

Harnessing mass spectrometry-based proteomics for continuous directed evolution

Katharina Belt¹, David Obe², Mark A. Wilson², A. Harvey Millar¹, Ulschan Bathe^{1,3,*}

¹ARC Centre of Excellence in Plant Energy Biology, School of Molecular Sciences, University of Western Australia, Crawley WA 6009, Australia

²Department of Biochemistry and Redox Biology Center, University of Nebraska, Lincoln, NE 68588, USA

³Biotechnology of Horticultural Crops, TUM School of Life Sciences, Technical University of Munich, Liesel-Beckmann-Str. 1, Freising 85354, Bavaria, Germany

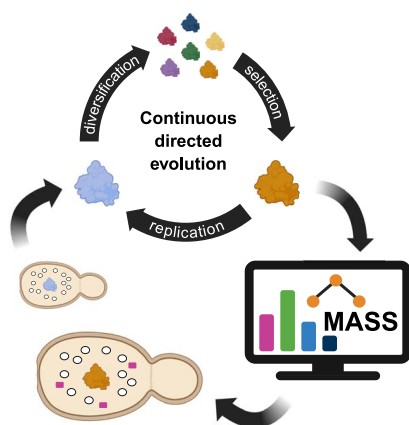
*Corresponding author. Biotechnology of Horticultural Crops, TUM School of Life Sciences, Technical University of Munich, Liesel-Beckmann-Str. 1, Freising 85354, Germany. E-mail: ulschan.bathe@tum.de

We dedicate this work to the memory of Andrew D. Hanson, whose pioneering contributions and scientific vision continue to inspire the field of synthetic biology.

Abstract

Continuous directed evolution is a powerful Synthetic Biology tool to engineer proteins with desired functions *in vivo*. Mimicking natural evolution, it involves repeated cycles of high-frequency mutagenesis, selection, and replication within platform cells, where the function of the target gene is tightly linked to the host cell's fitness. However, cells might escape the selection pressure due to the inherent flexibility of their metabolism, which allows for adaptation. Whole-proteome analysis as well as targeted proteomics offer valuable insights into global and specific cellular changes. They can identify modifications in the target protein and its interactors to help understand its evolution and network integration. Using the continuous evolution of the *Arabidopsis thaliana* methionine synthases AtMS1 and AtMS2 as an example, we show how mass spectrometry-based proteomics was able to assess the abundance of target enzymes, identify flaws in population construction, measure methionine metabolic adaptation, and allow informed decision-making in the evolution campaign.

Graphical Abstract



Keywords: spectrometry-based proteomics; continuous directed evolution; OrthoRep; *Arabidopsis* methionine synthase; enzyme longevity

Introduction

Directed evolution is one of Synthetic Biology's most outstanding achievements. It aims at engineering biomolecules, e.g. enzymes, with required phenotypes or functions [1]. Classical directed evolution was awarded the 2018 Nobel Prize [2], but continuous directed evolution (CDE) is considered even more powerful overcoming limitations in scale of experimentation and depth of

evolutionary search. Thus, it holds the potential to shape future medicine, agriculture and industry.

The principle of CDE is *in vivo* sequence diversification, selection and replication, and couples the function of the target gene to the host cell's fitness [3]. Diverse CDE tools allow directed evolution in different platform cell types, e.g. phage-based, bacterial, yeast or mammalian systems [4–10] to engineer antibodies,

Received: 23 July 2025; revised 9 October 2025; accepted 30 October 2025

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

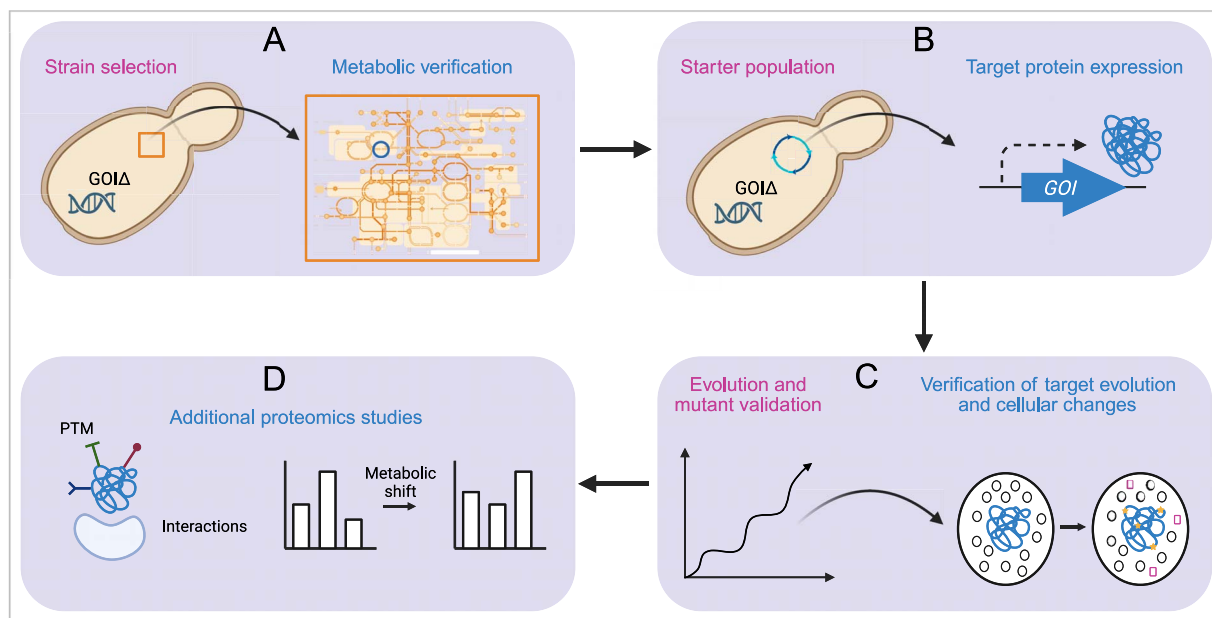


Figure 1. Proteomics in continuous directed evolution. Proteomics can be incorporated at critical stages of CDE to evaluate protein-level changes (A-D). A. The CDE workflow begins with baseline proteomic analysis of the initial strain to assess global protein expression. B. Targeted proteomics can validate expression and abundance of the protein of interest. C. Untargeted proteomic analysis can identify early protein shifts in CDE, e.g. protein-level adaptations. The best-performing strains are finally selected, and proteomic validation confirms the protein function associated with the observed traits (D). The figure was created in Biorender, Bathe, U. (2025) <https://BioRender.com/r1fhguw>.

substrate specificity, improved catalytic efficiency, and novel enzyme activities [11–14].

CDE is considered powerful because repeated cycles of mutagenesis and screening inside the platform organism allow efficient and scalable evolution. However, the host cells' metabolism is a complex and interconnected protein machinery, and CDE may cause trade-off effects in the proteome because the selection pressure is not selective enough or potentially affects other pathways. In the worst case, the platform cells escape the selection pressure entirely. In such cases, post-CDE evaluation would reveal that evolved phenotypes are not only or not exclusively connected to the target gene's function. To reduce the risk of an unsuccessful campaign and to make CDE even more powerful, whole-system analysis might be used to look beyond single target genes, transcripts, and metabolites. While whole genome and transcriptome sequencing would give some upstream insights into a cell's real condition, metabolomics and proteomics deliver more information about the platform's cellular changes because they are closer to the phenotype. Proteomics is not utilized as a standard tool in SynBio experiments thus far [15]; for example, we are aware of only one case of a CDE study utilizing whole-cell proteomics [16]. Yet, mass spectrometry-based proteomics can provide qualitative and quantitative insights into an organism's entire proteome or of specific protein targets and their relative or absolute abundance [17–19] (Fig. 1). For example, proteomics data can help to validate the platform strain (Fig. 1A) and the target protein expression (Fig. 1B and C), directly show interaction of the engineered target with other biomolecules and pathways, and potential trade-offs (Fig. 1C and D). These attributes make proteomics an invaluable tool for CDE, enabling the verification of intermediate steps, validation of final functions, and troubleshooting of challenges during the evolution process.

As a use case, we applied CDE to engineer enzyme longevity. Directed evolution has rarely been used to attempt extension of working lifetime of enzymes *in vivo*, i.e. to increase their catalytic-cycles-till-replacement (CCR) value [20, 21]. This unitless number describes how many reactions an enzyme can catalyze in

its *in vivo* lifetime, and CCR values in *Arabidopsis* vary from $<10^2$ to $>10^6$ [21]. Histidinol dehydrogenase (CCR = 33 000) has recently been engineered to extend its lifespan [22]. Another short-lived enzyme is cobalamin-independent methionine synthase from *Arabidopsis thaliana* (AtMS1 and AtMS2), which catalyze the generation of methionine (Met) from homocysteine and has a CCR that is ~ 100 -fold below the *Arabidopsis* median value of 4×10^5 . At least a part of the short-lived nature of AtMS may arise from its reactive homocysteine substrate and folate product that put it at risk of self-inactivation [21, 23]. Low CCRs may be associated with damage from catalytic misfires or spontaneous reactions between substrates or products and vulnerable residues in the active site region [21, 24]. Such self-inactivation might be reduced by replacing vulnerable residues [24, 25].

Throughout this study, we outline the stages of a CDE campaign where proteomics can be applied. We employed both targeted and untargeted proteomic approaches to monitor AtMS1 and AtMS2 during CDE. By measuring protein abundance and identifying trade-offs and cellular adaptations in the host, our study demonstrates how mass spectrometry-based proteomics can both identify flaws in population construction and provide insights into directed enzyme evolution within living systems. These results highlight the potential of proteomic strategies as powerful tools for troubleshooting and optimizing synthetic biology workflows, including enzyme lifespan extension.

Materials and methods

Plasmid construction

Gene sequences are given in Table S1. Primer sequences used for amplification and sequencing were purchased from Eurofins Genomics (Louisville, KY) and are listed in Table S2. Non-recoded cDNAs of *Arabidopsis* AtMS1 (AT5G17920) and AtMS2 (AT3G03780) were obtained from the *Arabidopsis* Biological Resource Center (<https://abrc.osu.edu/>), G13034 and C104738. MET6 was amplified from genomic DNA of *Saccharomyces cerevisiae* strain BY4742. ORF

of AtMS1, AtMS2 and MET6 were cloned into plasmids ArEc-TDH3 and GR-306MP as previously described [26]. Enzymes for polymerase chain reactions (PCR) and cloning were obtained from Thermo Fisher Scientific (Miami, FL).

Yeast strain construction and media

Yeast strain BY4742 *met6Δ* (MAT α ; *his3Δ1*; *leu2Δ0*; *lys2Δ0*; *ura3Δ0*; YER091c::kanMX4) was obtained from Euroscarf (Oberursel, Germany) and grown in YPD medium as previously described [26]. GA-Y319 was cultivated as described previously [26]. Plasmid ArEc-TDH3 containing AtMS1, AtMS2, or MET6 was transformed into BY4742 *met6Δ* using the lithium acetate (LiAc) method with minor modifications [27]. A 5-ml YPD preculture was inoculated with BY4742 *met6Δ* and grown to saturation at 30°C. Five millilitres of YPD per transformation was inoculated with the preculture to set an OD₆₀₀ of 0.3 and grown to an OD₆₀₀ of 0.6. Cells were pelleted by centrifugation (5 min, 1500 g), washed with 5 ml of Milli-Q water per 5 ml of culture, recentrifuged, washed with 150 μ l of TE/LiAc (10 mM Tris-HCl, 1 mM Na₂EDTA, pH 7.5, 100 mM LiAc), and then mixed with 50 μ l of TE/LiAc per 5 ml of culture. Ten millilitres of 11 mg/ml of single-stranded salmon sperm DNA were mixed with 300 ng of plasmid DNA and added to the cells. After adding 600 μ l of 50% w/v PEG 3350 plus 100 μ M of LiAc and 100 μ M of TE (10 mM Tris-HCl, 1 mM Na₂EDTA, pH 7.5), the mix was incubated for 45 min at 30°C with agitation, and afterwards for 20 min at 42°C. Cells were then pelleted by centrifugation (5 min, 700 g) and resuspended in 1 ml of Milli-Q water. Selection was done in synthetic complete (SC) minus histidine for 3 d at 30°C. Selection media were SC-His (6.7 g/l YNB with ammonium sulfate, 1.4 g/l Dropout mix synthetic minus histidine, 20 g/l Bacto Agar (when needed), 2% w/v glucose) for cells containing ArEc-TDH3 plasmids, SC-Leu (6.7 g/l YNB with ammonium sulfate, 1.4 g/l Dropout mix synthetic minus leucine, 20 g/l Bacto Agar (when needed), 2% w/v glucose) for GA-Y319 containing p1_MS, and SC-His-Trp-Leu-Met-Cys (6.7 g/l YNB with ammonium sulfate, 1.4 g/l Dropout mix synthetic minus histidine, tryptophan, leucine, Met and cysteine, 20 g/l Bacto Agar (when needed), 2% w/v glucose) for BY4742 *met6Δ* containing p1_MS. Strains designed for validation of mutant variants were grown in SC (6.7 g/l YNB with ammonium sulfate, 1.4 g/l Dropout mix synthetic minus histidine, tryptophan, uracil, Met, and cysteine, 20 g/l Bacto Agar (when needed), 2% w/v glucose). Transformation of GA-Y319 was done with Scal-digested GR-306MP harbouring AtMS1 or AtMS2 as previously described [28]. Single colonies were selected for total deoxyribonucleic acid (DNA) isolation and sequencing to confirm integration of the MS genes and the leucine selection marker in the p1 plasmid (1). Primers used are listed in Table S2. Sequence-verified clones were used as the donor strain for protoplast fusion and BY4742 *met6Δ* as recipient. Protoplast fusions were made as before (1). Single colonies were selected for total DNA isolation and sequencing to confirm integration of the MS mutant variants and the uracil selection marker in the p1 plasmid (1). Primers used are listed in Table S2.

Complementation and growth assays

For complementation assays, single colonies of BY4742 *met6Δ* transformed with ArEc-TDH3 containing AtMS1, AtMS2 or MET6 were grown in 5 ml SC-His at 30°C until saturation. Cells with ArEc-TDH3 only were used as negative control. An aliquot of the cultures was used to inoculate triplicate of 5 ml SC-His at a starting OD₆₀₀ of 0.05. Cultures were grown at 30°C for 5 d, and the OD₆₀₀ was monitored. Growth assays were done similarly but

in SC medium minus methionine, cysteine, histidine, tryptophane, and leucine plus 1 mM SeMet.

Evolution campaigns

Nine independent starting populations expressing AtMS1 or AtMS2 were grown for 29–35 serial passages on Met-free medium containing SeMet. As growth improved, populations were typically split into two or more subpopulations to guard against accidental loss and to enable different evolutionary trajectories from that point forward [1]. Three-millilitres of starter populations were set at OD₆₀₀ = 0.05 in SC minus methionine, cysteine, histidine, tryptophane, and leucine, and cultivated in Corning deep 24-well plates (Millipore Sigma, catalogue no. AXYPDW10ML24C). Plates were covered with Breathe-Easier Sealing Film, Diversified Biotech (Dedham, Massachusetts). Cultivation was performed at 30°C with humidification at 80% and agitation at 800 rpm in an Infors Multitron shaker (Sulzemoos, Germany). SeMet was added from the start at a concentration of 5 μ M. Populations were subcultured into fresh medium when they reached saturation. SeMet concentration in the medium was ramped up whenever the evolving populations reached saturation after two days of cultivation. Evolving populations were split at passages 6, 13, and 29. Bulk cultures were analysed for mutations by sequencing PCR amplicons containing the MS ORF and promoter.

DNA isolation and sequencing

Five-millilitre cultures of evolved populations were used for total DNA extraction as previously described or as follows (1). Cells from a 100 μ l aliquot of cultures were mixed with 100 μ l of 200 mM lithium acetate plus 1% SDS, and incubated for 10 min at 70°C. 300 μ l of ethanol was added, mixed, and centrifuged for 3 min at 15 000 $\times$ g. The pellet was washed with 500 μ l of 70% ethanol and centrifuged again. After drying for 5 min at 50°C, the pellet was resuspended in 100 μ l of milliQ water and centrifuged again for 15 sec. 90 μ l were then transferred to a new tube and used as template for PCR amplification and subsequent Sanger sequencing. Primers used are listed in Table S2.

Analysis of free amino acid pools

Precultures of unevolved and evolved populations were grown in 3 ml SC minus histidine, tryptophane, and leucine, and cultivated until saturation in Corning deep 24-well plates as described above. Aliquots of precultures were inoculated into 3 ml SC minus methionine, cysteine, histidine, tryptophane, and leucine at an OD₆₀₀ of 0.05, and cultivated as before until they reached late log phase. Per culture, a liquid amount, which equals around 300 Mio. cells, was harvested by centrifugation, and the cells were stored at –80°C until extraction. Free amino acids were extracted from the yeast cultures as described previously with the following modifications [29]. 500 μ l of 12:5:1 methanol-chloroform-water was added to the frozen cells and mixed thoroughly. Three separate samples were used to determine the spike recovery. They were split into two each, and one aliquot per sample was mixed with 50 μ l of a 2.5 mM Met solution. The downstream procedure was as for all other samples, but with half of the used volumes. The samples were incubated for 5 min at 50°C, mixed thoroughly again, and centrifuged for 1 min at 21 000 $\times$ g. The supernatant was transferred to a new tube, and the remaining cells were extracted again. 250 μ l of chloroform and 375 μ l of Milli-Q water were added to the pooled supernatants and mixed by inverting the tubes. The mix was centrifuged for 5 min at 21 000 $\times$ g, and 1 ml of the aqueous upper phase was recovered and dried in a speed vac. Derivatization of amino acids was

done as described previously with minor changes (4). The dried extracts were resolved in 100 μ l of Coupling solution (10:5:2:3 Acetonitrile: Pyridine: Triethylamine: H₂O), dried again, resolved in 100 μ l of Coupling solution, and mixed with 5 μ l PITC (Edman's Reagent; phenylisothiocyanate) (Thermo Fisher Scientific). The solution was then incubated for 5 min at room temperature, dried, and the extract was finally resolved in 125 μ l Sample buffer (0.071 g anhydrous sodium dihydrogen phosphate in 100 ml of Milli-Q water, adjusted pH to 7.4, mixed with 5.26 ml acetonitrile). Samples were centrifuged before high-performance liquid chromatography (HPLC) analysis. Reverse-phase HPLC analysis of the free Met pool was carried out as described [30] using a Waters™ (Milford, MA) 2695 Separations Module coupled to a Waters™ 2998 PDA Detector. Chromatographic separation was performed on a Supelco™ Discovery® C18 column (Merck KGaA, Darmstadt, Germany) at 25°C. The injection volume was 5 μ l. The flow rate was 1 ml/min, and the gradient was as follows. 10 min at 100% eluent A (131 mM sodium acetate, 3.38 mM triethylamine, adjusted to pH 5.7 with acetic acid, v/v 0.06% acetonitrile, 2.4 μ M EDTA), 30 s at 46% eluent B (3:2 acetonitrile:water), 2 min at 100% B, 8 min at 100% A. Detection was done at a wavelength of 254 nm. Met identification was done based on comparison with an authentic standard (Merck KGaA, Darmstadt, Germany), and concentrations were calculated based on a standard curve and spike recovery rates.

Protein expression in *Escherichia coli* and purification

For spike-in standard, *A. thaliana* AtMS1 and AtMS2 genes were synthesized with codon-optimization for expression in *E. coli* and then cloned into pET-15b vector at the NdeI and XhoI restriction sites. The pET-15b-AtMS1 and pET-15b-AtMS2 constructs were expressed in *E. coli* BL21(DE3) using standard protocols as previously described [31]. Briefly, cells were cultured at 37°C with shaking in LB medium containing 100 μ g/ml ampicillin and supplemented with 0.5 mM ZnSO₄ until the OD₆₀₀ of the culture reached 0.6. Protein overexpression was induced by adding filter-sterilized isopropyl β -D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.2 mM for AtMS1 and 0.4 mM for AtMS2. The culture was cooled, and cells were grown with shaking at 18°C overnight and then treated with 100 μ g/ml of chloramphenicol 1 hour prior to harvesting to help recover soluble protein from inclusion bodies [32]. Cells were collected via centrifugation at 4°C and stored at –80°C.

Harvested cells were resuspended in lysis buffer (300 mM sodium chloride, 50 mM HEPES, pH 7.4). The cells were then incubated with 1 mg/ml lysozyme at 4°C for 1 h, and then lysis was completed via sonication. The crude lysate was centrifuged (20 min, 30 000 $\times$ g, 4°C), and the pellet was discarded. The clarified supernatant was mixed with Ni²⁺-NTA resin (Millipore Sigma, cat. no. P6611) and incubated at 4°C on a nutator for 30 min. The slurry was used to pour a column, and AtMS1/2 was eluted using elution buffer (300 mM sodium chloride, 50 mM HEPES, pH 7.4, and 250 mM imidazole, pH 7.0). The eluted MS proteins were incubated with thrombin (MP Biomedicals, cat. no. 154163) at a ratio of 1.5 U/mg of protein and then dialyzed against 1 l of 50 mM ammonium bicarbonate at 4°C overnight. The protein was passed through a Ni²⁺-NTA column to remove any protein that still retained the hexahistidine tag, followed by benzamidine-sepharose resin (Cytiva, cat. no. 17512310) to remove thrombin. Protein purity was determined using Coomassie-stained SDS-PAGE. The purified MS proteins were concentrated to 17 mg/ml as determined by UV-visible spectrophotometry using extinction coefficients at 280 nm (ϵ_{280}) calculated by ProtParam (Expasy) [33]

and then flash-frozen in 50–100 μ l aliquots in liquid nitrogen for storage at –80°C. These were used to verify peptides for targeted MS analysis and as spike-in standards to allow absolute quantification.

Development of AtMS1 and AtMS2 standards for targeted proteomic analysis

To enable absolute quantification of Arabidopsis Met synthase in yeast lysates, purified AtMS1 and AtMS2 proteins were used as spike-in standards. They were expressed in *E. coli* and purified as described above. The identity and purity of the recombinant proteins were confirmed via SDS-PAGE and mass spectrometry. Known concentrations of recombinant AtMS1 and AtMS2 were spiked into *S. cerevisiae* BY4742 lysates that did not contain the construct expressing AtMS1 and AtMS2 for CDE. To ensure specificity, peptides unique to AtMS1 and AtMS2, distinct from the yeast MET6 protein, were selected based on *in silico* digestion and peptide uniqueness analysis. These peptides were used for downstream MRM assay development and quantification. A series of spike-in concentrations was used to construct a standard curve by correlating MS signal intensities with known amounts of AtMS1 and AtMS2. The standard curves enabled absolute quantification of the target proteins in experimental samples by comparing the MRM signal intensities of expressed AtMS1 and AtMS2 peptides to those of the spike-in standards.

Cultivation and harvest of evolution populations for proteomics analysis

Evolved yeast populations and their corresponding starter populations were grown in selective medium in Corning deep 24-well plates as described above. Per population, 12 cultures $\times$ 3 ml were harvested at an OD₆₀₀ between 2.5 and 3.7 and pooled. The cells were harvested by centrifugation at 4°C and washed with 15 ml ice-cold MilliQ water. The final pellet was resuspended in 1 ml ice-cold milliQ water, flash-frozen in liquid nitrogen, and freeze-dried using a Labconco™ FreeZone™ 2.5 l–50°C Benchtop Freeze Dryer (Catalogue no. 710201000).

Sample preparation for mass spectrometry analysis

Proteins of interest, Met synthases AtMS1 and AtMS2, were analysed from purified proteins and dried yeast extracts, including 84 protein samples derived from unevolved and evolved populations.

Protein extracts were prepared by incubating 10 mg dried yeast samples in 200 μ l SDS extraction buffer (7% SDS, 0.1 M Tris, 5 mM TCEP, 1x protease inhibitor, pH 7.5), followed by 10 min incubation at 95°C and 1000 rpm. Samples were transferred for a 1-min treatment in a sonicator water bath to enhance lysis. The samples were centrifuged for 10 min at max speed (~18 000 g), and the supernatant was transferred into a 96-well plate for further processing.

Protein quantification and digestion

Protein quantification was performed using a bicinchoninic acid (BCA) assay to estimate protein concentration. Proteins were digested using an SP3-aided protocol [34], employing trypsin at a 1:50 enzyme-to-substrate ratio. Digestion was carried out at 37°C for 18 h to ensure complete cleavage. From each sample, 10 μ l of the prepared solution was injected for analysis.

Mass spectrometry analysis

For untargeted analysis, protein samples were injected into an online nanoflow system (0.4 μ l/min) using a capillary column

(Picofrit, 50 μm tip opening, 75 μm diameter, New Objective, ICT36015030F-50) packed in-house with 15 cm of C18 silica material (3 μm , Dr. Maisch GmbH). This was connected to a Thermo Fusion mass spectrometer in-line with a Dionex Ultimate 3000 series UHPLC. The spray voltage was set to 2 kV, and the heated capillary temperature was maintained at 275°C. For data-dependent acquisition (DDA) experiments, the full MS resolution was set to 60 000 at m/z 200, with an AGC target of 300% and auto injection time. The mass range was set from 350 to 1500 m/z . The AGC target value for fragment spectra was set to 'standard' with a resolution of 15 000 and auto-injection time. An intensity threshold of 5×10^4 was applied, and the isolation width was set at 1.6 m/z . Normalized collision energy was set to 30%. The column flow rate was 0.4 $\mu\text{l}\cdot\text{min}^{-1}$ with the following elution profile: 0–0.5 min 2% (v/v) ACN, 0.1% (v/v) FA to 6% (v/v) ACN, 0.1% (v/v) FA; 0.5–25 min 6% (v/v) ACN, 0.1% (v/v) FA to 27% (v/v) ACN, 0.1% (v/v) FA; 25–26 min 27% (v/v) ACN, 0.1% (v/v) FA to 35% (v/v) ACN, 0.1% (v/v) FA; 26–26.1 min 35% (v/v) ACN, 0.1% (v/v) FA to 80% (v/v) ACN, 0.1% (v/v) FA; 26.1–26.5 min 80% (v/v) ACN, 0.1% (v/v) FA, 26.5–26.7 min 80% (v/v) ACN, 0.1% (v/v) FA to 2% (v/v) ACN, 0.1% (v/v) FA; 26.7–30 min 2% (v/v) ACN, 0.1% (v/v) FA.

Raw files were processed using the MaxQuant software (Version 2.4.0.0, [35]) using default settings. Trypsin was used as protease and up to two allowed missed cleavages. Cysteine carbamidomethylation was set up as a fixed modification, and oxidation of Met and N-terminal acetylation as a variable modification. Peptide identification required a minimum of seven peptides, and a maximum of five modifications was allowed. Match-between-run function was enabled. Global data normalization was done using the MaxQuant label-free quantitation (LFQ) algorithm with LFQ minimum ratio count set to two and fast LFQ enabled. The following fasta files were used: sequences for the evolved proteins, plus the reference fasta file UP000002311 for *S. cerevisiae* was downloaded from Uniprot. MaxQuant output files were used to further analyse and visualize in R.

For targeted analysis by multiple reaction monitoring, digested peptides were loaded onto an Aeris 3.6 μm peptide XB-C18 100 Å, LC column 50 $\times$ 2.1 mm (Phenomenex) using a Thermo UltiMate 3000 RSLCnano System coupled to an Thermo TSQ Altis Triple Quadrupole MS. The column was heated to 55°C and the column flow rate was 0.4 $\text{ml}\cdot\text{min}^{-1}$ with the following elution profile: 0–20.5 min 2% (v/v) acetonitrile, 0.1% (v/v) formic acid to 22% (v/v) acetonitrile, 0.1% (v/v) formic acid; 20.5–27.5 min 22% (v/v) acetonitrile, 0.1% (v/v) formic acid to 35% (v/v) acetonitrile, 0.1% (v/v) formic acid; 27.5–28.5 min 35% (v/v) acetonitrile, 0.1% (v/v) formic acid to 97% (v/v) acetonitrile, 0.1% (v/v) formic acid; 28.5–29.2 min 97% (v/v) acetonitrile, 0.1% (v/v) formic acid, 29.2–29.5 min 97% (v/v) acetonitrile, 0.1% (v/v) formic acid to 2% (v/v) acetonitrile, 0.1% (v/v) formic acid; 29.5–35 min 2% (v/v) acetonitrile, 0.1% (v/v) formic acid. The list of peptide transitions used for selected reaction monitoring (SRM-MS) is provided in Table S3. Peak area of ions for targeted peptides was determined using the Skyline software package version 24.1.0.199 (reference for Skyline: <https://pubmed.ncbi.nlm.nih.gov/20147306/>). Purified AtMS1 and AtMS2 proteins, which were expressed and extracted from *E. coli*, were used as internal standards. The protein standards were spiked into yeast lysates at known concentrations to generate a standard curve as described above. Absolute protein concentrations in the yeast samples were subsequently determined by comparing the MS signal intensities of recombinant proteins to those of the spiked-in standards. Further data analysis was performed using R.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD069101 and 10.6019/PXD069101 [36].

Results and discussion

Host selection and validation for directed evolution

The selection of a platform host depends on the target gene's origin and the destination organism, considering subcellular localization of the target protein, posttranslational modifications (PTM), intracellular pH, redox conditions, and oxygen levels. PCR and phenotypic tests (e.g. growth assays in defined media) are common techniques to validate the chosen background strain (e.g. a knockout mutant in *S. cerevisiae*). However, complementing these analyses with targeted proteomics can confirm strain suitability and verify the presence and absence of proteins (Fig. 1A). In addition, untargeted proteomics can reveal cellular changes that may influence selection during the evolution experiment.

We selected *S. cerevisiae* as the evolution host due to its 'plant-like' cellular conditions, including compartmentalization, metabolite and cofactor availability, and shared metabolic pathways. For our approach, the AtMS1 and AtMS2 genes were codon-optimized for yeast and expressed from the plasmid ArEc-TDH3 in a *met6Δ* knockout mutant (e.g. yeast Met synthase deleted) to perform complementation assays in a methionine-deficient (–Met) medium (Fig. S1). Targeted proteomics can support this validation step if the desired phenotype is not achieved, enabling the assessment of (i) protein abundance, (ii) solubility, and (iii) interactions with the cellular machinery.

Generation of evolution starter populations and verification

When selecting a CDE system, mutation rate and bias, and overall system design should align with the experiment's objectives. For this study, we utilized the OrthoRep system due to its ability to achieve an orthogonal replication mutation rate between 10^5 and 10^6 times higher than the genomic mutation rate [7, 37]. OrthoRep comprises an orthogonal error-prone polymerase expressed from a nuclear plasmid, and two cytosol-localized linear plasmids (p1 and p2) that facilitate self-replication, encode an RNA-polymerase, a marker gene, and the gene of interest (Fig. 2A). To assemble this system (Fig. 2B), AtMS1 or AtMS2 and the marker gene were first integrated into the p1 plasmid in the strain GA-Y319 [26]. Successful integration was confirmed by estimating the p1 plasmid size on an agarose gel (Fig. S2A) and by verifying the target gene via Sanger sequencing. In addition, we recommend that targeted proteomics analysis, as described earlier, be employed here to confirm the target protein's expression and functionality in the OrthoRep setup.

To transfer the OrthoRep machinery to the chosen platform host, we performed protoplast fusion between the obtained donor strain carrying the p1 plasmid and the *met6Δ* recipient strain with a plasmid expressing the error-prone polymerase (Fig. 2B) [26]. Protoplast fusion clones were verified by gel analysis (Fig. S2B) and Sanger sequencing of the target gene on p1 and complementation assays in –Met medium (Fig. S2B; Fig. 3). Proteomics can be employed here to assess whether the yeast protein machinery, including the CDE system and the protein of interest, shows any unexpected alterations (Fig. 1B; Fig. 2).

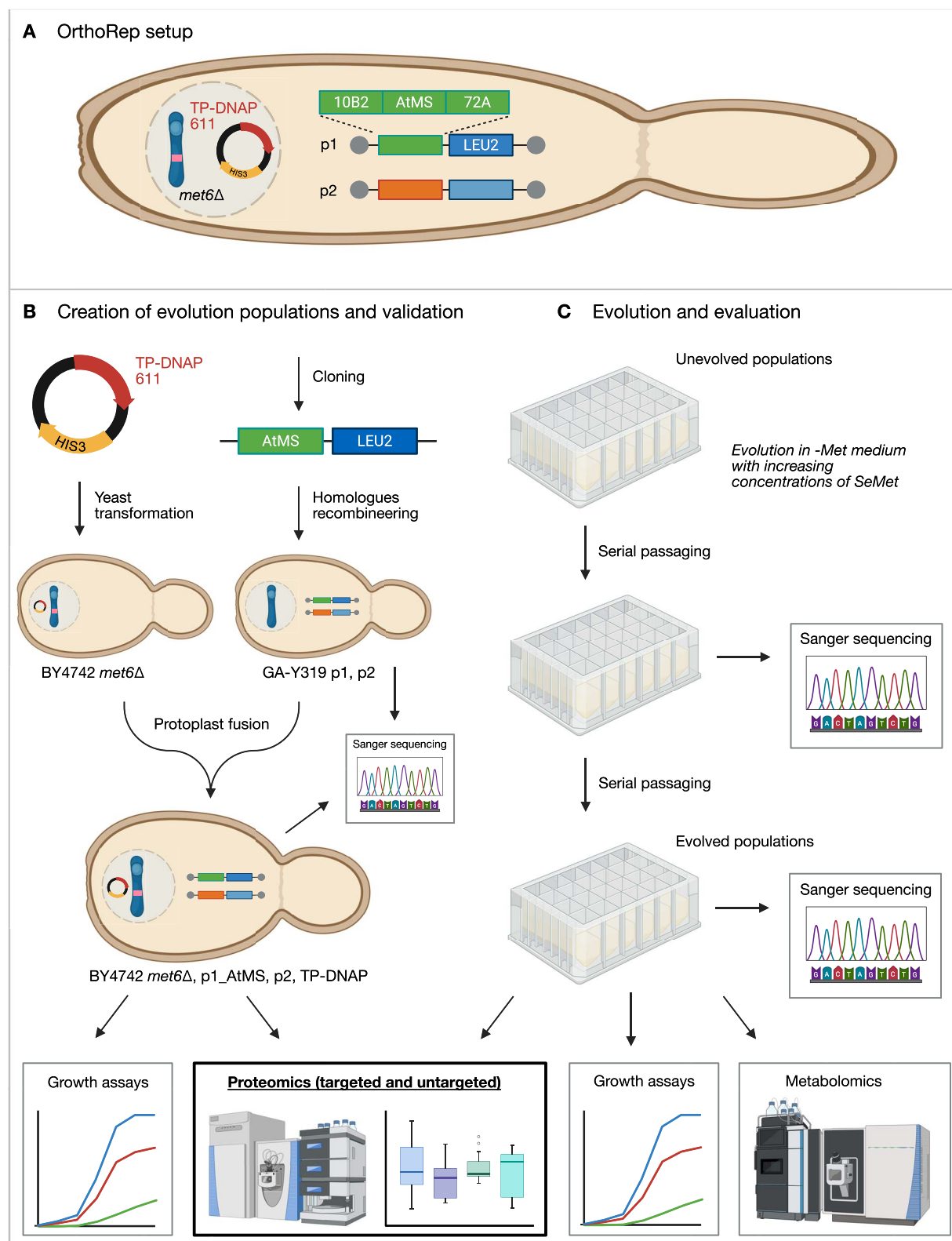


Figure 2. Assembly of OrthoRep cells and workflow of the presented study. A. Evolution populations were created in a BY4742 *met6Δ* yeast background. A nucleus-localized plasmid under control of a *HIS3* marker encoded the error-prone polymerase TP-DNAP1 611. p1 encodes the *AtMS* target gene and a *LEU2* selection marker. p2 encodes the RNA-polymerase and a p2-specific DNA-polymerase. B. Evolution starter populations were created and later verified by Sanger sequencing, growth assays, and proteomics (targeted and untargeted). C. Accumulated mutations in target genes in evolved populations were determined by Sanger sequencing, and cellular phenotypes were evaluated by growth assays, metabolomics, and proteomics (targeted and untargeted). The figure was created in BioRender. Bathe, U. (2025) <https://BioRender.com/gq2uf15>.

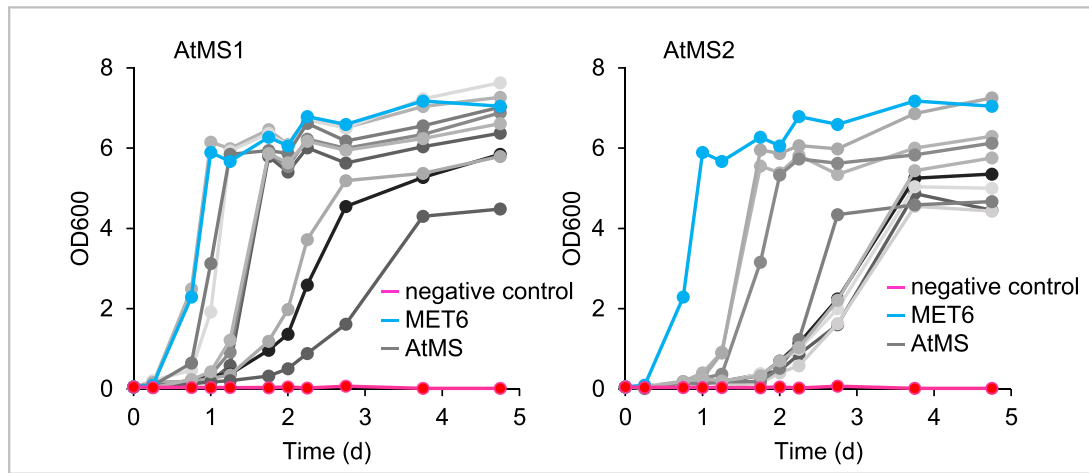


Figure 3. Test of evolution starter populations. Complementation of BY4742 *met6Δ* with AtMS1 and AtMS2 expressed from p1 after protoplast fusion. Strains having the p1 with no AtMS gene served as negative control but were supplemented with Met in the growth medium as positive control.

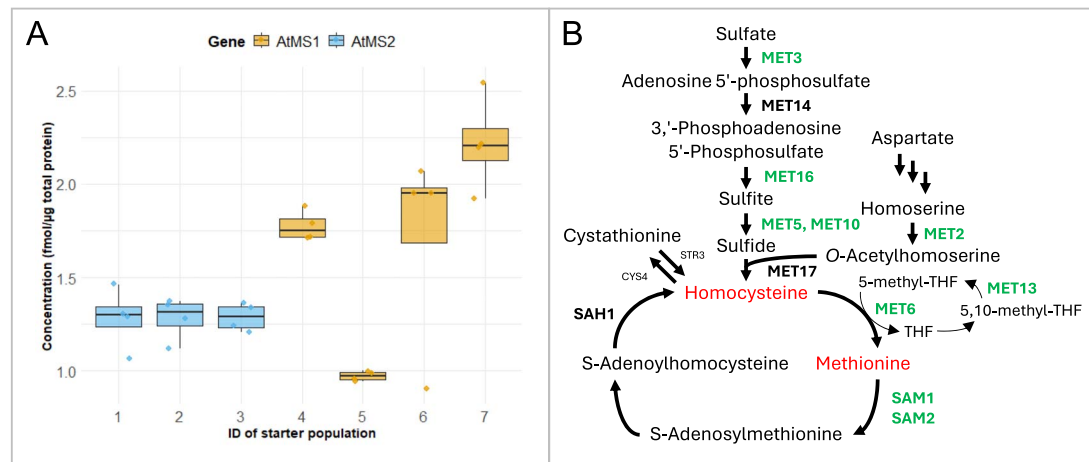


Figure 4. Proteomics of Arabidopsis AtMS enzymes expressed from p1 and MET related enzymes in unevolved yeast populations. A. Targeted protein analysis of Arabidopsis AtMS enzymes in evolution starter populations; $n=4$. B. Illustration of the Met biosynthesis pathway in yeast. Identified Met-related enzymes are highlighted in green. Additionally, MET18 (part of the cytosolic iron-sulphur protein assembly), MET22 (bisphosphate-3'-nucleotidase), MET8 (involved in the biosynthesis of siroheme), and METW were detected. The figure was modified from [59].

For our study, four AtMS1 and three AtMS2 starter populations were selected, and we first performed targeted proteomics using multiple reaction monitoring mass spectrometry (MRM-MS; Fig. 2B) to confirm the presence and accurately measure the abundance of AtMS1 and AtMS2 based on spike-in standards (Table S3). We confirmed the presence of AtMS1 and AtMS2 proteins, which were expressed between 1 and 2.5 fmol/ μ g total protein (Fig. 4A). While AtMS2 is uniformly expressed among the three investigated evolution starter populations, AtMS1 expression differs between the populations. It demonstrates that multiple populations should be used to start an evolution campaign; the differences most likely come from different copy numbers of p1 [7].

Untargeted proteomics was then used to investigate cellular changes in the host cells (Fig. 2B). We focused our analysis on Met biosynthesis and related pathways (Fig. 4B; Table S4; Fig. S3) that shows abundance of MET10, MET16, MET18, MET2, MET22, MET3, MET5, MET6, MET8, SAM1, SAM2, and METW in unevolved populations. Interestingly, the overall trend is that Met enzymes are more abundant in populations of AtMS2 compared to AtMS1. This illustrates that expression of foreign enzymes in a CDE

system can have different impacts on the host cells' metabolism and their adaptation, even though such enzymes share high sequence identity, e.g. AtMS1 and AtMS2 proteins are 92.81% identical.

Unexpectedly, we also found that MET6 was detected in the protoplast fusion clones, despite the evolution starter populations being generated from a *met6Δ* knockout strain (Fig. 2B). This likely resulted from chromosomal contamination during the fusion process. Rather than excluding these clones, we continued with them to showcase how proteomic analysis can reveal such unexpected events and add value to a CDE campaign, as detailed below.

Monitoring ongoing evolution

Running CDE experiments is considered hands-off and low-effort venture. Beneficial mutations in the target gene are selected when the target's performance is coupled to growth or another readable phenotype. Over time, these mutations accumulate, and the evolution is stopped when a desired phenotype is achieved. In practice, this means serial passaging (Fig. 2C), typically every couple of days, with growth phases in between. To evolve AtMS1

and AtMS2, we used a selection strategy in which the concentration of the toxic Met analogue L-selenomethionine (SeMet) in the medium was gradually raised from 5 μ M to 1 mM (Fig. S4). SeMet toxicity is believed to be attributed to the depletion of reduced thiol compounds [38]. The toxic effect can be out-competed by supplying Met or cysteine [38–40]. Therefore, we hypothesize that this strategy can favour increased Met synthase-driven flux to Met, which expands the free Met pool and so competes out the analogue's toxicity. A similar approach has been applied to other metabolites in directed evolution before [22, 41]. Although the starting populations contained MET6, therefore were not exclusively dependent on AtMS1 and AtMS2 expressed from p1, we argue that improved AtMS1 and AtMS2 can support flux towards Met production above and beyond the endogenous MET6 capability and therefore, can contribute to the anticipated goal of outcompeting the toxic effect of SeMet.

Starting populations expressing AtMS1 or AtMS2 were grown for 29–35 serial passages on Met-free medium containing SeMet. During the campaigns, it is recommended to collect sequencing data of the target gene to verify that mutations accumulate (Fig. 2C). Such information may also help to follow the evolutionary trajectory. While such analysis is limited to sequence information, proteomics is a complementary tool to monitor the actual phenotype (Fig. 1C). For example, targeted proteomics can be used to assess changes in the abundance of proteins of interest. Additionally, untargeted proteomics can help identify global cellular responses to the selection regime, revealing potential off-target effects (e.g. genomic mutations) or adaptive mechanisms in the host (e.g. epigenetic adaptation). Such data can provide critical feedback to guide adjustments in experimental conditions and ensure that the desired phenotypic traits are being optimized effectively.

Validation of evolved phenotypes

Once a desired phenotype is achieved, the evolution is stopped, and evolved sequences are verified (Fig. 2C). We halted the AtMS campaigns at a final SeMet concentration of 1 mM as complex secondary toxic effects are more likely at higher concentrations [38, 42]. In total, 72 populations were obtained. Figure 5A shows the growth of six representative evolved populations compared to the corresponding unevolved populations in 1 mM SeMet. To verify increased intracellular Met concentrations, we screened the 72 evolved populations in single replicates (Fig. 5B). Cellular free Met levels of 1–2 mM in most unevolved populations and a marked upshift in evolved populations towards levels as high as 10 mM were observed, proving that SeMet treatment can favour increased Met synthase-driven flux to Met. Therefore, we chose eight and six populations with AtMS1 and AtMS2, respectively, to measure free Met pools in triplicate to gain statistical confidence about our data (Fig. 5C). Consistent with our previous screening, we confirmed high Met-producing and low Met-producing populations. Free Met pools may differ slightly from the first evaluation of populations due to the timepoint of harvest.

To determine what changes accumulated in CDE, the mutated sequences were isolated from the populations and subjected to sequencing. The chosen AtMS populations from the Met pool measurements were used to detect mutations (Fig. S5). Mutant AtMS variants were recovered having a mix of nonsynonymous, synonymous, and promoter mutations. Most of them replaced the wildtype base in all gene copies of the population, proving selection was applied to AtMS1 and AtMS2 on p1.

As outlined above, proteomics can be applied to validate evolved phenotypes by directly assessing changes in abundance

of proteins as outcomes of the evolution process in a targeted and untargeted manner (Fig. 1C; Fig. 2). With that, broader cellular adaptations can be identified that might accompany the evolved phenotype. In our study, we hypothesized that longer AtMS enzyme lifetime (i.e. higher protein stability) would correlate with higher protein abundance in evolved populations. To further assess protein longevity, approaches such as pulse-chase SILAC and ^{15}N -labelling could be employed to directly measure turnover rates of AtMS1 and AtMS2 [43–45]. As an initial test for abundance, we used targeted proteomics to measure AtMS in evolved populations compared to unevolved ones (Fig. 6A). Lower AtMS protein levels were detected, indicating that accumulated major mutations did not correlate with evolved cellular phenotypes, e.g. elevated free Met pools and improved growth in medium with SeMet. While improved enzyme kinetics can, in principle, be the reason for increased Met flux at lower protein abundance, such improvements of central metabolic enzymes are likely modest [46, 47]. Instead, the presence of MET6 caused relaxed selection on AtMS1 and AtMS2. As protein expression from p1 can be considered a metabolic burden to the cell, expression might be reduced if the protein is not required for survival. Therefore, we think that lower AtMS1 and AtMS2 expressions in our settings could have resulted in a metabolic fitness benefit. If the evolution campaign had continued, AtMS expression might have been completely switched off, resembling natural gene loss as a process of adaptation [48].

We then used an untargeted proteomics approach to investigate broader host strain adaptation and investigated any changes specifically in the Met biosynthesis and related pathways of evolved populations (Table S4; Fig. S6; Fig. 6B and C). Principal component analysis revealed partial separation of evolved and unevolved populations, with AtMS1 and AtMS2 samples clustering in distinct regions (Fig. 6B). While evolved and unevolved populations showed some divergence along PC1 and PC2, there was also substantial overlap, particularly within AtMS2 samples. This indicates that although evolutionary pressure led to measurable adaptation in proteomic profiles, the extent of divergence was only modest. Notably, AtMS1 samples displayed greater dispersion, whereas AtMS2 samples clustered more tightly.

In total, 88 proteins showed significant differences in abundance across evolution and AtMS groups with a P value < 0.05 and $\log_2$ fold change > 2 . The heatmap of these proteins reveals distinct patterns separating evolved from unevolved populations (Fig. 6C; Table S5). These findings indicate a remodelling of protein abundance during CDE and suggest that proteomic adaptation involved both shared and AtMS-dependent responses.

When comparing Met-related enzymes, we identified adaptations in abundance that correlate with increased free Met pools, e.g. the abundance of Met biosynthesis enzymes MET3, MET16, MET10/MET5, MET2, and MET6 (along with related Met enzymes MET8 and METW) were increased in AtMS1 populations (Fig. S6). Interestingly, these enzymes were not increased in the AtMS2 populations. We therefore hypothesized that the increased free Met pool in AtMS2 populations came from improvements introduced into AtMS2. The isolated AtMS2 variants were tested in fresh cells, but no SeMet resistance was observed, and the origin of the improved phenotype in AtMS2 populations remains unclear. However, we demonstrated that the applied selection pressure affected the host's metabolic network and the proteome, and the findings highlight the complementary strengths of targeted and untargeted proteomics for a better understanding of evolved phenotypes and system-wide effects.

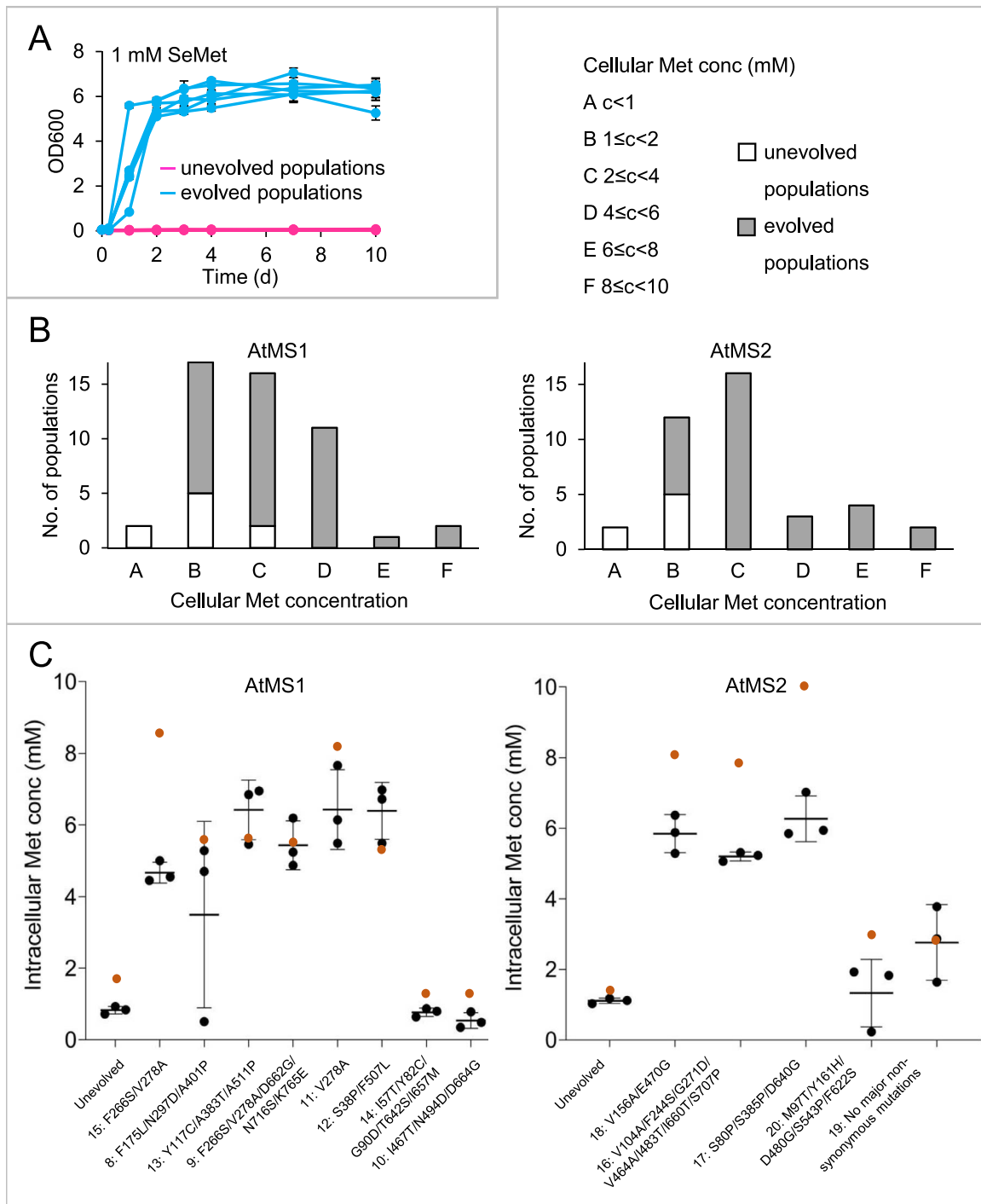


Figure 5. Analysis of evolved AtMS populations. A. Growth comparison of evolved and unevolved populations in presence of 1 mM SeMet. B. Free intracellular Met pools of all unevolved and evolved populations; $n = 1$. C. Evolved populations with the highest Met content from B (orange dots) were subjected again to free Met pool measurements; $n = 3$. Representative unevolved populations are given for comparison.

Directed evolution-acquired mutations in an enzyme might impede with its *in vivo* regulation, e.g. feedback inhibition, which can cause increased product formation. To our knowledge, there is no feedback inhibition of Met synthase, and we argue that there is likely none because accumulation of its reactive substrate homocysteine is toxic to the cell [49]. Instead, it is more likely that the Met biosynthesis is feedback-regulated earlier in the pathway [50].

In addition, one might consider other possible scenarios that can influence the outcome of an evolution campaign: (i) The copy number of the p1 plasmid can vary per cell and is dependent on the error-prone polymerase used [51], therefore, can account for differences between evolved populations or between unevolved and evolved populations. Here, qPCR can help and was used successfully in the past to verify p1 copy numbers [7, 52]. (ii) Synonymous mutations can accumulate and improve enzyme

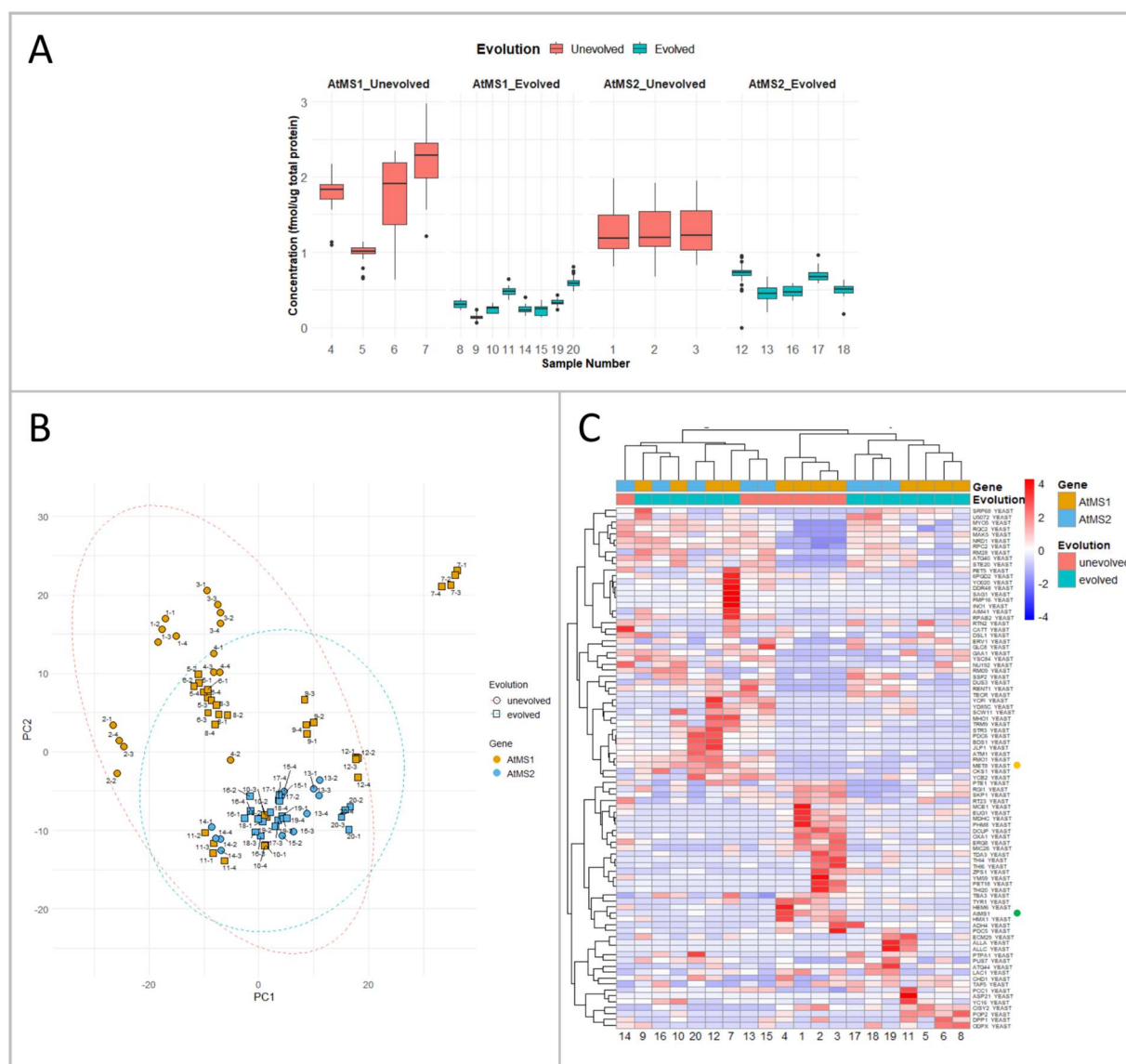


Figure 6. Protein analysis of AtMS1 and AtMS2 in evolved and unevolved yeast populations. A. Concentrations of AtMS1 and AtMS2 proteins in evolved and unevolved yeast populations (labelled 1–20), quantified by mass spectrometry; $n = 4$. B. Untargeted proteomics analysis of both evolved and unevolved populations expressing AtMS1 or AtMS2. Left, PCA plot illustrates clustering of AtMS1 and AtMS2 samples as well as evolved and unevolved populations. Right, heatmap illustrating row-scaled abundances of significantly changing proteins ($\log_2 FC > 2$ and $P < 0.05$ determined by t-test). AtMS1 (green dot) and MET8 (yellow dot) were among the proteins most significantly changing in abundance based on selected thresholds.

expression to meet the host cell's preference for codon usage (even though the target gene has been codon-optimized) [26, 28]. Non-synonymous mutations in evolved sequences should be tested independently of evolved synonymous mutations to verify their contribution to the evolved phenotype. (iii) Promoter mutations may accumulate and improve expression, in particular in OrthoRep, as the promoter is subjected to hypermutation too [53]. (iv) False-positives might arise in an evolved population due to metabolic crosstalk between cells that allows 'cheaters' to benefit from truly evolved cells [54, 55].

Proteomics offers an unbiased approach to confirm gene deletions and to detect unexpected protein expressions or compensatory changes during evolution experiments. While adding tags like FLAG or HA are useful for targeted detection in many contexts, our label-free method provides a comprehensive view without requiring genetic modification and avoids the tag itself or its

co-evolution affecting results in CDE. Mass spectrometry-based proteomics provides a flexible and unbiased readout for CDE experiments, as it is applicable to virtually any target protein once a reference proteome (or the sequence of a mutated variant) is available [22, 56, 57]. In addition, targeted MS approaches such as MRM allow absolute or near-absolute quantification with high precision and broad dynamic range, surpassing the semi-quantitative nature and antibody dependence of Western blotting. Importantly, proteomics enables simultaneous measurement of the evolving target protein together with its pathway context in the same experiment, ensuring internally consistent comparisons across enzymes and conditions. Our study did not aim for near complete proteome coverage but rather an economical use that could be employed across populations and timepoints in a campaign; the relatively short 26 min gradient and 30 min acquisition time inevitably limited the number of identifications compared

with longer gradients uses in deeper yeast proteomics [58]. Future studies employing extended gradients or data-independent acquisition (DIA) could substantially increase proteome depth and completeness if required, offering complementary insights to the high-throughput MS1-based workflow used in our study. Though, the proteomics approach used here still demonstrated high sensitivity and statistical confidence, enabling the detection and quantification of even very low protein abundances, down to femtomoles per microgram of the yeast background proteome.

Previously, mass spectrometry was used in CDE that evolved compact zinc finger degrons to enable selective, potent, and tunable protein degradation. The authors employed peptide mass spectrometry to confirm the absence of off-target protein degradation, an essential step to ensure the specificity of their evolved constructs [16].

Beyond phenotype validation: additional proteomics approaches

To further enhance the insights gained from CDE, proteomics analyses can be integrated at other stages in the process (Fig. 1D). First, post-translational modification analysis can provide valuable information on how target proteins are regulated or modified in the host or during CDE, potentially revealing novel mechanisms of adaptation (e.g. change in phosphorylation, acetylation, oxidation or glycosylation patterns). Second, interaction proteomics, such as co-immunoprecipitation coupled with mass spectrometry, can be employed to identify changes in protein–protein interaction networks. This is particularly useful for understanding how the evolved phenotype integrates into the cellular machinery. Third, proteomics-based quantification of metabolic enzymes can link phenotypic changes to metabolic shifts, providing a more holistic understanding of how evolution reshapes cellular physiology. These additional analyses can deepen the mechanistic understanding of evolved phenotypes and inform strategies for optimizing the evolutionary process further.

Conclusion

CDE experiments can greatly benefit from implementing mass spectrometry-based proteomics because it uncovers cellular changes which complements what can be detected on the DNA or RNA levels. As shown in our study, it can give valuable insights into the actual phenotype of the platform cells during a campaign. Using *Arabidopsis* AtMS1 and AtMS2 as an example, we demonstrated at what checkpoints both targeted and untargeted proteomics can be applied, what the delivered value would be, and outlined other potential applications during and after evolution. We discovered the value of proteomics to reveal unexpected flaws in population construction through the retention of some MET6 capacity; information that was important to interpret our campaign strategy based on SeMet-dependent selection. It also revealed how different target proteins subjected to evolution can have varying effects on the host cell, even though they share high sequence similarity. Using systems-level analysis, we further identified changes in protein expression of diverse pathways in yeast as a result of the applied selection. Combined, our results demonstrate the benefits of complementing continuous evolution with proteomics analysis.

Acknowledgements

We are exceptionally thankful for advice and guidance by Andrew D. Hanson from the University of Florida. Dr. Hanson has been and

will always be a source of inspiration. We thank Elke Ströher from the WA Proteomics Facility at the University of Western Australia for her support with mass spectrometry analysis.

CRedit statement

Katharina Belt (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing—original draft, Writing—review & editing), David Obe (Investigation, Methodology), Mark A Wilson (Conceptualization, Formal analysis, Funding acquisition, Methodology, Supervision, Writing—original draft, Writing—review & editing), A Harvey Millar (Conceptualization, Formal analysis, Funding acquisition, Supervision, Writing—original draft, Writing—review & editing), Ulschan Bathe (Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Validation, Visualization, Writing—original draft, Writing—review & editing).

Supplementary data

Supplementary data is available at SYNBIO online.

Conflict of interest: The authors declare no competing financial interests.

Funding

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—project number 455236359 to U.B., and the Elite Network of Bavaria—project NFG New tools for a sustainable plant production—CDE to increase CO₂ fixation and biomass in crops to U.B. The proteomics work was supported by the Australian Research Council (project FL200100057 to A.H.M) and infrastructure funding from the Western Australian State Government in partnership with the Australian Federal Government, through Bioplatforms Australia as part of the National Collaborative Research Infrastructure Strategy (NCRIS). M.A.W is partially supported by the Department of Energy (DOE) grant DE-SC0020153.

Data availability

All data generated in this study are available in the main text, in the figures or in the supplemental information. Data are also available on bioRxiv: <https://doi.org/10.1101/2025.03.03.641153>. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD069101 and <https://www.ebi.ac.uk/pride/archive/projects/PXD069101> [36].

Material availability statement

All material used in this study is available upon request and signing a MTA.

References

1. Molina RS, Rix G, Mengiste AA et al. In vivo hypermutation and continuous evolution. *Nature Reviews Methods Primers* 2022;**2**:37. <https://doi.org/10.1038/s43586-022-00119-5>

2. Arnold FH. Innovation by evolution: bringing new chemistry to life (Nobel lecture). *Angew Chem Int Ed* 2019;**58**:14420–6. <https://doi.org/10.1002/anie.201907729>
3. García-García JD, Joshi J, Patterson JA et al. Potential for applying continuous directed evolution to plant enzymes: an exploratory study. *Life (Basel)* 2020;**10**:179. <https://doi.org/10.3390/life10090179>
4. Tian R, Zhao R, Guo H et al. Engineered bacterial orthogonal DNA replication system for continuous evolution. *Nature Chem Biology* 2023;**19**:1504–12. <https://doi.org/10.1038/s41589-023-01387-2>
5. Moore CL, Papa LJ 3rd, Shoulders MD. A Processive protein chimera introduces mutations across defined DNA regions in vivo. *J Am Chem Soc* 2018;**140**:11560–4. <https://doi.org/10.1021/jacs.8b04001>
6. Wang Q, You J, Li Y et al. Continuous evolution of protein through T7 RNA polymerase-Guided Base editing in *Corynebacterium glutamicum*. *ACS Synth Biol* 2025;**14**:216–29. <https://doi.org/10.1021/acssynbio.4c00606>
7. Ravikumar A, Arzumanyan GA, Obadi MKA et al. Scalable, continuous evolution of genes at mutation rates above genomic error thresholds. *Cell*. 2018;**175**:1946–1957.e13. <https://doi.org/10.1016/j.cell.2018.10.021>
8. Esvelt KM, Carlson JC, Liu DR. A system for the continuous directed evolution of biomolecules. *Nature*. 2011;**472**:499–503. <https://doi.org/10.1038/nature09929>
9. Halperin SO, Tou CJ, Wong EB et al. CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. *Nature*. 2018;**560**:248–52. <https://doi.org/10.1038/s41586-018-0384-8>
10. Zhao N, Zhou J, Tao T et al. Evolved cytidine and adenine base editors with high precision and minimized off-target activity by a continuous directed evolution system in mammalian cells. *Nat Commun* 2024;**15**:8140. <https://doi.org/10.1038/s41467-024-52483-3>
11. Morrison MS, Wang T, Raguram A et al. Disulfide-compatible phage-assisted continuous evolution in the periplasmic space. *Nat Commun* 2021;**12**:5959. <https://doi.org/10.1038/s41467-021-26279-8>
12. Packer MS, Rees HA, Liu DR. Phage-assisted continuous evolution of proteases with altered substrate specificity. *Nat Commun* 2017;**8**:956. <https://doi.org/10.1038/s41467-017-01055-9>
13. Rosenthal RG, Diana Zhang X, Đurđić KI et al. Controlled continuous evolution of enzymatic activity screened at ultrahigh throughput using drop-based microfluidics. *Angew Chem Int Ed* 2023;**62**:e202303112. <https://doi.org/10.1002/anie.202303112>
14. Schulz-Mirbach H, Wichmann P, Satanowski A et al. New-to-nature CO(2)-dependent acetyl-CoA assimilation enabled by an engineered B(12)-dependent acyl-CoA mutase. *Nat Commun* 2024;**15**:10235. <https://doi.org/10.1038/s41467-024-53762-9>
15. Vaccari NA, Zevallos-Aliaga D, Peeters T et al. Biosensor characterization: formal methods from the perspective of proteome fractions. *Synth Biol* 2025;**10**:ysaf002. <https://doi.org/10.1093/synbio/ysaf002>
16. Mercer JAM, DeCarlo SJ, Roy Burman SS et al. Continuous evolution of compact protein degradation tags regulated by selective molecular glues. *Science*. 2024;**383**:eadk4422. <https://doi.org/10.1126/science.adk4422>
17. Duncan O, Trösch J, Fenske R et al. Resource: mapping the *Triticum aestivum* proteome. *Plant J* 2017;**89**:601–16. <https://doi.org/10.1111/tpj.13402>
18. Nelson CJ, Millar AH. Protein turnover in plant biology. *Nat Plants* 2015;**1**:15017. <https://doi.org/10.1038/nplants.2015.17>
19. Angel TE, Aryal UK, Hengel SM et al. Mass spectrometry-based proteomics: existing capabilities and future directions. *Chem Soc Rev* 2012;**41**:3912–28. <https://doi.org/10.1039/c2cs15331a>
20. Tivendale ND, Hanson AD, Henry CS et al. Enzymes as parts in need of replacement - and how to extend their working life. *Trends Plant Sci* 2020;**25**:661–9. <https://doi.org/10.1016/j.tplants.2020.02.006>
21. Hanson AD, McCarty DR, Henry CS et al. The number of catalytic cycles in an enzyme's lifetime and why it matters to metabolic engineering. *PNAS*. 2021;**118**:e2023348118. <https://doi.org/10.1073/pnas.2023348118>
22. Oliveira-Filho ER, Gayen AK, Leong BJ et al. Continuous Directed Evolution of a Plant Histidinol Dehydrogenase to Extend Lifespan. *bioRxiv*. 2025;2025.07.01.662441. <https://doi.org/10.1101/2025.07.01.662441>
23. Lerma-Ortiz C, Jeffries JG, Cooper AJ et al. 'Nothing of chemistry disappears in biology': the top 30 damage-prone endogenous metabolites. *Biochem Soc Trans* 2016;**44**:961–71. <https://doi.org/10.1042/bst20160073>
24. Bathe U, Leong BJ, McCarty DR et al. The moderately (D)efficient enzyme: catalysis-related damage in vivo and its repair. *Biochemistry*. 2021;**60**:3555–65. <https://doi.org/10.1021/acs.biochem.1c00613>
25. Moellering RE, Cravatt BF. Functional lysine modification by an intrinsically reactive primary glycolytic metabolite. *Science*. 2013;**341**:549–53. <https://doi.org/10.1126/science.1238327>
26. García-García JD, Van Gelder K, Joshi J et al. Using continuous directed evolution to improve enzymes for plant applications. *Plant Physiol* 2022;**188**:971–83. <https://doi.org/10.1093/plphys/kiab500>
27. Mount RC, Jordan BE, Hadfield C. Transformation of lithium-treated yeast cells and the selection of auxotrophic and dominant markers. *Methods Mol Biol* 1996;**53**:139–45. <https://doi.org/10.1385/0-89603-319-8:139>
28. Gelder KV, Oliveira-Filho ER, García-García JD et al. Directed evolution of aerotolerance in sulfide-dependent thiazole synthases. *ACS Synth Biol* 2023;**12**:963–70. <https://doi.org/10.1021/acssynbio.2c00512>
29. Hanson AD, Nelsen CE, Everson EH. Evaluation of free Proline accumulation as an index of drought resistance using two contrasting barley cultivars. *Crop Sci* 1977;**17**:720–26. <https://doi.org/10.2135/cropsci1977.0011183X001700050012x>
30. Bobbitt DR. Protein sequencing protocols. Series: methods in molecular biology, Vol. 64 edited by Bryan John smith Celltech therapeutics, U.K. Humana press: Totowa, NJ. 1997. xiv + 375 pp. ISBN 0-89603-353-8. *J Am Chem Soc* 1997;**119**:8581–2. <https://doi.org/10.1021/ja9657824>
31. Lakshminarasimhan M, Madzalan P, Nan R et al. Evolution of new enzymatic function by structural modulation of cysteine reactivity in *Pseudomonas fluorescens* isocyanide hydratase. *J Biol Chem* 2010;**285**:29651–61. <https://doi.org/10.1074/jbc.M110.147934>
32. Carrió MM, Villaverde A. Protein aggregation as bacterial inclusion bodies is reversible. *FEBS Lett* 2001;**489**:29–33. [https://doi.org/10.1016/s0014-5793\(01\)02073-7](https://doi.org/10.1016/s0014-5793(01)02073-7)
33. Wilkins MR, Gasteiger E, Bairoch A et al. Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol* 1999;**112**:531–52. <https://doi.org/10.1385/1-59259-584-7:531>
34. Hughes CS, Moggridge S, Müller T et al. Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc* 2019;**14**:68–85. <https://doi.org/10.1038/s41596-018-0082-x>
35. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and

- proteome-wide protein quantification. *Nat Biotechnol* 2008;**26**:1367–72. <https://doi.org/10.1038/nbt.1511>
36. Perez-Riverol Y, Bandla C, Kundu DJ et al. The PRIDE database at 20 years: 2025 update. *Nucleic Acids Res* 2025;**53**:D543–d553. <https://doi.org/10.1093/nar/gkae1011>
37. Rix G, Williams RL, Hu VJ et al. Continuous evolution of user-defined genes at 1 million times the genomic mutation rate. *Science*. 2024;**386**:eadm9073. <https://doi.org/10.1126/science.adm9073>
38. Kitajima T, Chiba Y. Selenomethionine metabolism and its toxicity in yeast. *Biomolecular Concepts* 2013;**4**:611–6. <https://doi.org/10.1515/bmc-2013-0033>
39. Hu Y, Chai X, Men J et al. Does methionine status influence the outcome of Selenomethionine supplementation? A comparative study of metabolic and selenium levels in HepG2 cells. *Nutrients*. 2022;**14**:3705. <https://doi.org/10.3390/nu14183705>
40. Ip C. Differential effect of dietary methionine on the biopotency of selenomethionine and selenite in cancer chemoprevention. *J Natl Cancer Inst* 1988;**80**:258–62. <https://doi.org/10.1093/jnci/80.4.258>
41. Zhou Y, Chen J, Pu W et al. Development of a growth-coupled selection platform for directed evolution of heme biosynthetic enzymes in *Corynebacterium glutamicum*. *Frontiers in Bioengineering and Biotechnology* 2023;**11**:1236118. <https://doi.org/10.3389/fbioe.2023.1236118>
42. Plateau P, Saveanu C, Lestini R et al. Exposure to selenomethionine causes selenocysteine misincorporation and protein aggregation in *Saccharomyces cerevisiae*. *Sci Rep* 2017;**7**:44761. <https://doi.org/10.1038/srep44761>
43. Bogenhagen DF, Haley JD. Pulse-chase SILAC-based analyses reveal selective oversynthesis and rapid turnover of mitochondrial protein components of respiratory complexes. *J Biol Chem* 2020;**295**:2544–54. <https://doi.org/10.1074/jbc.RA119.011791>
44. Cao H, Duncan O, Millar AH. Protein turnover in the developing *Triticum aestivum* grain. *The New phytologist* 2022;**233**:1188–201. <https://doi.org/10.1111/nph.17756>
45. Sabatier P, Lechner M, Guzmán UH et al. Global analysis of protein turnover dynamics in single cells. *Cell*. 2025;**188**:2433–2450.e21. <https://doi.org/10.1016/j.cell.2025.03.002>
46. Bar-Even A, Noor E, Savir Y et al. The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry*. 2011;**50**:4402–10. <https://doi.org/10.1021/bi2002289>
47. Newton MS, Arcus VL, Gerth ML et al. Enzyme evolution: innovation is easy, optimization is complicated. *Curr Opin Struct Biol* 2018;**48**:110–6. <https://doi.org/10.1016/j.sbi.2017.11.007>
48. Sharma V, Hecker N, Roscito JG et al. A genomics approach reveals insights into the importance of gene losses for mammalian adaptations. *Nat Commun* 2018;**9**:1215. <https://doi.org/10.1038/s41467-018-03667-1>
49. Hasan T, Arora R, Bansal AK et al. Disturbed homocysteine metabolism is associated with cancer. *Exp Mol Med* 2019;**51**:1–13. <https://doi.org/10.1038/s12276-019-0216-4>
50. Usuda Y, Kurahashi O. Effects of deregulation of methionine biosynthesis on methionine excretion in *Escherichia coli*. *Appl Environ Microbiol* 2005;**71**:3228–34. <https://doi.org/10.1128/aem.71.6.3228-3234.2005>
51. Piseri A, Liu CC. The role of plasmid copy number and mutation rate in evolutionary outcomes. *Nat Ecol Evol* 2025;**9**:1694–4. <https://doi.org/10.1038/s41559-025-02792-7>
52. Javanpour AA, Liu CC. Genetic compatibility and extensibility of orthogonal replication. *ACS Synth Biol* 2019;**8**:1249–56. <https://doi.org/10.1021/acssynbio.9b00122>
53. Zhong Z, Ravikumar A, Liu CC. Tunable expression systems for orthogonal DNA replication. *ACS Synth Biol* 2018;**7**:2930–4. <https://doi.org/10.1021/acssynbio.8b00400>
54. Van Gelder K, Gayen AK, Hanson AD. Mirages in continuous directed enzyme evolution: a cautionary case study with planitized bacterial THI4 enzymes. *Plant Biotechnol J* 2025;**23**:1070–2. <https://doi.org/10.1111/pbi.14563>
55. Trivedi VD, Mohan K, Chappell TC et al. Cheating the cheater: suppressing false-positive enrichment during biosensor-guided biocatalyst engineering. *ACS Synth Biol* 2022;**11**:420–9. <https://doi.org/10.1021/acssynbio.1c00506>
56. Yunus IS, Lee TS. Applications of targeted proteomics in metabolic engineering: advances and opportunities. *Curr Opin Biotechnol* 2022;**75**:102709. <https://doi.org/10.1016/j.copbio.2022.102709>
57. Ebhardt HA, Root A, Sander C et al. Applications of targeted proteomics in systems biology and translational medicine. *Proteomics*. 2015;**15**:3193–208. <https://doi.org/10.1002/pmic.201500004>
58. Hebert AS, Richards AL, Bailey DJ et al. The one hour yeast proteome. *Mol Cell Proteomics* 2014;**13**:339–47. <https://doi.org/10.1074/mcp.M113.034769>
59. Takagi H. Metabolic regulatory mechanisms and physiological roles of functional amino acids and their applications in yeast. *Biosci Biotechnol Biochem* 2019;**83**:1449–62. <https://doi.org/10.1080/09168451.2019.1576500>