products increased, while the level of methionine-derived products decreased compared with MAM1. However, only in two of these seven strains did HPhe production decrease, while an increase was seen in the remaining five strains. Similarly, HPhe level was increased in all strains expressing MAM1 variants with lowered levels of leucine-derived products. Two strains - A290S and GAG - showed unchanged composition of chain-elongated amino acids, although the production in GAG was dramatically reduced. Strains harbouring constructs with IPMS2 or the two MAM1 variants (G259A and A290T) did not produce any of the chain-elongated amino acid products (Figure 5). The composition of chain-elongated amino



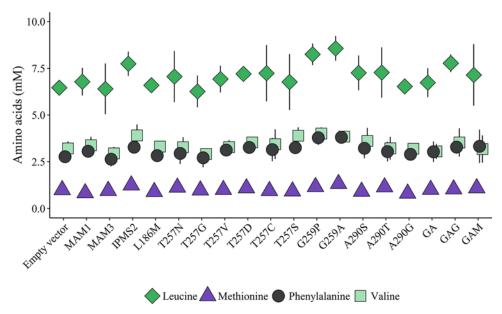


Figure 4. Levels of protein amino acid substrates in *E. coli* expressing the chain elongation pathway with different MAM variants

GA, GAG and GAM represent MAM variants with multiple mutations; GA, T257G/G259A; GAG, T257G/G259A/A290G; GAM, T257G/G259A/L186M. Levels of leucine, phenylalanine and valine are statistically unchanged between the constructs when comparing to the average mean of all cultures. Methionine levels vary between cultures, although the differences are small. Data represent average and standard deviation of three biological replicates each grown in three technical replicates.

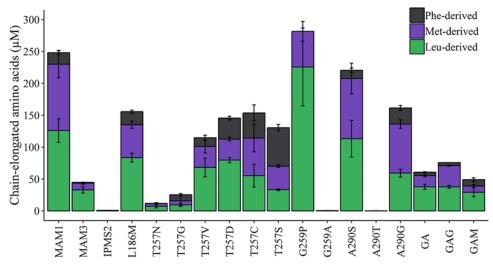


Figure 5. Production of chain-elongated products in all MAM variants

Products are grouped according to precursor amino acid: methionine (purple), phenylalanine (grey) and leucine (green). Methionine-derived products include methionine elongated from one to six cycles, leucine-derived products include leucine elongated once or twice and phenylalanine-derived product is only elongated once. Table shows the percentage of total production within each group of compounds. Data represent average and standard deviation of three biological replicates each grown in three technical replicates.

acids produced by the MAM1 variants indicates that the targeted substitutions have changed the substrate specificity of almost all of them. Overall, MAMa homologues and MAM1 variants mimicking MAMa produced higher levels than either MAMb homologues, MAM1 variants mimicking MAMb or IPMS enzymes (Figure 5).

The number of recursive cycles is affected by the mutations in MAM1 (Figure 6 and Supplementary Figure S7). The



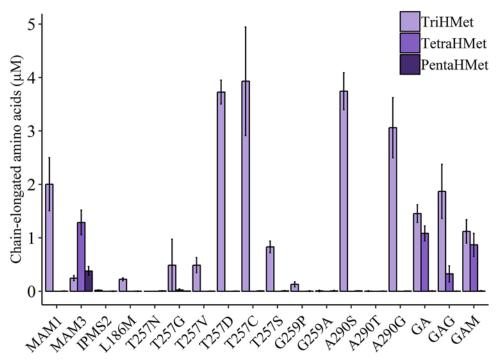


Figure 6. Production of amino acids elongated with more than two methylene groups

MAM3 is the only MAM variant that produces five and six times elongated products. HexaHM levels are below the limit of quantification and thus not included in the graph. Data represent average and standard deviation of three biological replicates each grown in three technical replicates.

native MAM1 enzyme produces predominantly methionine- and leucine-derived products that have been elongated once or twice and low levels of the thrice-elongated methionine (TriHM). This is consistent with previously published results [20]. MAM3 produces a larger variety of products – mostly of once-elongated products (HM and HL), almost no thrice-elongated products (TriHM and TriHL), and substantial amounts of both four and five times elongated methionine (Tetra- and PentaHM). In addition, the six times elongated methionine (HexaHM) was detected below the limits of quantification. Four single mutation MAM1 variants (T257D/C, A290S/G) produced more TriHM than the native MAM1. The GA, GAG and GAM variants produced TetraHM. These were the only variants besides MAM3 that were able to generate products elongated more than three times. Interestingly, MAM1 variants carrying the single mutation (either T257G or G259A) – that are combined in the GA double variant – did not produce amino acids elongated more than twice. Thus, reducing size of the residue Thr257 alone does not allow for bigger substrate – on the contrary, it reduces the number of cycles. When the T257G substitution was combined with a larger residue on Gly259, the ability to elongate methionine three times was restored and even increased to four rounds. This suggests that the binding or interaction with the substrate might have been compromised in the single mutations, but restored when combining them.

# Analysis of *Arabidopsis mam1* and *mam3* mutants for HPhe-derived glucosinolate

Noticeably, both MAM1 and MAM3 were able to produce HPhe, with MAM1 having nearly 20-fold higher levels than MAM3 (Figure 7A). The HPhe levels were affected by mutations in the residues Thr257 and Gly259. Dependent on the substitution, mutations in residue Thr257 can either decrease or increase specificity towards phenylalanine, with T257S showing nearly 3-fold higher HPhe production than the native MAM1. Additionally, mutating Gly259 eliminates production of HPhe. HPhe production in functional variants with mutations in Leu186 and Ala290 were similar to that of MAM1.

The results were contradicting with the previously proposed role of MAM3 [10], but could potentially be an effect of expressing the enzymes in a heterologous system. To investigate this further and determine the role of MAM1 and MAM3 in production of the HPhe-derived 2-phenylethyl glucosinolate *in planta*, we performed glucosinolate



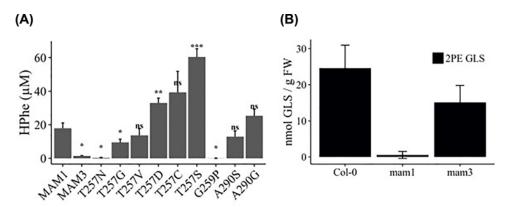


Figure 7. Production of HPhe in E. coli by MAM1 variants and the effect of MAM mutants on HPhe-derived glucosinolates in A. thaliana

(A) Production of HPhe in MAM variants with changed specificity toward phenylalanine. Data represent average and standard deviation of three biological replicates each grown in three technical replicates. Tukey's HSD (Honestly Significant Difference) was used for statistical analysis. Variants marked with asterisk denotes significant changes: *P*-values: 0.01–0.05 = \*; *P*-values: 0.001–0.01 = \*\*\*; *P*-values: 0-0.001 = \*\*\*. (B) Glucosinolate analysis in *A. thaliana* ecotype Col-0 and knockout mutants *mam1* and *mam3*. Data represent average and standard deviation of 3–7 biological replicates.

analysis on leaves of *mam1* (TU1 [43] and *mam3* (SALK\_007222) mutants. 2-Phenylethyl glucosinolate was absent from the *mam1* and present in the *mam3* mutant, thus demonstrating that MAM1 is responsible for production of this glucosinolate in *A. thaliana* (Figure 7B and Supplementary Figure S8).

## Protein expression assessed by targeted proteomics

Protein levels of the chain elongation enzymes expressed in *E. coli* were monitored by targeted proteomics of proteotypic peptides representing the individual proteins. Seven proteins were monitored: MAM1, MAM3, IPMS2, BCAT4, LSU1, SSU3 and IPMDH1 (Figure 8 and Supplementary Figure S9). LSU1 and IPMDH1 were present at similar levels in all strains (Figure 8C,D), while both BCAT4 and SSU3 had altered expression in some strains. In strains harbouring MAM3 or the MAM1 variants T257G, GA and GAM, BCAT4 was lower than in the MAM1-containing strain (Figure 8A). In strains harbouring the MAM1 variants T257N, T257D, G259A and A290T, SSU3 protein levels were higher than in the strain harbouring MAM1 (Figure 8E). No chain-elongated amino acid products were found in strains expressing G259A and A290T and only low levels in T257N (Figure 5). However, the T257D had good production levels. Thus, higher SSU3 levels do not block chain elongation, and not all constructs resulting in low levels or no chain-elongated products have elevated SSU3 levels, as evident from the analysis of IPMS2 and T257G.

Regarding the protein levels of MAM1 variants, only one construct resulted in changed levels compared to the native MAM1 (Figure 8A). T257S showed higher MAM1 levels, although this did not result in correspondingly higher levels of chain-elongated products (Figure 5). As these results represent a relative quantification of protein levels, MAM3 (Figure 8D) and IPMS2 (Figure 8F) cannot be directly compared with the MAM1 variants. The approach, however, allows us to confirm expression of all biosynthetic enzymes.

## **Discussion**

We set out to obtain a better understanding of MAM1 substrate specificity and product profile, and thereby the number of recursive cycles of the chain elongation pathway in glucosinolate biosynthesis. Amino acid residues in MAM1 were selected for mutagenesis based on MAM-specific polymorphic residues identified in a phylogenetic analysis of 57 MAM protein sequences from 12 Brassicaceae species as well as the distance of these to the substrate. The latter was estimated from a MAM1 homology model and a TtHCS structure co-crystallised with a substrate [48]. The natural variation seen in MAM-specific polymorphic residues were mimicked in the MAM1 variants generated in the present study. Three of the residues mutated in MAM1 based on our selection criteria were predicted previously to be interacting with the substrate: two in the active site (Thr257 and Gly259) and one in the substrate pocket (Leu186) [12]. This validated our approach and selection criteria for identification of potentially important residues. We identified a fourth residue (Ala290) that was included in the analysis.



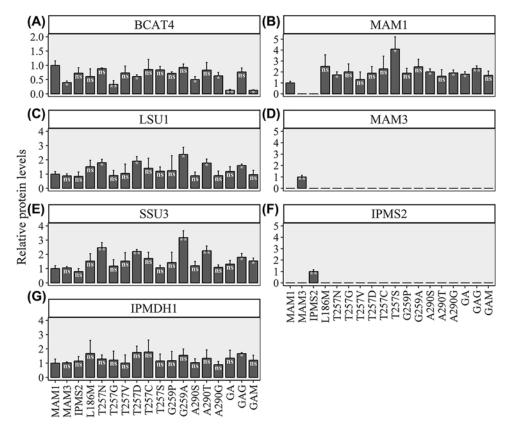


Figure 8. Protein levels in *E. coli* strains expressing the chain elongation pathway genes
(A) BCAT4, (B) MAM1, (C) LSU1, (D) MAM3, (E) SSU3, (F) IPMS2 and (G) IPMDH1. A representative peptide for each protein was used for relative quantification. Data represent average and standard deviation of three biological replicates.

We monitored the effect of protein engineering MAM1 by analysing the level and profile of chain-elongated amino acid products in medium of *E. coli* strains expressing the chain elongation pathway. As the Thr257 residue was highly polymorphic, it is represented by more substitution mutations than the other targeted residues in the library of MAM1 variants. The lowest producing strain harboured the T257N MAM1 variant, which suggests that introducing an amide group at the bottom of the active site may result in sterical hindrance, thereby lowering activity of the enzyme. The T257G variant produced severely reduced levels of elongated products compared with other MAM1 variants as well as MAM3. When the T257G substitution was performed in *Mt*IPMS it caused loss of activity in *E. coli* [12], which is in agreement with our results. MAM1 variants with substitutions T257V, T257D, 257C and T257S all produced a similar range of total products, but with different profiles of chain-elongated amino acids (Figure 5). T257S had an almost 3-fold increase in HPhe production compared with MAM1 (Figure 7A) and produced more phenylalanine-derived products than leucine- and methionine-derived products combined. The T257C substitution mimics MAMc, which is found in e.g. *C. sativa*, which produces glucosinolates derived from methionine derivatives from nine cycles of chain elongation [50]. Together with T257D, A290G and A290S, T257C had the highest TriHM production amongst the tested MAM1 variants including native MAM synthases. The production of TriHM as well as DHM in the T257C MAM1 variant confirms that homologues of MAMc have evolved to be iterative enzymes.

The IPMS2 mimicking G259P variant produced more chain-elongated leucine than methionine products. The increase in leucine-derived products was almost 2-fold. It was also the only MAM variant unable to elongate pheny-lalanine. The substitution has replaced glycine with the much bulkier proline. Such a substitution would normally be expected to reduce activity of an enzyme. However, proline can also introduce a bend in the  $\beta$ -sheet, which in this case could make room for the methylene branch on the leucine side chain [51].

The MAM1 variants G259A and A290T and IPMS2 did not generate chain-elongated amino acid products in *E. coli*, but the presence of the proteins was confirmed by targeted proteomics. The IPMS2 enzyme only elongated valine, and thus was not expected to produce other products. When the A290T mutation is introduced into MAM1 of *A. thaliana*, it generates a *mam1* knockout phenotype [9], which is in agreement with our data.



In *A. thaliana*, glucosinolates derived from chain-elongated leucine products are only seen in mutants that are altered in their branched chain amino acid metabolism [49,52]. However, when MAM1 variants are expressed together with the chain elongation enzymes in *E. coli*, chain-elongated leucine products were produced by all the *E. coli* strains. This could be explained by a loss of specificity when moving enzymes to heterologous hosts, but it could also suggests that leucine biosynthesis and methionine chain elongation are physical separated at the cellular level in the native host *A. thaliana*. Such model was previously proposed [53] and would allow for a localised high methionine-to-leucine ratio and thus hinder by-product formation. Indeed, feeding MAM1 with methionine reduced the proportion of leucine-derived by-products [20]. However, such physical separation into distinct localisations of the respective enzymes at the cellular level has yet to be proven.

We provide conclusive *in planta* evidence that MAM1 and not MAM3 as hitherto proposed is responsible for production of the HPhe-derived 2-phenylethyl glucosinolate (Figure 7A). The data are in agreement with the observation that the only MAM present in *Barbarea vulgaris* that produces high levels of 2-phenylethyl glucosinolate is a MAM1 homologue [54].

## **Acknowledgments**

AP would like to thank the Department of Plant and Environmental Sciences, University of Copenhagen for a PhD stipend. The authors like to thank Dan Kliebenstein and Meike Burow for providing seeds of the *A. thaliana mam1* and *mam3* knockout lines, respectively.

### **Author Contribution**

AP, LGH and NM contributed equally to this work that is based on an idea coined by LGH, NM and BAH. NM and OM modelled the MAM1 structures and identified residues targeted for mutation. LGH and NM designed the MAM1 variants and cloned the constructs. AP performed the experiments with assistance from CC on the LC-MS analysis. AP, CC and BAH analysed the data and wrote the paper based on a draft written by AP. All authors commented on the manuscript.

## **Funding**

This work was supported by the Danish National Research Foundation [grant number DNRF99]; and the Novo Nordisk Foundation [grant number NNF14OC0011253].

#### **Competing Interests**

The authors declare that there are no competing interests associated with the manuscript.

#### **Abbreviations**

BCAT4, branched-chain aminotransferase 4; DHL, dihomoleucine; DHM, dihomomethionine; HCS, homocitrate synthase; HeptaHM, heptahomomethionine; HexaHM, hexahomomethionine; HL, homoleucine; HM, homomethionine; HPhe, homophenylalanine; ICD, isocitrate dehydrogenase; IPMDH, isopropylmalate dehydrogenase; IPMI, isopropylmalate isomerase; IPMS, isopropylmalate synthase; LSU1, IPMI large subunit 1; MAM, methylthioalkylmalate synthase; PentaHM, pentahomomethionine; SSU3, IPMI small subunit 3; TetraHM, tetrahomomethionine; TFA, trifluoroacetic acid; TriHL, trihomoleucine; TriHM, trihomomethionine.

### References

- 1 Halkier, B.A. and Gershenzon, J. (2006) Biology and biochemistry of glucosinolates. Annu. Rev. Plant Biol. 57, 303–333, https://doi.org/10.1146/annurev.arplant.57.032905.105228
- 2 Jensen, L.M., Halkier, B.A. and Burow, M. (2014) How to discover a metabolic pathway? An update on gene identification in aliphatic glucosinolate biosynthesis, regulation and transport. Biol. Chem. 395, 529–543, https://doi.org/10.1515/hsz-2013-0286
- 3 de Kraker, J.-W., Luck, K., Textor, S., Tokuhisa, J.G. and Gershenzon, J. (2007) Two Arabidopsis genes (IPMS1 and IPMS2) encode isopropylmalate synthase, the branchpoint step in the biosynthesis of leucine. Plant Physiol. 143, 970–986, https://doi.org/10.1104/pp.106.085555
- 4 Benderoth, M., Pfalz, M. and Kroymann, J. (2009) Methylthioalkylmalate synthases: genetics, ecology and evolution. *Phytochem. Rev.* 8, 255–268, https://doi.org/10.1007/s11101-008-9097-1
- 5 Benderoth, M., Textor, S., Windsor, A.J., Mitchell-Olds, T., Gershenzon, J. and Kroymann, J. (2006) Positive selection driving diversification in plant secondary metabolism. *Proc. Natl. Acad. Sci. U.S.A.* 103, 9118–9123, https://doi.org/10.1073/pnas.0601738103
- 6 Frantom, P.A., Birman, Y., Hays, B.N. and Casey, A.K. (2014) An evolutionarily conserved alternate metal ligand is important for activity in α-isopropylmalate synthase from *Mycobacterium tuberculosis*. *Biochim. Biophys. Acta* **1844**, 1784–1789, https://doi.org/10.1016/j.bbapap.2014.07.013
- 7 Fahey, J.W., Zalcmann, A.T. and Talalay, P. (2001) The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. Phytochemistry **56**, 5–51, https://doi.org/10.1016/S0031-9422(00)00316-2



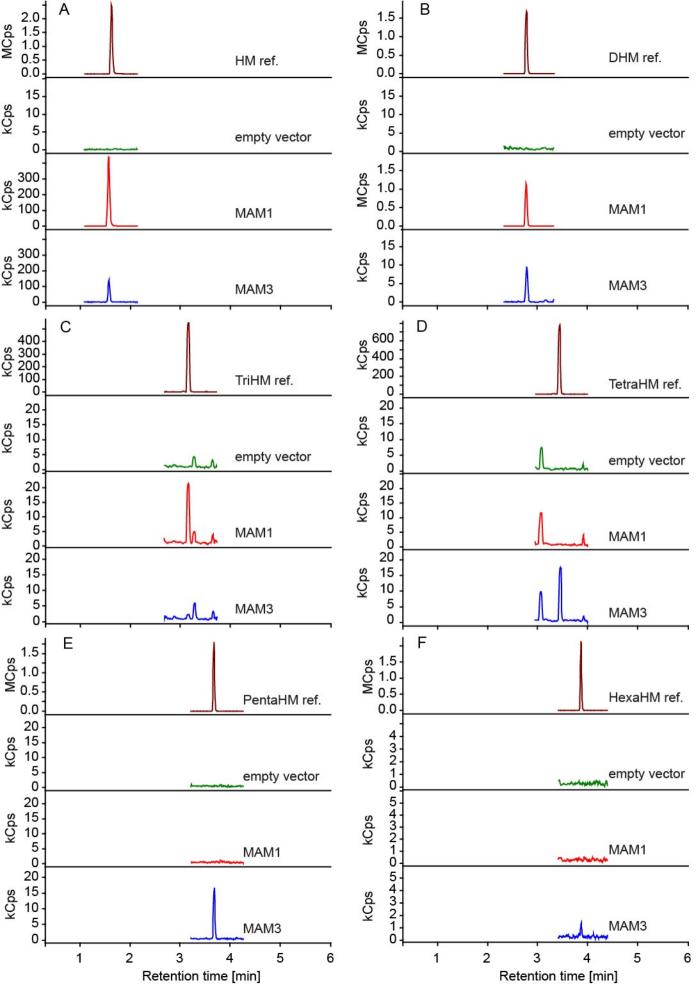
- 8 Burow, M., Halkier, B.A. and Kliebenstein, D.J. (2010) Regulatory networks of glucosinolates shape *Arabidopsis thaliana* fitness. *Curr. Opin. Plant Biol.* 13, 348–353, https://doi.org/10.1016/j.pbi.2010.02.002
- 9 Kroymann, J., Textor, S., Tokuhisa, J.G., Falk, K.L., Bartram, S., Gershenzon, J. et al. (2001) A gene controlling variation in Arabidopsis glucosinolate composition is part of the methionine chain elongation pathway. *Plant Physiol.* **127**, 1077–1088, https://doi.org/10.1104/pp.010416
- 10 Textor, S., de Kraker, J.-W., Hause, B., Gershenzon, J. and Tokuhisa, J.G. (2007) MAM3 catalyzes the formation of all aliphatic glucosinolate chain lengths in *Arabidopsis*. *Plant Physiol*. **144**, 60–71, https://doi.org/10.1104/pp.106.091579
- 11 Textor, S., Bartram, S., Kroymann, J., Falk, K.L., Hick, A., Pickett, J.A. et al. (2004) Biosynthesis of methionine-derived glucosinolates in *Arabidopsis thaliana*: recombinant expression and characterization of methylthioalkylmalate synthase, the condensing enzyme of the chain-elongation cycle. *Planta* 218, 1026–1035, https://doi.org/10.1007/s00425-003-1184-3
- 12 de Kraker, J.-W. and Gershenzon, J. (2011) From amino acid to glucosinolate biosynthesis: protein sequence changes in the evolution of methylthioalkylmalate synthase in *Arabidopsis*. *Plant Cell* **23**, 38–53, https://doi.org/10.1105/tpc.110.079269
- 13 Kumar, G., Johnson, J.L. and Frantom, P.A. (2016) Improving functional annotation in the DRE-TIM metallolyase superfamily through identification of active site fingerprints. *Biochemistry* **55**, 1863–1872, https://doi.org/10.1021/acs.biochem.5b01193
- 14 Koon, N., Squire, C.J. and Baker, E.N. (2004) Crystal structure of LeuA from *Mycobacterium tuberculosis*, a key enzyme in leucine biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 8295–8300, https://doi.org/10.1073/pnas.0400820101
- 15 Huisman, F.H.A., Koon, N., Bulloch, E.M.M., Baker, H.M., Baker, E.N., Squire, C.J. et al. (2012) Removal of the C-terminal regulatory domain of α-isopropylmalate synthase disrupts functional substrate binding. *Biochemistry* **51**, 2289–2297, https://doi.org/10.1021/bi201717j
- 16 Huisman, F.H.A., Squire, C.J. and Parker, E.J. (2013) Amino-acid substitutions at the domain interface affect substrate and allosteric inhibitor binding in α-isopropylmalate synthase from *Mycobacterium tuberculosis*. *Biochem. Biophys. Res. Commun.* 433, 249–254, https://doi.org/10.1016/j.bbrc.2013.02.092
- 17 Casey, A.K., Hicks, M.A., Johnson, J.L., Babbitt, P.C. and Frantom, P.A. (2014) Mechanistic and bioinformatic investigation of a conserved active site helix in α-isopropylmalate synthase from *Mycobacterium tuberculosis*, a member of the DRE-TIM metallolyase superfamily. *Biochemistry* **53**, 2915–2925, https://doi.org/10.1021/bi500246z
- 18 Kumar, G. and Frantom, P.A. (2014) Evolutionarily distinct versions of the multidomain enzyme α-isopropylmalate synthase share discrete mechanisms of V-type allosteric regulation. *Biochemistry* **53**, 4847–4856, https://doi.org/10.1021/bi500702u
- 19 Kliebenstein, D.J., Gershenzon, J. and Mitchell-Olds, T. (2001) Comparative quantitative trait loci mapping of aliphatic, indolic and benzylic glucosinolate production in *Arabidopsis thaliana* leaves and seeds. *Genetics* **159**, 359–370
- 20 Mirza, N., Crocoll, C., Erik Olsen, C. and Ann Halkier, B. (2016) Engineering of methionine chain elongation part of glucoraphanin pathway in *E. coli*. *Metab. Eng.* **35**, 31–37, https://doi.org/10.1016/j.ymben.2015.09.012
- 21 Crocoll, C., Mirza, N., Reichelt, M., Gershenzon, J. and Halkier, B.A. (2016) Optimization of engineered production of the glucoraphanin precursor dihomomethionine in *Nicotiana benthamiana*. *Front. Bioeng. Biotechnol.* **4**, 14, https://doi.org/10.3389/fbioe.2016.00014
- 22 Mikkelsen, M.D., Olsen, C.E. and Halkier, B.A. (2010) Production of the cancer-preventive glucoraphanin in tobacco. *Mol. Plant* 3, 751–759, https://doi.org/10.1093/mp/ssq020
- 23 Al-Shehbaz, I.A., Beilstein, M.A. and Kellogg, E.A. (2006) Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview. *Plant Syst. Evol.* **259**, 89–120, https://doi.org/10.1007/s00606-006-0415-z
- 24 Beilstein, M.A., Al-Shehbaz, I.A. and Kellogg, E.A. (2006) Brassicaceae phylogeny and trichome evolution. Am. J. Bot. 93, 607–619, https://doi.org/10.3732/ajb.93.4.607
- 25 Zhang, J., Wang, X., Cheng, F., Wu, J., Liang, J., Yang, W. et al. (2015) Lineage-specific evolution of methylthioalkylmalate synthases (MAMs) involved in glucosinolates biosynthesis. Front. Plant Sci. 6, 18
- 26 Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. J. Mol. Biol. 215, 403–410, https://doi.org/10.1016/S0022-2836(05)80360-2
- 27 Sali, A. and Blundell, TL. (1993) Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234, 779–815, https://doi.org/10.1006/jmbi.1993.1626
- 28 Nour-Eldin, H.H., Hansen, B.G., Nørholm, M.H.H., Jensen, J.K. and Halkier, B.A. (2006) Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. *Nucleic Acids Res.* **34.** e122, https://doi.org/10.1093/nar/gkl635
- 29 Geu-Flores, F., Nour-Eldin, H.H., Nielsen, M.T. and Halkier, B.A. (2007) USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Res.* **35**, e55, https://doi.org/10.1093/nar/gkm106
- 30 Jander, G., Norris, S.R., Joshi, V., Fraga, M., Rugg, A., Yu, S. et al. (2004) Application of a high-throughput HPLC-MS/MS assay to *Arabidopsis* mutant screening; evidence that threonine aldolase plays a role in seed nutritional quality. *Plant J.* **39**, 465–475, https://doi.org/10.1111/j.1365-313X.2004.02140.x
- 31 RStudio Team (2015) RStudio: Integrated Development for R, RStudio, Inc., Boston, MA
- 32 Wickham, H. (2016) ggplot2: Elegant Graphics for Data Analysis, Springer-Verlag, New York
- 33 Garnier, S., Ross, N., Rudis, B., Sciaini, M. and Scherer, C. (2018) viridis: Default Color Maps from Matplotlib, http://CRAN.R-project.org/package=viridis
- 34 Wickham, H. (2011) The split-apply-combine strategy for data analysis. J. Stat Softw. 40, 1–29, https://doi.org/10.18637/jss.v040.i01
- 35 Wickham, H. (2007) Reshaping data with thereshape package. J. Stat. Softw. 21, 1-20, https://doi.org/10.18637/jss.v021.i12
- 36 MacLean, B., Tomazela, D.M., Shulman, N., Chambers, M., Finney, G.L., Frewen, B. et al. (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, 966–968, https://doi.org/10.1093/bioinformatics/btq054
- 37 Wessel, D. and Flügge, UI. (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal. Biochem.* **138**, 141–143, https://doi.org/10.1016/0003-2697(84)90782-6

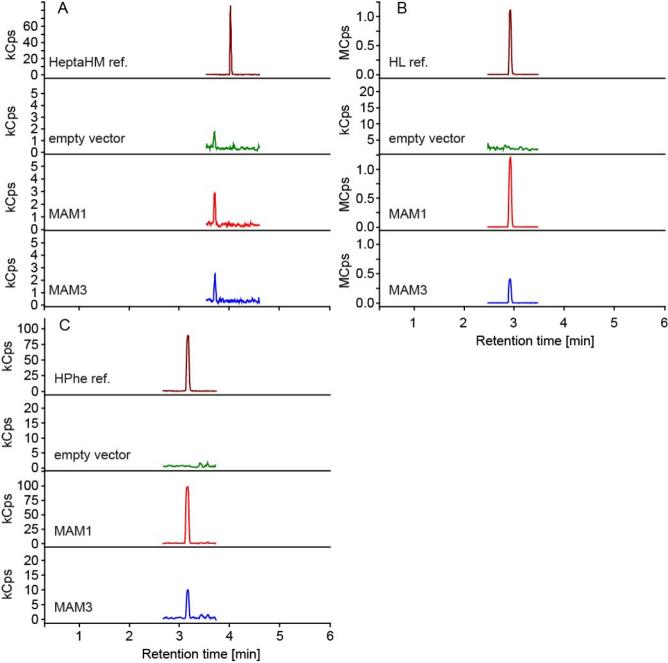


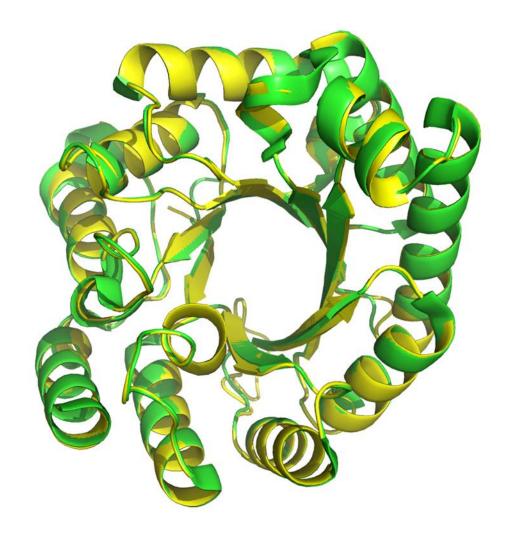
- 38 Batth, T.S., Keasling, J.D. and Petzold, C.J. (2012) Targeted proteomics for metabolic pathway optimization. Methods Mol. Biol. 944, 237-249
- 39 Petersen, A., Crocoll, C. and Halkier, B.A. (2018) De novo production of benzyl glucosinolate in *Escherichia coli*. *Metab. Eng.* **54**, 24–34, https://doi.org/10.1016/j.ymben.2019.02.004
- 40 Percy, A.J., Chambers, A.G., Yang, J., Domanski, D. and Borchers, C.H. (2012) Comparison of standard- and nano-flow liquid chromatography platforms for MRM-based quantitation of putative plasma biomarker proteins. *Anal. Bioanal. Chem.* 404, 1089–1101, https://doi.org/10.1007/s00216-012-6010-y
- 41 Batth, T.S., Singh, P., Ramakrishnan, V.R., Sousa, M.M.L., Chan, L.J.G., Tran, H.M. et al. (2014) A targeted proteomics toolkit for high-throughput absolute quantification of *Escherichia coli* proteins. *Metab. Eng.* 26, 48–56, https://doi.org/10.1016/j.ymben.2014.08.004
- 42 de Mendiburu, F. (2017) agricolae: Statistical Procedures for Agricultural Research, http://CRAN.R-project.org/package=agricolae
- 43 Haughn, G.W., Davin, L., Giblin, M. and Underhill, E.W. (1991) Biochemical genetics of plant secondary metabolites in *Arabidopsis thaliana*: the glucosinolates. *Plant Physiol.* **97**, 217–226, https://doi.org/10.1104/pp.97.1.217
- 44 Hogge, L.R., Reed, D.W., Underhill, E.W. and Haughn, G.W. (1988) HPLC separation of glucosinolates from leaves and seeds of *Arabidopsis thaliana* and their identification using thermospray liquid chromatography/mass spectrometry. *J. Chromatogr. Sci.* 26, 551–556, https://doi.org/10.1093/chromsci/26.11.551
- 45 Kliebenstein, D.J., Lambrix, V.M., Reichelt, M., Gershenzon, J. and 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, 681–693, https://doi.org/10.1105/tpc.13.3.681
- 46 Crocoll, C., Halkier, B.A. and Burow, M. (2016) Analysis and quantification of glucosinolates. In *Current Protocols in Plant Biology* (Stacey, G., Birchler, J., Ecker, J., Martin, C.R., Stitt, M. and Zhou, J.-M., eds), pp. 385–409, John Wiley & Sons, Inc., Hoboken, NJ, U.S.A.
- 47 Jensen, L.M., Jepsen, H.S.K., Halkier, B.A., Kliebenstein, D.J. and Burow, M. (2015) Natural variation in cross-talk between glucosinolates and onset of flowering in *Arabidopsis*. Front. Plant Sci. **6**, 697, https://doi.org/10.3389/fpls.2015.00697
- 48 Okada, T., Tomita, T., Wulandari, A.P., Kuzuyama, T. and Nishiyama, M. (2010) Mechanism of substrate recognition and insight into feedback inhibition of homocitrate synthase from *Thermus thermophilus*. *J. Biol. Chem.* **285**, 4195–4205, https://doi.org/10.1074/jbc.M109.086330
- 49 Field, B., Cardon, G., Traka, M., Botterman, J., Vancanneyt, G. and Mithen, R. (2004) Glucosinolate and amino acid biosynthesis in *Arabidopsis. Plant Physiol.* **135**, 828–839, https://doi.org/10.1104/pp.104.039347
- 50 Russo, R., Galasso, I. and Reggiani, R. (2014) Variability in glucosinolate content among Camelina species. Am. J. Polym. Sci. 05, 294-298
- 51 Reardon, D. and Farber, G.K. (1995) The structure and evolution of alpha/beta barrel proteins. FASEB J. 9, 497–503, https://doi.org/10.1096/fasebj.9.7.7737457
- 52 Binder, S., Knill, T. and Schuster, J. (2007) Branched-chain amino acid metabolism in higher plants. *Physiol. Plant.* **129**, 68–78, https://doi.org/10.1111/j.1399-3054.2006.00800.x
- 53 Imhof, J., Huber, F., Reichelt, M., Gershenzon, J., Wiegreffe, C., Lächler, K. et al. (2014) The small subunit 1 of the *Arabidopsis* isopropylmalate isomerase is required for normal growth and development and the early stages of glucosinolate formation. *PLoS One* **9**, e91071, https://doi.org/10.1371/journal.pone.0091071
- 54 Liu, T., Zhang, X., Yang, H., Agerbirk, N., Qiu, Y., Wang, H. et al. (2016) Aromatic glucosinolate biosynthesis pathway in *Barbarea vulgaris* and its response to *Plutella xylostella* infestation. *Front. Plant Sci.* **7**, 83

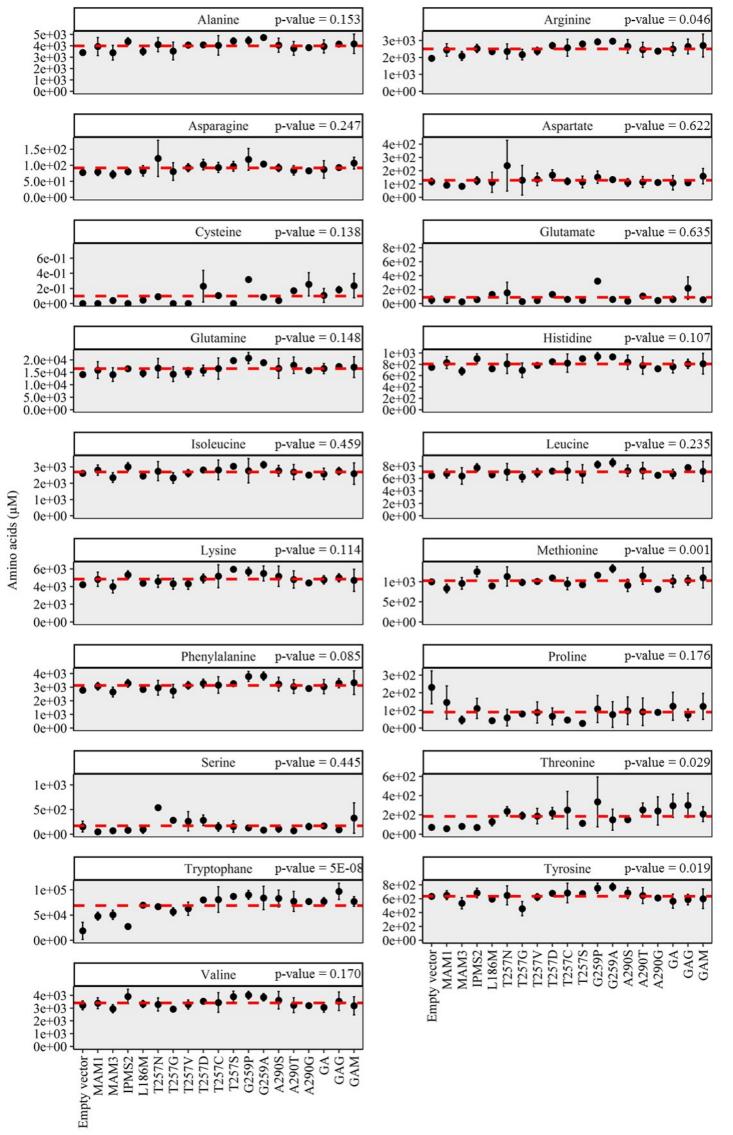
Arabidopsis thaliana Col-0 (NP_001330047.1) AtMAM1	LTIGA 5	
Arabidopsis thaliana mt-0 (AJ486884.1) AtmMAM1	LTIGA 5	
Arabidopsis thaliana sorbo (CAD31141.1) AtsMAM1	LTIGA 5	
Arabidopsis thaliana Br-0 (1001 genomes 6904) AtbMAM1	LSIGA 5	
Brassica oleracea var. oleracea (XP 013618041.1) BooMAM1.1	VACPS 5	
Brassica napus (XP_013678219.1) BnMAM1.1	VACPS 5	
Brassica rapa (XP_009130133.1) BrMAM1.1	VACPS 5	
Eruca vesicaria subsp. sativa (AGS49163.1) EvsMAM	LACPS 5	
Arabidopsis thaliana Col-0 (NP_173285.2) AtIPMS1	HNIPS 5	
Arabidopsis thaliana Col-0 (NP 177544.1) AtIPMS2	HNIPS 5	
Brassica napus (XM 013817468.2) BnMAM1.2	INIGS 5	
Brassica rapa (XM 009138151.2) BrMAM1.2	INIGS 5	
Brassica rapa subsp. pekinensis (ACR10246.1) BrpMAM.1	INIGS 5	
Raphanus sativus (XM_018606131.1) RsMAM1.2	LNLGS 5	
Arabidopsis thaliana Ler-0 (Q8VX04.1) AtMAM2	LNVAS 5	
Arabidopsis cebennensis (CAJ55509.1) AcMAMa	LNVAS 5	
Arabidopsis lyrata lyrata (XP_002872029.1) AllMAM2	LNVAS 5	
Arabidopsis lyrata petraea (CAJ55506.1) AlpMAMa	LNVAS 5	
Eutrema salsugineum (XP_006394566.2) EsMAM2	LNVAS 5	
Camelina sativa (XP_010493409.1) CsMAM2.1	LNVAS 5	
Camelina sativa (XP_010421098.1) CsMAM2.2	LNVAS 5	
Camelina sativa (XP_010454580.1) CsMAM1	LNVAS 5	
Boechera divaricarpa (CAJ55512.1) BdMAMa		
1		
Capsella rubella (XP_006289820.1) CrMAM2	LNVAS 5	
Brassica oleracea var. oleracea (XP_013593685.1) BooMAM2	LNVAS 5	
Brassica napus (XP_013730921.2) BnMAM2	LNVAS 5	
Raphanus sativus (XP_018477323.1) RsMAM2	LNVAS 5	
Eutrema salsugineum (XP_006394567.1) EsMAM1	LNIAS 5	
Brassica rapa (XM 009138149.2) BrMAM1.4	LNLAS 5	
Brassica rapa subsp. pekinensis (ACR10248.1) BrpMAM.2	LNLAS 5	
Brassica oleracea var. oleracea (XM 013763569.1) BooMAM1.2	FTLAS 5	
Brassica rapa (XM 009146283.2) BrMAM1.3	FTLAS 5	
Raphanus sativus (XP 018457526.1) RsMAM1.1	FTLAS 5	
Arabidopsis lyrata lyrata (CAJ55517.1) AllMAMc.1	CCIAS 5	
Arabidopsis lyrata subsp. lyrata (XP_020879218.1) AllMAM3.2	CCIAS 5	
Arabidopsis lyrata subsp. lyrata (EFH48290.1) AllMAMc.2	CCIAS 5	
Arabidopsis lyrata petraea (CAJ55508.1) AlpMAMc	CCIAS 5	
Camelina sativa (XP_019083909.1) CsMAM3.3	CCIAS 5	
Capsella rubella (XP_006289994.1) CrMAM1	CCIAS 5	
Arabidopsis cebennensis (CAJ55511.1) AcMAMc	CCIAS 5	
Boechera divaricarpa (CAJ55514.1) BdMAMc	CCIAS 5	
Camelina sativa (XP 010493411.1) CsMAM3.1	CCIAS 5	
Camelina sativa (XP_010454582.1) CsMAM3.7	CCIAS 5	
Capsella rubella (XM_006289411.2) CrMAM3.2	CCVAG 5	
Camelina sativa (XM 010473736.2) CsMAM3.4	GGFAG 5	
· · · · · · · · · · · · · · · · · · ·		
Camelina sativa (XM_010418504.2) CsMAM3.5	GGFAG 5	
Capsella rubella (XM_006301203.2) CrMAM3.3	IGFAG 5	
Camelina sativa (XM_010430647.2) CsMAM3.6	IGFAG 5	
Boechera divaricarpa (CAJ55513.1) BdMAMb	IGFAA 5	
Arabidopsis lyrata lyrata (XP_002872030.1) AllMAM3.1	LGFAA 5	
Arabidopsis lyrata subsp. petraea (CAJ55507.1) AlpMAMb	LGFAT 5	
Arabidopsis cebennensis (CAJ55510.1) AcMAMb	IGFAT 5	
Arabidopsis thaliana Col-0 (NP 197693.1) AtMAM3	MGFAA 5	
Arabidopsis thaliana Br-0 (1001 genomes 6904) AtbMAM3	MGFAA 5	
Arabidopsis thaliana Ler-0 (1001 genomes 7213) AtlMAM3	MGFAA 5	
Arabidopsis thaliana sorbo (AJ486943.1) AtsMAM3	MGFAA 5	
Camelina sativa (XP_010421099.1) CsMAM3.2	TGFAA 5	
Camelina sativa (XP_010493410.1) CsMAM3.8	TGFAA 5	
Capsella rubella (XP_006290200.1) CrMAM3.1	TGFAA 5	

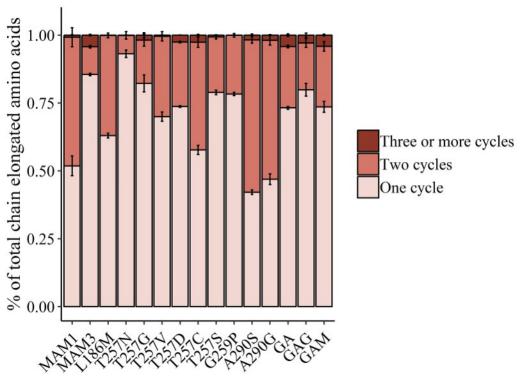
- 1		1 2	3	4	5	6 7	8	9	10	11	12 13	14	15	16	17 18	19	20 21	22	23	24	25 2	26 27	7 28	29	30 31	32	33 34	35	36 37	38	39 40	41 42	43	44 45	46	47 48	8 49	50	51 52	53 F	4 55	56 57
AtMAM1	1	100	,00 99,	,41 95,45	92,11	92,11 91,3	2 85,85	86,05	82,09	87,82	85,32 75,5	55 79,4	49 79,53	79,96	79,61 79,84	79,65	73,59 79	,06 75	,73 77,60	77,80	0 77,21 7	77,80 79	9,02 77,76	75,49	76,67 73,9	2 75,98	69,85 77	,56 77,56	76,97 77,1	76,97	79,33 75,59	77,17 75	83 72,05	71,65 70,	67 71,71	72,62 72	2,82 65	,94 60,89	60,95 60,	75 76,32 7	5,93 76,22	45,31 44,
AtmMAM1	2	0	99,	,41 95,45	92,11	92,11 91,3	2 85,85	86,05	82,09	87,82	85,32 75,5	55 79,4	49 79,53	79,96	79,61 79,84	79,65	73,59 79	,06 75	,73 77,60	77,80	77,21 7	77,80 79	9,02 77,76	75,49	76,67 73,9	2 75,98	69,85 77	10 (2) WO (10 (2)	100000000000000000000000000000000000000	76,97	79,33 75,59	77,17 75	83 72,05	71,65 70,	67 71,71		100000	,94 60,89	60,95 60,	75 76,32 7	5,93 76,22	45,31 44,
AtsMAM1	3	3	3	95,26	91,91	91,91 91,1	2 86,05	86,05	81,89	87,82	85,32 75,5	55 79,4	49 79,53	80,15	79,61 80,04	79,84	73,77 79	,26 75	,73 78,00	78,19	9 77,60 7	78,19 79	9,02 78,15	75,69	76,86 74,1	2 76,17	70,22 77	76 77,76	77,17 77,3	36 77,17	79,53 75,79	77,36 75	83 72,24	71,85 70,	87 72,10	72,62 72	2,82 65	,94 60,89	60,95 60	75 76,52 7	6,13 76,41	45,31 44,
AtMAM2	4	23	23	24	93,69	93,69 92,9	0 87,03	87,23	82,87	88,41	86,69 76,6	65 80,6	66 80,12	81,09	79,41 79,26	79,06	73,22 78	.47 74	.95 78,59	78,78	8 78,19 7	78,78 80	0,00 78,74	75,10	76,47 73,9	2 75,59	69,85 75	79 75,98	75,20 74,6	74,41	77,56 74,41	75,79 74	66 71,06	70,67 70,	47 71,51	71,46 71	1,65 65	.35 61,28	61,54 61,	75,34	4,95 75,24	44,69 43,
AIIMAM2	5	40	40	41 32		99.21 96.8	4 88,80	88,61	84,84	90,37	88.45 77.5	57 81.4	45 81,73	82,06	80,63 78,91	78,71	72.73 77	.73 74	.41 78.04	78.24	4 77.65 7	78.24 80	0.43 77.80	74.76	76,52 73,7	8 75,44	69.54 76	62 76,82	76,03 75,6	75,44	78,19 73,48	75.05 74	31 69,94	69.16 68.	96 70.59	71,51 71	1,71 65	.03 60.97	61,22 61	75,39	5.00 75.68	45.00 44.
AlpMAMa	6	40	40	41 32	4	96.8	4 89.00	88.80	85.04	90.37	88.65 77.7	76 81.6	64 81.93	82.24	80.23 79.10	78,91	72,91 77	93 74	.61 78.04	78.24	4 77.65 7	78.24 80	0.43 77.80	74.56	76.32 73.5	8 75.24	69.54 76	77. 1	76,03 75,6	75.44	78.39 73.48	75.05 74	31 69.94	69 16 68	96 70.78	71.51 71	1.71 65	03 60.97	61.42 61	22 75.00	4.61 75.49	45.16 44
AcMAMa	7	44	44	45 36	16	16	88.02	87.82	83,86	89 39	87.87 76.8	84 80 6	66 80.75	81 31	79,84 78,12		71,82 76		22 77.45	77.65	5 77.06 7	77.65 80	0.23 77.41	74 56	75 93 73 1	9 75.44	69,54 75			74 66	77.60 73.08	1.5,55	12 69.55	68.76 68	57 70 59	71.32 71	151 64	83 61 17	60.83 60.	63 74.22	3.83 74.90	44.69 44.
CsMAM2.1	8	72	72	71 66	57	56 6	1	97.42	92.06	93.45	92 13 77	39 814	45 80 94	80.93	78.54 78.52		72 18 77	15 73	63 76.67	76.86	6 76.27 7	76.86 76	6 91 77 67	73 39	75.34 71.8	2 74.07	69 17 75		75.64 75.8	76.03	76.62 72.69	73.92 72	94 69 17	68 38 68	18 70.61	71.07 71	1.26 64	82 59.88	60.79 60	79 73.87	3.28 74.07	45.62 44
CsMAM2.2	0	71	71	71 65	58	57 6	2 13	51,12	92.86	93.85	92 91 77 3	39 814	45 80.94	80.75	78 74 79 10	78.91	73.09 77	73 73	63 77.45	77.65	5 77.06 7	77.65 77	7 69 78 46	73 30	75.54 72.2	74.66	69.54 74	.85 75,05	75,64 76,0	76,03	77.41 72.89	74 12 73	53 69 17	68 38 68	18 70,41	71.65 71	1.84 65	42 60.67	61 39 61	39 74.07	3.48 74.46	45.62 44
CsMAM1	10	01	01	02 97	77	76 9	2 40	26	32,00	89.09	87.99 73.5	53 773	34 77.01	77,20	78,31 75,64	0.717.0	69,29 74	07 75	,56 73.14	73,33	2 22422 2	73 33 7	2,99 73,91	70.45	72.02 69.9	9 72.02	66,79 72	trace constitute	300 March 300 M	19 73 29	74.26 69.74	70.79 70	00 66 60	65.81 66	00 67.86	70.75 70	0.75 65	11 50.06	60.20 60	120 69.74	0 16 70 57	44.20 43
10000000000	10	80	60	62 50	40	40 5	4 22	24	EE	0.7550	92 32 78	12 82 2	34 77,01	81.87	80.31 80.27	1010000000	100000000000000000000000000000000000000	91 75	20 78.82	1911	100000000000000000000000000000000000000	HARRING TOO	100 000 100 100 100 100 100 100 100 100	74 17	76,32 73,1	0 75.24	20000000	,83 76,03	100.000.000	75,20	78,59 73,87	70,78 70	12 60.06	60.17 60	77 71 20	70,75 70	2,62 66	21 61.84	61.00 61	70 75 05	4.06 74.66	44,20 45,
BdMAMa	10	75	75	75 60	F0	FO 6	2 40	26	61	20	76,5	12 02,2	23 02,31	70.70	70.47 70.40	78.21	72.10 76	100	25 76 27	79,02	75.00		6.61 77.95	70.54	74.46 72.2	73,24	60.20 75	Marie	2000000	76,23	77.10 72.02	100520	66 60.00	69,17 60,	11 71,20	70.66 70	0.05 64	44 50.00	60.43 60	75,05	4,20 74,00	45,78 45,
CrMAM2	12	122	122 4	122 427	100	101 10	6 123	100	144	110		93.1	15 07.00	07.00	75.46 75.41	10,01	74.05 75	,00 73	04 70.70	70,50	70,90	72.25	0,01 77,95	60.70	70,93 67.7	70.00	60.22 72	10 70,34	73,73 73,3	75,34	72,11 70,52	71,99 72	14 64.04	64.24 62	04 65 40	60.40 60	0,63 64	00 57.05	58.04 57	7.06 60.50	1,02 12,02	43,40 44,
BooMAM2	13	105	105 1	105 00	122	04 0	0 123	123	144	04	128	93,	01,22	07,02	79,18 79,34	10,20	72.44 70	,41 /1	,01 12,19	76.70	6 76 47 7	77.45 76	2,00 72,50	72.00	70,83 07,7	70,28	69,33 72	93 76 13	73,11 72,	72,01	75,11 70,53	71,32 71	00 60.04	67.45 67	00,19	68,49 68	8,67 62	92 51,95	56,04 57,	71.00	72.06	45,00 42,
BnMAM2	14	105	105 1	105 99	95	94 9	9 95	95	110	91	100	20	91,73	84,89	79,18 79,34	13,14	73,14 79	,34 74	,00 70,00	70,70	70,17	77.47 70	0,41 76,32	73,29	74,46 71,1	0 73,39	68,37 75	,93 70,13	77,10 75,8	75,73	76,91 74,17	75,00 74	00,24	67,45 67,	00 08,09	74.65 74	2,39 66	74 60.67	01,49 01,	71,98	73,06	45,10 44,
RsMAM2	15	104	104 1	104 101	93	92 9	8 9/	400	117	07	103	09 4	42	84,40	3.0100		77.00 70	,90 /5	70,77	70,91	7 70,38 7	70.07	0,02 70,08	70.50	72.64 70.6	70.00	70.00 75	,13 /0,33	75.00 74.0	75,74	75,00 70,05	70,59 75	,00 08,32	07,52 00,	93 08,77	(1,00) 7	1,84 05	,74 00,07	50.40 51,	71,96 7	1,57 73,83	45,31 45,
EsMAM2	10	107	107 1	100 101	96	404 40	102	103	122	400	140	34 00	01 83	405	80,41 79,44		70.00 70	44 75	70,59	70,78	75.24	75.00	5.00 75.00	74.00	73,04 70,6	72,90	10,02 /5	,23 75,42		72 72.50	75,96 72,05	72,00 72	10 07,00	00,85 66,	60 08,22	75.04	5.04 62	90 00 00	59,10 59,	100 71,40 /	72,86	44,29 43,
EsMAM1	17	104	104 1	104 105	99	101 10	3 109	108	108	100	110 13	34 10	07 106	105	//,19	77,00	70,96 76	,41 /5	,94 /5,/3	75,93	7 75.00	0,93 /	5,20 75,00	71,29	72,85 70,1	2 73,53	00,91 /4	74,90	74,51 73,7	3 /3,53	77,09 72,55	73,19 /4	05 70.15	09,09 68,	00 00 75	70,04 /5	0.04 68	04 57.00	50,19 64,	,99 /1,62 /	73,83	44,53 44,
BOOMAM1.1	18	103	103 1	102 106	108	107 11	2 110	107	124	101	111 13	34 10	05 102	110	117	99,80	90,61 92	,28 87	,52 /6,27	76,47	7 75,88 7	0,67 75	5,69 74,71	71,76	73,53 70,3	72,96	67,58 76	08 76,27	76,67 76,0	76,27	77,06 73,92	74,95 74	,95 /0,45	69,86 69,	08 68,75	70,02 70	0,21 63	57,06	58,43 58	43 72,21 7	1,62 /2,90	45,00 45,
BnMAM1.1	19	104	104 1	103 107	109	108 11	3 111	108	125	102	112 13	35 10	103	111	118 1		90,42 92	,08 87	,33 /6,08	76,27	75,69 7	75	0,49 /4,51	/1,5/	73,33 70,3	9 72,76	67,40 76	,08 76,27	76,47 75,8	76,08	70,86 73,73	74,76 74	/6 /0,45	69,86 69,	08 68,75	70,02 70	0,21 63	57,06	58,43 58	43 72,02 7	1,43 /2,71	45,00 45,
BrMAM1.1	20	145	145 1	144 147	150	149 15	5 153	148	168	144	154 13	38 14	48 146	121	160 51	52	86	,3/ 80	,85 70,07	70,26	69,71 7	0,44 69	9,53 68,80	66,42	67,88 64,9	66,91	66,36 69	70,07	70,44 69,7	69,89	70,44 67,70	68,67 68	65,03	64,48 63,	63,45	64,98 65	5,16 58	,11 52,89	53,83 54	66,30 6	5,76 67,70	42,48 42,
BjIPMS1	21	107	107 1	106 110	114	113 11	9 117	114	132	108	119 13	34 10	06 102	111	121 39	40	74	86	,34 76,27	76,47	7 75,88 7	76,67 75	5,69 75,49	71,96	73,53 70,2	72,76	67,58 75	,49 75,69	76,27 75,2	29 75,49	76,27 74,12	75,34 75	,15 70,84	70,25 69,	28 68,75	69,83 70	0,02 62	,43 56,87	57,84 57	,65 71,62 7	1,23 73,49	45,16 44,
EvsMAM	22	124	124 1	124 128	131	130 13	2 135	135	120	127	137 15	58 13	30 127	137	121 63	64	104	69	73,37	73,57	7 72,98 7	73,77 72	2,19 71,99	70,22	72,19 68,2	72,89	66,67 74	,07 74,26	74,07 73,4	18 73,48	74,26 71,79	73,03 73	,23 68,64	67,85 68,	05 67,32	69,88 70	0,08 63.	,84 58,22	57,34 57	,14 67,72 6	7,52 69,22	43,44 43,
AllMAMc.1	23	114	114 1	112 109	112	112 11	5 119	115	137	108	121 14	48 12	20 118	125	124 121	122	164 1	121 1	135	99,40	98,61 9	99,20 93	3,85 88,47	83,73	85,71 82,3	4 83,04	75,14 80	,00 80,20	80,00 79,8	79,80	81,58 78,53	79,37 77	,98 70,24	69,84 69,	44 70,89	69,57 69	9,77 64	,55 59,49	59,33 59	,13 72,92 7	2,33 74,41	44,06 42,
AIIMAM3.2	24	113	113 1	111 108	111	111 11	4 118	114	136	107	120 14	47 11	19 117	124	123 120	121	163 1	120 1	134 3		99,20 9	99,40 94	4,05 89,07	84,13	86,31 82,7	4 83,43	75,51 80	,20 80,40	80,20 80,0	00,08	81,78 79,13	79,96 78	,17 70,83	70,04 70,	04 71,49	69,77 69	9,96 64	,95 59,49	59,33 59	,13 73,12 7	2,53 74,61	44,38 42,
AIIMAMc.2	25	116	116 1	114 111	114	114 11	7 121	117	139	110	123 15	50 12	22 120	127	126 123	124	166 1	123 1	137 7	4	4 9	98,81 93	3,45 88,47	83,73	85,71 83,2	3 82,64	74,77 79	,60 79,80	79,60 79,4	79,41	81,19 78,53	79,37 77	,58 70,24	69,44 69,	44 70,89	69,19 69	9,38 65	,34 59,84	59,68 59	,48 72,53 7	1,94 74,02	44,06 42,
AlpMAMc	26	113	113 1	111 108	111	111 11	4 118	114	136	107	120 14	45 11	17 116	123	123 119	120	162 1	119 1	133 4	- 3	3 6	94	4,25 88,87	84,13	86,31 82,5	4 83,23	75,32 80	,20 80,40	80,20 80,0	00,08	81,78 79,13	79,96 78	,57 70,63	69,84 69,	84 71,09	69,96 70	0,16 64	,75 59,49	59,13 58	,93 73,52 7	2,92 75,00	44,53 42,
AcMAMc	27	107	107 1	107 102	100	100 10	1 118	114	138	108	120 14	49 12	21 119	126	127 124	125	167 1	124 1	141 31	30	0 33	29	86,31	81,94	83,93 80,3	6 80,31	72,78 78	,26 78,46	78,06 77,8	77,67	79,05 76,39	77,03 76	,04 69,11	68,32 68,	12 70,36	69,05 69	9,25 64	,03 58,98	59,21 59	,01 72,73 7	2,13 73,62	43,28 42,
BdMAMc	28	113	113 1	111 108	113	113 11	5 113	109	132	106	112 14	49 12	21 118	129	127 129	130	171 1	125 1	142 58	55	5 58	56	69	84,49	86,68 82,3	84,66	76,78 77	,14 77,34	77,53 77,3	77,34	80,52 78,31	79,16 77	,76 75,73	74,90 74,	27 75,78	68,48 68	8,68 67	,56 61,76	62,45 62	,24 74,50 7	3,90 73,12	43,59 42,
CsMAM3.1	29	125	125 1	124 127	129	130 13	0 136	136	151	132	141 16	65 13	37 134	147	147 144	145	184 1	143 1	151 82	80	0 82	80	91 78		94,43 86,4	8 82,64	76,07 75	,49 75,69	75,10 74,5	74,51	77,08 77,58	78,61 75	,45 69,25	68,65 68,	25 67,92	65,38 65	5,57 60	,79 55,58	55,56 55	,36 69,96 6	9,37 70,28	43,75 43,
CsMAM3.7	30	119	119 1	118 120	120	121 12	3 126	125	143	121	131 15	59 13	31 128	141	139 135	136	176 1	135 1	141 72	69	9 72	69	81 67	28	89,4	6 84,42	77,37 76	,88 77,08	76,48 76,4	18 76,48	78,66 78,97	80,00 76	,83 70,44	69,64 69,	25 69,90	67,31 67	7,50 62	,38 56,95	56,75 56	,55 71,54 7	J,75 71,65	43,91 43,
CsMAM3.3	31	133	133 1	132 133	134	135 13	7 144	142	159	137	142 17	76 14	48 141	157	153 151	151	192 1	152 1	161 89	87	7 84	88	99 89	68	53	79,09	72,36 73	,72 73,91	73,52 73,7	72 73,72	75,30 73,61	74,26 72	,28 66,27	65,48 65,	48 66,53	66,02 66	6,21 65	,62 59,88	60,33 60	,13 67,19 6	3,40 67,72	42,16 41,
CrMAM1	32	123	123 1	122 125	126	127 12	6 133	130	143	127	134 16	52 13	36 133	145	135 139	140	182 1	140 1	138 86	84	4 88	85	100 77	88	79 10	6	84,83 74	,95 75,15	75,54 75,1	75,15	78,11 78,13	78,77 79	,92 69,98	69,18 68,	39 69,84	68,29 68	8,48 63	,55 59,06	57,68 57	,49 69,35 6	3,76 70,33	43,15 41,
CrMAM3.2	33	164	164 1	162 164	166	166 16	6 168	166	181	163	168 16	59 17	73 169	164	180 177	178	185 1	177 1	181 134	132	2 136	133	147 124	129	122 14	9 81	69	,94 70,13	70,87 70,1	70,50	72,73 74,21	75,00 74	,95 66,36	65,61 65,	23 67,16	61,79 61	1,97 57	,65 53,60	52,90 52	71 64,33 6	3,59 65,06	40,95 39,
AIIMAM3.1	34	114	114 1	113 123	119	119 12	3 126	128	142	123	127 15	51 12	23 121	132	129 122	122	165 1	125 1	132 101	100	0 103	100	110 115	124	117 13	3 127	162	99,60	96,22 92,0	91,65	88,47 83,50	84,52 82	,34 73,61	72,82 72,	22 72,08	70,23 70	0,43 64	,36 58,32	57,94 57	,74 72,78 7	2,19 73,87	43,66 44,
AlpMAMb	35	114	114 1	113 122	118	118 12	2 125	127	141	122	126 15	50 12	22 120	131	128 121	121	164 1	124 1	131 100	99	9 102	99	109 114	123	116 13	2 126	161	2	96,61 92,0	91,65	88,47 83,50	84,52 82	,34 73,81	73,02 72,	42 72,28	70,43 70	0,62 64	,55 58,51	58,13 57	,94 72,98 7	2,39 74,07	43,82 44,
AcMAMb	36	117	117 1	116 126	122	122 12	124	124	139	119	124 14	46 11	17 115	132	130 119	120	162 1	121 1	132 101	100	0 103	100	111 113	126	119 13	4 124	157	19 17	93,6	93,24	88,47 83,10	83,93 82	,94 73,41	72,62 71,	83 71,88	70,62 70	0,82 65	,15 58,51	58,53 58	,33 73,18 7	2,39 74,07	43,97 44,
AtMAM3	37	116	116 1	115 129	124	124 12	8 123	122	137	120	126 15	51 12	23 122	135	134 122	123	166 1	126 1	135 102	10	1 104	101	112 114	129	119 13	3 126	161	40 40	32	99,20	88,67 83,50	84,13 82	,34 72,82	72,02 71,	43 70,89	70,04 70	0,23 64	,75 58,71	58,33 57	,94 73,18 7	2,58 74,66	43,44 43,
AtsMAM3	38	117	117 1	116 130	125	125 12	9 122	121	136	121	126 15	52 12	24 123	136	135 121	122	165 1	125 1	135 102	101	1 104	101	113 114	129	119 13	3 126	159	42 42	34	4	88,47 83,30	83,93 82	,34 72,82	72,02 71,	43 70,89	70,04 70	0,23 64	,75 58,51	58,13 57	,74 73,18 7	2,58 74,66	43,44 43,
BdMAMb	39	105	105 1	104 114	111	110 11	4 119	115	131	109	117 14	46 11	18 119	128	124 117	118	162 1	121 1	131 93	92	2 95	92	106 98	116	108 12	5 111	147	58 58	58 5	57 58	87,08	88,29 84	73,41	72,82 72,	62 71,68	70,82 71	1,01 65	,35 59,69	58,73 58	,53 72,98 7	2,39 74,26	43,12 43,
CsMAM3.2	40	124	124 1	123 130	135	135 13	7 139	138	154	133	143 10	50 13	32 130	149	140 133	134	177 1	132 1	143 108	105	5 108	105	119 108	113	106 13	3 110	138	83 83	85 8	84	65	96,79 89	,18 73,75	73,15 72,	95 71,20	66,73 66	6,93 63	,00 58,30	56,31 56	,11 71,43 7	0,83 72,92	43,28 43,
CsMAM3.8	41	116	116 1	115 123	127	127 12	9 133	132	149	128	138 15	56 12	28 124	146	137 128	129	172 1	126 1	137 104	101	1 104	101	116 104	108	101 13	0 107	134	78 78	81 8	80 81	59 16	90	,60 74,00	73,40 73,	00 71,06	67,96 68	8,16 63	,47 58,97	57,40 57	,20 72,48 7	1,88 74,16	43,84 43,
CrMAM3.1	42	123	123 1	123 129	131	131 13	2 138	135	153	132	140 15	57 12	29 127	147	132 128	129	172 1	127 1	136 111	110	0 113	108	121 111	124	117 14	0 101	134	89 89	86 8	89 89	77 54	47	74,00	73,20 72,	69,86	67,70 67	7,90 63	,07 58,58	57,20 57	,00 70,10 6	9,70 72,33	43,75 43,
CsMAM3.4	43	142	142 1	141 147	153	153 15	5 156	156	168	152	157 19	90 16	62 160	173	153 151	151	192 1	149 1	159 150	147	7 150	148	156 117	155	149 17	0 151	180	132	134 13	137	134 131	130 1	30	99,15 95,	73 81,65	64,01 64	4,20 63	,81 57,97	58,28 58	,07 69,08 6	8,27 67,59	41,85 41,
CsMAM3.5	44	144	144 1	143 149	157	157 15	9 160	160	172	156	161 19	94 16	66 164	177	157 154	154	195 1	152 1	163 152	151	1 154	152	160 121	158	153 17	4 155	184	136	138 14	141	137 134	133 1	34 4	94,	80,80	63,42 63	3,62 63	,18 57,56	58,49 58	,28 68,67 6	7,87 67,19	41,54 41,
CsMAM3.6	45	149	149 1	148 150	158	158 16	0 161	161	171	158	162 19	96 16	68 167	177	160 158	158	199 1	157 1	162 154	151	1 154	152	161 124	160	155 17	4 159	186 1	40 139	142 14	144	138 135	135 1	36 20	24	82,28	63,23 63	3,42 62	,97 57,14	56,39 56	,18 67,87 £	7,47 66,40	41,22 41,
CrMAM3.3	46	144	144 1	142 145	150	149 15	0 149	150	162	146	150 18	39 16	60 158	170	159 160	160	201 1	160 1	166 147	144	4 147	146	150 117	162	152 16	9 152	176 1	41 140	142 14	17 147	143 144	145 1	51 87	91	84	65,56 65	5,76 65	,69 59,63	58,79 58	,58 69,48 F	9,08 68,97	42,41 41,
BnMAM1.2	47	141	141 1	141 147	147	147 14	8 149	146	148	142	152 17	73 14	44 146	163	123 155	155	194 1	156 1	153 157	156	159	155	160 162	179	169 17	4 163	209 1	53 152	151 15	154	150 171	165 1	66 185	188 1	89 177	99	9,80 84	,10 70,92	64,66 64	, <mark>46 64,85 6</mark>	4,27 66,80	43,82 43,
BrMAM1.2	48	140	140 1	140 146	146	146 14	7 148	145	148	141	151 17	72 14	43 145	162	122 154	154	193 1	155 1	152 156	155	5 158	154	159 161	178	168 17	3 162	208 1	52 151	150 15	153	149 170	164 1	65 184	187 1	88 176	1	83	,94 70,78	64,53 64	,33 65,05 6	4,47 66,99	43,97 43,
BrpMAM.1	49	173	173 1	173 176	178	178 17	9 178	175	172	171	181 20	01 17	70 173	198	155 189	189	230 1	192 1	179 179	177	7 174	178	182 157	198	190 16	5 183	227	80 179	176 17	78 178	175 185	183	85 173	176 1	77 164	79	80	80,79	74,65 74	,41 61,17 F	0,76 61,58	39,03 38,
RsMAM1.2	50	201	201 2	201 199	201	201 20	0 205	201	201	195	206 23	30 20	00 201	226	185 222	222	261 2	223 2	211 207	207	7 204	207	210 187	227	220 19	5 208	251 2	213 212	212 21	212	206 211	208 2	210 203	205 2	07 195	146	147	83	75,24 75	,00 56,46 5	6,06 57,34	36,66 35,
BrMAM1.4	51	198	198 1	198 195	197	196 19	9 198	195	197	192	201 22	27 19	96 192	218	173 212	212	253 2	215 2	212 205	205	5 202	206	206 181	224	218 19	0 212	252 2	212 211	209 21	10 211	208 218	213 2	199	198 2	08 197	176	177 1	108 103	99	,50 57,23 5	6,83 57,51	35,16 35,
BrpMAM.2	52	199	199 1	199 196	198	197 20	0 198	195	197	193	202 22	28 19	97 193	218	174 212	212	252 2	216 2	213 206	206	6 203	207	207 182	225	219 19	1 213	253 2	213 212	210 21	213	209 219	214 2	215 200	199 2	09 198	177	178 1	109 104	2	57,03	6,63 57,31	35,01 34,
BooMAM1.2	53	121	121 1	120 126	126	128 13	2 133	132	154	127	142 17	72 14	44 143	153	145 142	143	185 1	145 1	164 137	136	6 139	134	138 127	152	144 16	6 156	193 1	137	136 13	136	137 144	139 1	51 154	156 1	60 152	181	180 1	193 219	213	214	8,19 88,91	44,84 44,
BrMAM1.3	54	123	123 1	122 128	128	130 13	4 136	135	157	131	145 17	74 14	46 145	156	148 145	146	188 1	147 1	165 140	139	9 142	137	141 130	155	148 17	0 159	197 1	41 140	140 13	139	140 147	142	53 158	160 1	62 154	184	183 1	195 221	215	216 9	88,12	45,47 43,
RsMAM1.1	55	122	122 1	121 127	125	126 12	9 133	131	151	130	140 16	67 13	39 134	146	134 139	140	178 1	136 1	157 130	129	9 132	127	134 136	151	144 16	4 151	189	33 132	132 12	29 129	131 137	131 1	40 164	166 1	70 157	171	170 1	194 218	215 7	216 56	60	45,94 45,
AtIPMS1	56	350	350 3	350 354	352	351 35	4 348	348	356	347	350 37	79 35	51 350	371	355 352	352	390 3	351 3	362 358	356	6 358	355	363 361	360	359 36	9 365	398 3	359	358 36	362	364 363	360 3	371	373 3	75 368	359	358 3	889 406	413	114 353	349 346	86,
AtIPMS2	57	355	355 3	355 360	356	355 35	8 356	356	359	352	357 38	35 35	57 352	375	354 353	353	393 3	355 3	361 370	368	370	367	371 368	365	364 37	5 375	408 3	358 357	356 36	362	365 364	362 3	62 373	375 3	76 373	362	361 3	394 412	414	415 358	360 350	87
STATE OF THE STATE OF	10000	28	. 28	- ATTO-655	2000 500	Section 1	200000	Section .	STOCKERS	5000 Billion	SIRVE SIRVE	Maria Street	- 100 mm	NE DEWONE	ARTON ARTON	NACOCK .	M000000 30	Control (	SHOWS	SHOPW	31100000	allowards S	HARRIOT STOTAL	MINORAL	and the same of th	THE STATE OF THE S	A 900 H	Market Street,	ARCON MAN	end 2 Milotope 2	STRONGERS STRONGERS	allegade g Alle	omes a Michigan	1 SECONDS 1970	STREET, STREET,	Alfonotes S	10000000	0.000	The second secon	2000	COLUMN TO THE PARTY OF THE PART	

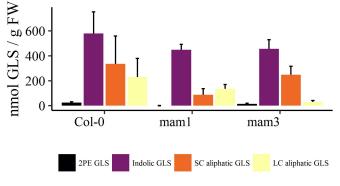


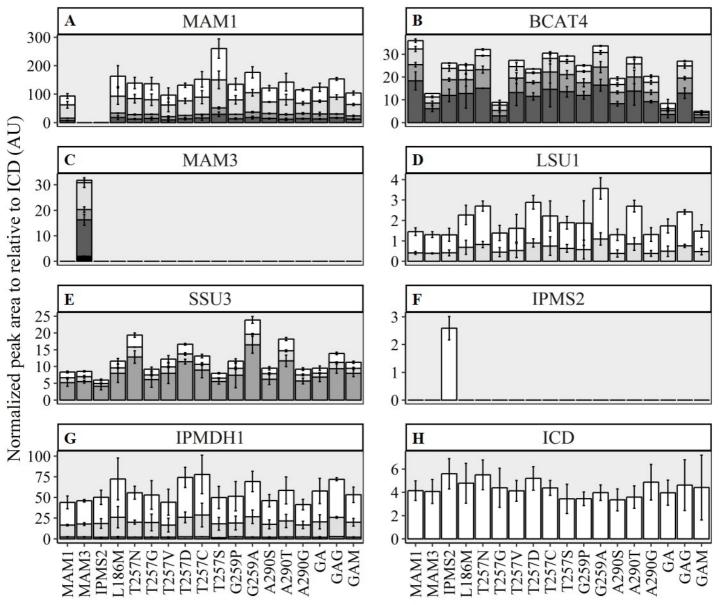












# **Supplementary Figure legends**

**Supplementary Figure S1**. Alignment of 57 MAM and 2 IPMS full protein sequences. According to *A. thaliana* MAM1 numbering, only Leu186, Thr257, Ile258, Gly259 and Ala290 residues are shown.

**Supplementary Figure S2**. Matrix from the alignment of MAM and IPMS enzyme sequences. Upper part indicates sequence similarity in percentage and the lower part shows differences between the protein sequences. Two MAM sequences have been removed from the matrix, as the comparison indicated low sequence similarity due to poor quality of the sequence data.

**Supplementary Figure S3**. LC-MS traces of homomethionine (HM), dihomomethionine (DHM), trihomomethionine (TriHM), tetrahomomethionine (TetraHM), pentahomomethionine (PentaHM) and hexahomomethionine (HexaHM). A representative sample of MAM1 and MAM3 containing strains as well as an empty vector control are shown. For each compound, the pure sample is shown in the top panel.

**Supplementary Figure S4**. LC-MS traces of heptahomomethionine (HeptaHM), homoleucine (HL) and homophenylalanine (HPhe). A representative sample of MAM1 and MAM3 containing strains as well as an empty vector control are shown. For each compound, the pure sample is shown in the top panel.

**Supplementary Figure S5.** Homology model of MAM1 from *A. thaliana*. (yellow) superimposed on the crystal structure of *N. meningitis* IPMS (green). Only the  $(\beta/\alpha)_8$  catalytic barrel is shown. The RMSD value is 0.26 Å.

**Supplementary Figure S6.** Protein amino acids monitored in this study. Each graph shows levels of one amino acid in the media of *E. coli* strains expressing MAM variants and an empty vector control. The mean of all cultures is indicated by the red line. ANOVA was used to investigate changes in the individual amino acids between the cultures. p-values are shown in upper right corner of each graph. Data represents average and standard deviation of three biological replicates each grown in three technical replicates.

**Supplementary Figure S7**. Produced chain elongated amino acids were grouped according to how many methylene groups were added to the methionine, phenylalanine or leucine structures (one, two or three and above). Bars represents the percentage of each group of products relative to the total production within each culture. Data represents average and standard deviation of three biological replicates each grown in three technical replicates.

**Supplementary Figure S8.** Glucosinolate analysis in *A. thaliana* ecotype Col-0 and knockout mutants mam1 and mam3. Glucosinolates (GLS) were grouped into aromatic (2PE) GLSs, indole GLSs and short chain (SC – one to three elongation rounds) and long chain (LC – four elongation rounds and above) aliphatic GLSs. Data represents average and standard deviation of 3-7 biological replicates.

**Supplementary Figure S9**. Protein levels in *E. coli* strains expressing the chain elongation pathway genes, A) *BCAT4*, B) *MAM1*, C) *LSU1*, D) *MAM3*, E) *SSU3*, F) *IPMS2* and G) *IPMDH1* and H) *E. coli* housekeeping gene *ICD*. A-G are normalized the ICD, while the ICD is not normalized. Different colours indicate individual peptides within each protein. Data represents average and standard deviation of three biological replicates.

# **Supplementary Tables**

**Supplementary Table S1.** List of primers used in this study. Lower case letters denotes the USER tail used for cloning. Uracils are marked in blue. Mutated nucleotides are marked in red.

Name	Sequence	Description
MAM1_fwd	ggcttaauATGCATCACCACCACCATCACAACTATGTGCGTGTATTC	MAM1 forward primer
BCAT4_rev	ggtttaauTCAGCCCTGGCGGTCAAT	BCAT4 reverse primer
MAM3_fwd	${\tt ggcttaau} ATGCATCACCACCACCATCACGCTGAGTCCAAAAAGGTGGC$	MAM3 forward primer
MAM3_rev	acaaaattauTTCTAGAGGGGTTATACAACAGCGGAAATC	MAM3 reverse primer
IPMS2_fwd	ggcttaauATGCATCACCACCACCATCACATGGAGTCTTCGATTCTCA	IPMS2 forward primer
IPMS2_rev	$a caa aa atta {\color{blue} u} TTCTAGAGGGGTCAGGCAGGGACTTCGTTG$	IPMS2 reverse primer
L186_fwd	acatctactaguGACATTCACATGAAATATAAG	Fusion primer for L186-mutant
L186M_rev	actagtagatguGAATACCATTATCCTTGGCCTC	Fusion primer for L186M-mutant
T257/G259_fwd	acggtagggauCAACATGCCACATGAATACGG	Fusion primer for T257-mutant
T257N_rev	atccctaccguATCACCGATGTTCACCACCGTAA	Fusion primer for T257N-mutant
T257G_rev	atccctaccguATCACCGATGCCCACCACCGTAA	Fusion primer for T257G-mutant
T257V_rev	atccctaccguATCACCGATGACCACCACCGTAA	Fusion primer for T257V-mutant
T257D_rev	atccctaccguATCACCGATGTCCACCACCGTAA	Fusion primer for T257D-mutant
T257C_rev	atccctaccguATCACCGATGCACACCACCGTAA	Fusion primer for T257C-mutant
T257S_rev	atccctaccguATCACCGATGCTCACCACCGTAA	Fusion primer for T257S-mutant
G259P_rev	atccctaccguATCTGGGATGGTCACCACCGTAA	Fusion primer for G259P-mutant
G259A_rev	atccctaccguATCGGCGATGGTCACCACCGTAA	Fusion primer for G259A-mutant
A290S_fwd	acgatgttgtcguCAGCGTTCATTGTCACAACGAC	Fusion primer for A290S-mutant
A290T_fwd	acgatgttgtcguCACCGTTCATTGTCACAACGAC	Fusion primer for A290T-mutant
A290G_fwd	acgatgttgtcguCGGCGTTCATTGTCACAACGAC	Fusion primer for A290G-mutant
A290_rev	acgacaacatcguCAATTCCAGGGGTGTTTGCTTTG	Fusion primer for A290-mutant
GA_rev	atccctaccguATCGGCGATGCCCACCACCGTAA	Fusion primer for T257G/G259A-mutant

**Supplementary Table S2.** Transitions and response factors used to quantify levels of chain elongated amino acids. Response factors of methionine- and leucine-derived products were reported in (20).

Amino acid	Q1	Q3	CE	Internal standard	Q1	Q3	Response factor
	180.1	91.1 <sup>Qt</sup>	-22	13C,15N-Phe	176.2	129.2	1.61
HPhe	180.1	134.1	-8				
	180.1	117.1	-13				
HM	164.2	70.3 <sup>Qt</sup>	-16	13C,15N-Val	124.1	77.2	0.35
DHM	178.2	84.2 <sup>Qt</sup>	-14	13C,15N-Phe	176.2	129.2	1.05
DHM	178.2	132.1	-10				
TriHM	192.2	98.2 <sup>Qt</sup>	-14	13C,15N-Phe	176.2	129.2	1.25
TTITIVI	192.2	146.1	-9				
TetraHM	206.2	112.2 <sup>Qt</sup>	-13	13C,15N-Phe	176.2	129.2	0.91
Tett attivi	206.2	160.1	-10				
PentaHM	220.3	126.2 <sup>Qt</sup>	-16	13C,15N-Phe	176.2	129.2	0.41
1 chtainvi	220.3	174.1	-10				
НехаНМ	234.4	140.2 <sup>Qt</sup>	-16	13C,15N-Phe	176.2	129.2	0.36
HEXAIINI	234.4	188.1	-11				
HeptaHM	248.2	188.1 <sup>Qt</sup>	-10	13C,15N-Phe	176.2	129.2	n.d.
періаніч	148.2	154.1	-16				
HL	146.2	100.1	10	13C,15N-Phe	176.2	129.2	0.88
DHL	160.2	114.1	-10	13C,15N-Phe	176.2	129.2	1.00*
TriHL	174.2	128.1	-10	13C,15N-Phe	176.2	129.2	1.00*

<sup>\*</sup>Response factors for DHL and TriHL were assumed to be 1.00. CE = collision energy. Q = quadrupole. Qt = quantifier transition

**Supplemental Table S3:** Transitions used for Selected Reaction Monitoring (SRM) for targeted proteomics on peptides of methionine chain elongation pathway proteins. Retention times and parent ion to fragment ion transitions are given. All peptides were selected in Q1 as the double-charged ions [M+H]<sup>2+</sup> and fragment ions in Q3 as single-charged ions [M+H]<sup>+</sup>. Collision energies (CE) for natural non-labelled (light) and isotopically labelled (heavy) peptides were identical.

				Light p	eptide	Heavy p	eptide	_
				Q1	Q3	Q1	Q3	CE
Protein name	AGI code	Peptide	RT	$[M+H]^{2+}$	$[M+H]^{+}$	$[M+H]^{2+}$	$[M+H]^{+}$	[eV]
BCAT4	At3g19710	TGEETLAAK	1.60	460.25	761.42	464.25	769.42	-13
			1.60	460.25	818.44	464.25	826.44	-15
			1.60	460.25	632.38	464.25	640.38	-14
		LYETLSDIQTGR*	8.30	698.36	776.4	703.36	786.4	-19
			8.30	698.36	990.53	703.36	1000.53	-26
			8.30	698.36	1119.57		1129.57	-23
		SITNYCPVWIPLAEAK	18.70	931.49	927.54	935.49	935.54	-3
			18.70	931.49	1123.67	935.49	1131.67	-29
			18.70	931.49	628.38	935.49	636.38	-41
		GNVVSTPTIAGTILPGVTR	16.50	927.03	529.32	932.03	539.32	
			16.50	927.03	642.4	932.03	652.4	
MAM1	At5g23010	SLGFNDIQFGCEDGGR*	13.60	886.39	897.36	891.39	907.36	
			13.60		1025.42		1035.42	
			13.60	886.39	404.2	891.39	414.2	
		DGEQSPGGSLTPPQK	2.50	749.37	981.55	753.37	989.55	
			2.50		1068.58		1076.58	
			2.50	749.37	469.29	753.37	477.29	
		STYEILSPEDIGIVK	17.00	832.45	957.54	836.45	965.54	
			17.00	832.45	870.51	836.45	878.51	-25
			17.00		1070.62		1078.62	
		SGNASLEEVVMALK	17.50	724.38	918.51		1126.63	
			17.50		1118.63		1039.59	
			17.50		1031.59	728.38	926.51	-23
MAM3	At5g23020	GESLMDGVYTK	6.70	600.29	813.4	604.29	821.4	
			6.70		1013.51		1021.51	-16
			6.70	600.29	682.35	604.29	690.35	
		ALVVNGAEISSEK	5.70	658.86	934.46	662.86	942.46	
			5.70		1033.53		1041.53	
			5.70	658.86	820.42	662.86	828.42	
		SGNAPLEEVVMALK	18.70		1128.65		1136.65	
			18.70	729.39	918.51	733.39	926.51	-25
		OWNER OPERATOR HAT	18.70	729.39	588.4	733.39	596.4	
		STYEILSPEDVGIVK*	15.50	825.44	943.52	829.44	951.52	
			15.50	825.44	856.49	829.44	864.49	
		DOEOGDO A ALEBROY	15.50		1056.61		1064.61	
		DGEQSPGAALTPPQK	3.40	748.38	469.29	752.38	477.29	
			3.40	748.38	372.24	752.38	380.24	-41

			2.40	740.20	070.57	752.20	007.57	2.4
TD) 100			3.40	748.38	979.57	752.38	987.57	-24
IPMS2	At1g/4040	GTYEIMSPEEIGLER*	15.40		1102.53		1112.53	-28
			15.40	889.94	942.5	894.94	952.5	-32
IPMI-LSU1	At4g13430	FILDGEMPSYLQAK*	16.90	806.42	806.45	810.42	814.45	-26
			16.90	806.42	937.5	810.42	945.5	-24
			16.90		1123.56		1131.56	-24
		VWMDVYALPVPGAGGK	18.50	830.44	682.4	834.44	690.4	-21
			18.50	830.44	486.28	834.44	494.28	-37
			18.50	830.44	261.17	834.44	269.17	-49
IPMI-SSU3	At3g58990	LGSFALNGLPK	12.60	558.83	712.45	562.83	720.45	-19
			12.60		1003.57		1011.57	-18
			12.60	558.83	641.41	562.83	649.41	-17
		EDGSSLLINHTTR	3.90	721.87	854.49	726.87	864.49	-29
			3.90	721.87	967.58	726.87	977.58	-29
			3.90	721.87	741.41	726.87	751.41	-28
		NCVATGEIFPLESEVR*	15.10	910.95	829.45	915.95	839.45	-30
			15.10	910.95	976.52	915.95	986.52	-30
			15.10	910.95	490.27	915.95	500.27	-37
IPMDH1	At5g14200	LSDAILLGAIGGYK*	17.00	695.91	891.54	699.91	899.54	-19
			17.00	695.91	778.46	699.91	786.46	-20
			17.00	695.91	1004.63	699.91	1012.63	-22
		AGSLEGLEFDFK	14.20	656.83	855.44	660.83	863.44	-20
			14.20	656.83	984.48	660.83	992.48	-18
			14.20	656.83	685.33	660.83	693.33	-17
		IEDAVVDALNK	6.05	593.83	944.52	597.83	952.52	-17
			6.05	593.83	659.39	597.83	667.39	-17
			6.05	593.83	758.45	597.83	766.45	-17
ICD	JW1122	GPLTTPVGGGIR*	6.37	562.83	655.40	567.83	665.40	-20
			6.37	562.83	857.49	567.83	867.49	-21
			6.37	562.83	756.44	567.83	766.44	-19

<sup>\*=</sup> peptides represented in Figure 8.

**Supplemental Table S4:** Pairwise t-test of the methionine levels from media of *E. coli* strains expression the chain elongation pathway with 18 different MAM variants and an empty vector control. p-values shown.

	Empty vector	MAM1	MAM3	IPMS2	L186M	T257N	T257G	T257V	T257D	T257C	T257S	G259P	G259A	A290S	A290T	A290G	GA	GAG
MAM1	0.1316	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MAM3	0.7101	0.2507	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IPMS2	0.0254	0.0004	0.5510	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
L186M	0.3535	0.0103	0.5755	0.0023	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T257N	0.2301	0.0088	0.1192	0.2755	0.0372	-	-	-	-	-	-	-	-	-	-	-	-	-
T257G	0.8806	0.1727	0.8245	0.0178	0.4355	0.1784	-	-	-	-	-	-	-	-	-	-	-	-
T257V	0.9379	0.1136	0.6532	0.0305	0.3152	0.2609	0.8196	-	-	-	-	-	-	-	-	-	-	-
T257D	0.3792	0.0199	0.2139	0.1591	0.0752	0.7432	0.3045	0.4223	-	-	-	-	-	-	-	-	-	-
T257C	0.6634	0.2773	0.9491	0.0087	0.6195	0.1055	0.7753	0.6080	0.1920	-	-	-	-	-	-	-	-	-
T257S	0.5230	0.3757	0.7885	0.0051	0.7698	0.0700	0.6245	0.4740	0.1332	0.8379	-	-	-	-	-	-	-	-
G259P	0.1362	0.0040	0.0655	0.4266	0.0185	0.7637	0.1024	0.1569	0.5307	0.0572	0.0366	-	-	-	-	-	-	-
G259A	0.0043	4.9e-05	0.0016	0.4823	0.0003	0.0773	0.0029	0.0053	0.0383	0.0013	0.0007	0.1385	-	-	-	-	-	-
A290S	0.4230	0.4693	0.6657	0.0033	0.8977	0.0494	0.5140	0.3799	0.0974	0.7125	0.8695	0.0251	0.0005	-	-	-	-	-
A290T	0.1735	0.0057	0.0862	0.3536	0.0254	0.8680	0.1323	0.1985	0.6218	0.0758	0.0492	0.8930	0.1075	0.0342	-	-	-	-
A290G	0.0920	0.8519	0.1836	0.0002	0.4347	0.0054	0.1229	0.0786	0.0126	0.2048	0.2851	0.0024	2.8e-05	0.3639	0.0035	-	-	-
GA	0.8826	0.0993	0.6039	0.0357	0.2834	0.2909	0.7659	0.9444	0.4632	0.5604	0.4326	0.1776	0.0064	0.3438	0.2232	0.0681	-	-
GAG	0.7299	0.0666	0.4745	0.0552	0.2058	0.3888	0.6206	0.7891	0.5911	0.4364	0.3272	0.2475	0.0106	0.2541	0.3053	0.0446	0.8432	-
GAM	0.3673	0.0189	0.2059	0.1656	0.0718	0.7602	0.2943	0.4096	0.9821	0.1847	0.1278	0.5454	0.0403	0.0931	0.6377	0.0119	0.4498	0.5758