microLife, 2026, **7**, uqag005
https://doi.org/10.1093/femsml/uqag005
Published: 14 February 2026
Research Article

Proteome characterization of single gene deletion mutants lacking zinc finger μ-proteins in *Haloferax volcanii*

Pedro H.C. Franco[1,‡], Deniz Üresin[2,‡], Andreas Borst[2], Leona A. Rusling[3,4], Jörg Soppa[2,*], Julian D. Langer[1,3,*]

[1]Proteomics, Max-Planck Institute of Biophysics, 60438 Frankfurt am Main, Germany
[2]Institute of Molecular Biosciences, Goethe University, 60438 Frankfurt am Main, Germany
[3]Mass Spectrometry and Proteomics, Max-Planck Institute of Brain Research, 60438 Frankfurt am Main, Germany
[4]Institute of Pharmaceutical Chemistry, Goethe University, 60438 Frankfurt am Main, Germany
*Corresponding authors. Julian D. Langer, Max-Planck Institute of Biophysics, Max-von-Laue Straße 3, 60438, Frankfurt am Main, Germany. E-mail: julian.langer@biophys.mpg.de; Jörg Soppa, Biozentrum, Campus Riedberg, Gebäude N230, Raum 0.01b, Max-von-Laue-Str. 60438 Frankfurt am Main, Germany. E-mail: soppa@bio.uni-frankfurt.de
Editor: Kai Papenfort
‡These authors have contributed equally.

Abstract

Microproteins (≤70 amino acids) have important and often essential roles in all kingdoms of life, influencing cell motility, regulation of membrane transport and as transcription factors. In the halophilic archaeon and model system *Haloferax volcanii* a significant number of μ-proteins were predicted to be zinc finger proteins. Here we used mass spectrometry-based proteomics to systematically investigate the impact of single gene deletions of 19 zinc finger μ-proteins on the proteome of *H. volcanii* grown in synthetic medium with glucose as sole carbon and energy source. We employed a state-of-the-art dia-PASEF acquisition strategy, detecting over 3400 proteins across the 19 deletion strains and the wild type. The comprehensive proteome coverage enabled a systematic analysis of proteome remodeling. We found that in 11 out of the 19 mutants the proteome remodeling involved proteins annotated to play a role in cell motility, matching swarming and growth rate phenotypes we observed for these strains. Taken together, our data provide the most comprehensive proteome coverage of *H. volcanii* to date, and the effect of 19 different zinc-finger μ-proteins deletion strains on the proteome of this organism. The combined data (available via ProteomeXchange with identifier PXD066008) provide a valuable resource for future research in the field.

Keywords *Haloferax volcanii*, proteomics, dia-PASEF, small proteins, cell motility, zinc-finger proteins

Introduction

Proteins of less than 70 amino acids (aa), hereby defined as μ-proteins, comprise an understudied group of proteins found in all domains of life (Storz et al. 2014, Kushwaha et al. 2022, Burton et al. 2024). Their small size makes them difficult to be annotated during genomic sequencing and to be studied by conventional analytical workflows (Ahrens et al. 2022). Nevertheless, μ-proteins are now known to play key roles in many organisms, including cell motility (Hoffman et al. 2022), antibiotic resistance (Hobbs et al. 2012), and regulation of membrane transporters (Wright et al. 2024).

Using transcriptomics and proteomics, recent studies have expanded the annotation of μ-proteins of multiple organisms, from model prokaryotes to humans. Given its high sensitivity and resolution, Ribo-Seq is routinely employed in the identification of small open-reading-frames (sORFs), e.g. in *Escherichia coli* (Hemm et al. 2020), *Staphylococcus aureus* (Fuchs et al. 2021), SARS-CoV-2 (Finkel et al. 2021), and humans (Sandmann et al. 2023). Complementarily, MS-based proteomics has been widely used for obtaining evidence on the protein level of μ-proteins in these organisms and others (Cassidy et al. 2016, Sandmann et al. 2023, Meier-Credo et al. 2023, Franco et al. 2025).

In *Haloferax volcanii*, a model organism for Haloarchaea [whose biology and applications are extensively reviewed in (Pohlschroder et al. 2025)], more than 400 μ-proteins have been annotated and multiple candidates have been characterized (Hadjeras et al. 2023, Üresin et al. 2024). In a previous work, we observed that a higher percentage of μ-proteins in *H. volcanii* contained dual C(P)XCG motifs—a family of predicted zinc finger proteins—in comparison to the global proteome (Üresin et al. 2024). Additionally, we have experimentally shown that the majority of these μ-proteins annotated as "zinc finger proteins" do actually bind zinc, while others might be zinc free or complexed to other metals (Üresin et al. 2024).

Zinc finger proteins are capable of interacting with a wide range of biomolecules, from nucleic acids to soluble and membrane proteins (Matthews and Sunde 2002, Eom et al. 2016, Üresin et

Received: 19 July 2025. **Revised:** 3 February 2026. **Accepted:** 8 February 2026

al. 2023). In addition, they play important roles in human cells, which include transcription regulation, apoptosis, and autophagy (Li et al. 2022). *Haloferax volcanii* mutants lacking zinc finger μ-proteins have shown distinct phenotypic differences from the wild type, including growth adaptation, biofilm formation and swarming (Nagel et al. 2019). The fact that nearly all mutants exhibited at least one phenotype and the variance of phenotypes indicated that these proteins play important roles also in various important haloarchaeal biological processes. It should be noted that the widely-used "swarm plate assay" applied in that study de facto quantifies swimming and chemotaxis in plates with very low agar concentration, while bona fide swarming (e.g. by Myxobacteria) occurs only on solid support and is another type of movement.

In this work, we studied, using mass spectrometry (MS)-based proteomics, the proteome remodeling caused by the single gene deletion of 19 zinc finger μ-proteins in *H. volcanii* cells grown in synthetic medium with glucose as sole carbon and energy source media (Fig. 1). For the first time for this organism, we acquired MS data using the dia-PASEF workflow on a timsTOF instrument (Meier et al. 2020). This approach utilizes ion-mobility separation and ion accumulation, increasing sensitivity while reducing spectral complexity. dia-PASEF thus enables acquisition of comprehensive datasets in short measurement times.

Furthermore, we used the generated differential expression data to better understand the phenotypes we observed in phenotypical analysis for a large number of these strains.

The acquisition of large-scale datasets for *H. volcanii* and deriving mutant strains, paired with a phenotypical understanding of each mutant, is beneficial for community efforts such as the Archaeal Proteome Project (ArcPP) (Schulze et al. 2020, Schulze et al. 2021). The project compiles MS datasets from different studies which can be reviewed and reanalyzed to generate important insight into *H. volcanii* metabolism.

Materials and methods

Strains, media, and culture conditions

All deletion mutants were derived from the *H. volcanii* strain H26 (Allers et al. 2004), which was used as wild type in this study. Cell extracts were isolated from *H. volcanii* wild type and mutant cultures grown in synthetic glucose medium (Dambeck and Soppa 2008) with a mixture of trace elements added as previously described (Üresin et al. 2024) to the mid-exponential growth phase. The medium contained 2.14 M NaCl, 135 mM KCl, 220 mM $MgCl_2$, 40 mM $MgSO_4$, 9 mM $CaCl_2$, 1 mM K_2HPO_4, 10 mM NH_4Cl, 0.1% BME Vitamins 100x solution, 25 mM Glucose, 50 μg/ml Uracil, 100 mM MOPS pH 7.2, 50 mM Tris/HCl pH7.2, 3.5 μM $ZnSO_4$, 1.5 μM $MnCl_2$, 50 μM H_3BO_3, 8.5 μM $CoCl_2$, 0.5 μM $CuCl_2$, 1 μM $NiCl_2$, 1 μM Na_2MoO_4, and 8 μM $FeSO_4$.

Generation of in frame deletion mutants

The Pop-In-Pop-Out method (Allers et al. 2004) was used for the generation of in-frame deletion mutants. Each strain harbors the deletion of a single zinc-finger μ-protein gene: *HVO_0546, HVO_0695, HVO_0767, HVO_0885, HVO_1118, HVO_1352, HVO_1359, HVO_1533, HVO_1677, HVO_2142, HVO_2400, HVO_2523, HVO_2805A, HVO_2901, HVO_2982, HVO_2983A, HVO_A0089, HVO_A0254A*, and *HVO_A0457A*. Throughout the text, each deletion strain will be referred to without the "HVO_" tag, e.g. Δ*0546*.

For each deletion mutant, two overlapping PCR fragments (∼500 bp) were generated, with the first containing the upstream region and the first codons of the gene and the second containing the last codons and the downstream region. By PCR, both fragments were fused and the resulting fragment was inserted into the pMH101 plasmid (Hammelmann and Soppa 2008). Sequences for all oligonucleotides used in this study can be found in Supporting Table 1. Plasmid sequences were verified via sequencing. After transformation of *H. volcanii* wild type (H26), the first selection step was growing the strains in medium that does not contain uracil, as H26 lacks the *pyrE2* gene which was included in the plasmids. Plasmid-containing clones ("Pop-In clones") were then grown in medium containing uracil and 5-FOA, selecting for Pop-Out variants. Pop-In- and Pop-Out clones were identified by colony PCR. Since *H. volcanii* is polyploid (Breuert et al. 2006), Pop-Out clones were frequently checked again for any wild type alleles via 40-cycle PCR during the course of this study.

Phenotypic characterization of mutants

With the exception of Δ *0885* and Δ*A0457A*, all mutants have been phenotypically characterized (Üresin et al., in preparation, Nagel et al. 2019). In short, the following phenotypic traits have been characterized: growth in various media, cell and colony morphology, biofilm formation, and the velocity of movement on swarm plates. It should be noted that the widely-used swarm plates in fact measure swimming/chemotaxis on plates with very low agar concentration, while bona fide swarming occurs only on solid support.

Whole cell extraction

Cell extracts were isolated from *H. volcanii* wild type and mutant cultures grown as described above to the mid-exponential growth phase (4–5 × 10^8 cells/ml). Three different isolation protocols were tested. The first one was based on the methanol-chloroform extraction (MCE) method after Wessel & Flügge (Wessel 1984), the second, based on the acid guanidinium thiocyanate-phenol-chloroform extraction (AGPC) method adapted from Chomczynski & Sacchi (Chomczynski 1987) and the last one involved cell lysis with SDS.

For the MCE method, 2 ml per culture were pelleted by centrifugation (2 min, 8000 rpm). The supernatant was removed and the pellet resuspended in 0.2 ml lysis buffer (10 μl DNase (1 mg ml^{-1}), 10 μl Protease inhibitor mix, and 1 ml aqua bidest). After incubation at room temperature for 30 min, 0.8 ml methanol, 0.2 ml chloroform, and 0.6 ml aqua bidest were added and the solution was mixed after every addition. A phase separation was achieved via centrifugation at 4000 rpm for 15 min at room temperature. The upper phase was removed, then 0.6 ml methanol were added, mixed and centrifuged again at 4000 rpm for 5 min at room temperature. The supernatant was removed and the pellet dried at room temperature. The protein extract was then solubilized in 200 μl of solution buffer (25 mM Tris/HCl pH 7.2) at 65°C.

For the AGPC method, 2 ml per culture were pelleted by centrifugation (2 min, 8000 rpm). The supernatant was removed and the pellet was weighed. 1 ml of TRIzol reagent was added per

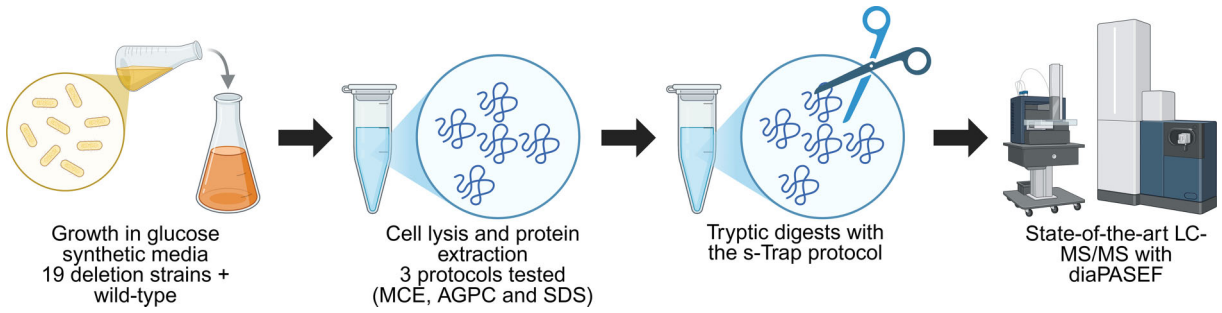

Figure 1 Optimized mass spectrometry-based proteomics workflow for the study of *H. volcanii*. Cells grown in glucose synthetic media were lysed following three different protocols: methanol-chloroform extraction (MCE), acid guanidinium thiocyanate-phenol-chloroform extraction (AGPC), and SDS-based cell lysis. The most efficient protein extract, based on the protein yield, was digested following the s-Trap protocol, and peptide samples were analyzed in timsTOF instruments using dia-PASEF. Figure was generated using BioRender.com

100 mg of pellet. Cell lysis was performed by incubation at 65°C for 15–20 min. Afterwards the incubation was continued for 5 min at room temperature before adding 0.2 ml chloroform per ml of TRIzol reagent and vortexing. After another 5 min incubation step at room temperature the phase separation was performed via centrifugation for 15 min at 4°C. The upper phase was removed and 0.3 ml ethanol per ml of TRIzol reagent was added. DNA was pelleted by centrifugation (2000x g, 5 min, 4°C) and the supernatant was transferred to a new reaction tube. 1.5 ml isopropanol were added per ml of TRIzol reagent to precipitate the proteins. After 15 min of incubation at room temperature, a 15 min centrifugation step at 4°C was performed. The supernatant was removed and 2 ml of 300 mM guanidinium hydrochloride in 95% ethanol were added and the pellet was shredded mechanically. After 20 min of incubation at room temperature and a 15 min centrifugation step at 4°C, the supernatant was removed. To remove any extant phenol, the last steps, beginning from the addition of guanidinium hydrochloride, were repeated 4–6 times. The phenol free protein pellet was finally washed twice with 2 ml of ice-cold acetone and then dried. Finally, it was solved in 200 µl of solution buffer (see above).

For cell lysis with SDS 0.5 ml per culture were transferred to an Amicon Ultra-0.5 Centrifugal Filter Unit (Merck KGaA, Darmstadt, Germany) with a cut-off of 30 kDa and then concentrated to around 100 µl via centrifugation (14 000 rpm at room temperature). 400 µl of lysis solution (aqua bidest with a protease inhibitor mix and DNase) were added and the solution was mixed for lysis of the halophilic cells via osmotic shock. The centrifugation step to around 100 µl followed by the addition of 400 µl of lysis solution was repeated twice. After the third repetition, 400 µl of 2% (w/v) SDS in 100 mM Tris/HCl pH 7.5 were added and the solution was mixed and then incubated for 15 min at 55°C and concentrated to 200 µl via centrifugation. The extraction protocol was repeated four times to obtain identical protein amounts for comparison to the previous two protocols. The pooled extract was then again concentrated to 200 µl via centrifugation.

The efficiencies of the three methods were estimated by loading identical sample volumes (12.5 µl of each extract, mixed with 4 µl loading buffer) on a gel, separation by SDS-PAGE, and staining the gel with Coomassie brilliant blue. Quality criteria were the amount of isolated protein, the number and intensity of different bands, the size distribution of the isolated proteins, and the reproducibility of the method. In addition, the amounts of protein extracted by each protocol were estimated by UV-280 absorption.

Databases and software

Annotations and gene and protein sequences were retrieved from the HaloLex database (Pfeiffer et al. 2008). The HaloLex *H. volcanii* FASTA file version 230 328 from March 28th, 2023 was used for all proteomics analysis. Oligonucleotides for cloning were designed with the CloneManager 8.0 software (Sci Ed, Colorado, USA).

Sample preparation for mass spectrometry

The *H. volcanii* MCE-protein extracts were digested for MS analysis following the s-Trap protocol (HaileMariam et al. 2018) with minor adaptations and purified using solid-phase extraction. Briefly, a volume of cell lysate containing 100 µg of proteins was diluted to a final volume of 100 µl with 50 mM ammonium bicarbonate (ABC) and proteins were reduced by addition of 4 µl 0.5 M tris(2-carboxyethyl) phosphine (TCEP) for 30 min at 37°C with mild agitation. Cysteine alkylation was performed with 8 µl 0.5 M iodoacetamide (IAA) for 30 min in the dark with mild agitation. The samples were then acidified with 12 µl 12% phosphoric acid, diluted with 750 µl of the S-Trap binding buffer [1 M tetraethylammonium bicarbonate (TEAB) Buffer and Methanol (10:90)] and transferred to S-Trap mini columns (Protifi). SDS removal was conducted by three washing cycles with 400 µl S-Trap binding buffer prior to digestion. MS-grade Trypsin (Serva) was added in an enzyme to protein ratio 1:50 using 125 µl digestion buffer (50 mM ABC, pH 8.5) as media and the column incubated overnight at 37°C with light agitation. Peptides were eluted stepwise in 80 µl of 50 mM ABC, 0.2% formic acid, and 50% acetonitrile (ACN). The pooled fractions were diluted 1:1 with 0,1% trifluoroacetic acid (TFA) and desalted with C18-SPE cartridges (Biotage). After equilibration with 2 ml ACN, 1 ml 50% ACN/1% acetic acid and 2 ml 0.1% TFA the samples were loaded onto the cartridge, washed with 2 ml 0.1% TFA and eluted with 1 ml 80% ACN/0.1% TFA. The eluted fractions were dried using an Eppendorf concentrator (Eppendorf) and stored at -20°C.

Liquid chromatography and mass spectrometry

Initial replicates of each deletion strain and two wild type replicates were reconstituted in 5% ACN with 0.1% formic acid (FA) and separated on a ReproSil C18-PepSep analytical column (particle

size = 1,9 μm, ID = 75 μm, L = 15 cm, Bruker Corporation, Billerica, USA) using a nano-HPLC (nanoElute, Bruker Corporation) at a temperature of 60°C. Peptides were separated by a gradient of water (buffer A: 100% H_2O and 0.1% FA) and acetonitrile (buffer B: 100% ACN, 0.1% FA) with a constant flow rate of 500 nL/min. The gradient went from 2% to 35% buffer B in 17.8 min. All solvents were LC-MS grade and purchased from Riedel-de Häen/Honeywell (Seelze, Germany). Eluting peptides were analyzed on a hybrid trapped ion mobility spectrometer (TIMS)—quadrupole time of flight (Q-ToF) mass spectrometer timsTOF Pro II (Bruker Daltonics) coupled to the nano-HPLC by a CaptiveSpray nano-electrospray ion source. Data was acquired using data independent acquisition utilizing Parallel Accumulation Serial Fragmentation (dia-PASEF). MS scans were carried out in positive mode from 100–1700 Th and ion mobility window ranged from 0.85 to 1.3 V.s/cm^2, with a ramp time of 100 ms. For dia-PASEF, 21 MS/MS windows with a width of 25 Da were set from 475 to 1000 Da and 0.85 V.s/cm^2 to 1.27 V.s/cm^2. Estimated cycle time was of 0.95 s. Stepped collision energy was set linearly from 20 eV at 0.6 V.s/cm^2 to 59 at 1.6 V.s/cm^2.

The second batch of biological replicates and three additional wild type replicates were reconstituted and peptides were separated as described above with the following changes to account for a next-generation instrument and boost identification rates. Peptides were separated by a gradient of water (buffer A: 100% H_2O and 0.1% FA) and acetonitrile (buffer B: 100% ACN, 0.1% FA) with a constant flow rate of 800 nL/min. The gradient went from 2% to 38% buffer B in 21 min. Eluting peptides were analyzed on a hybrid trapped ion mobility spectrometer (TIMS)—quadrupole time of flight (Q-ToF) mass spectrometer timsTOF HT (Bruker Daltonics) coupled to the nano-HPLC by a CaptiveSpray nano-electrospray ion source. Data was acquired using data independent acquisition utilizing Parallel Accumulation Serial Fragmentation (dia-PASEF). MS scans were carried out in positive mode from 100–1700 Th and ion mobility window ranged from 0.85 to 1.3 V.s/cm^2, with a ramp time of 100 ms. For dia-PASEF, 21 MS/MS windows with a width of 25 Da were set from 475 to 1000 Da and 0.85 V.s/cm^2 to 1.27 V.s/cm^2. Estimated cycle time was of 0.95 s. Stepped collision energy was set linearly from 20 eV at 0.6 V.s/cm^2 to 59 at 1.6 V.s/cm^2.

Mass spectrometric data analysis

DIA raw files were processed with the open-source software DIA-NN (Demichev et al. 2020, 2022) (version 1.9) using a library-free approach. The predicted library was generated by in silico digesting the *H. volcanii* proteome with Trypsin/P. Deep learning-based spectra- and RT-prediction was enabled. Peptide length range was set to 7–35 amino acids, missed cleavages to 2 and precursor charges to 1–5. Methionine oxidation and *N*-terminus acetylation were set as variable modifications, and cysteine carbamidomethylation as a fixed modification. The maximum number of variable modifications per peptide was limited to 3. MS1 and MS2 accuracies were set to 15.0 ppm, isotopologues and match-between-runs were enabled, while shared spectra were disabled. Protein inference was performed using genes with the heuristic protein inference option enabled. The neural network classifier was set to single-pass mode and the quantification strategy was selected as "QuantUMS (high precision)." Normalization was disabled and only employed in downstream processing. After the database search, the output from DIA-NN was imported to R using the diann-rpackage. Data was filtered using a 1% FDR on the protein and peptide level, and maintaining only proteins with at least one proteotypic peptide, as recommended by the original publication (Demichev et al. 2020). For stringent data analysis and comparison to the ArcPP repository, raw data were reprocessed using the diann-rpackage using a 0.5% FDR, 1% PEP and a minimum of 2 PSMs as filters for protein identification, as recommended by their original work (Schulze et al. 2020).

Differential expression analysis and data processing

The database-search output from DIA-NN was analyzed using the MS-DAP R package (Koopmans et al. 2023) for differential expression analysis with the following parameters. Only proteins present in both replicates with a minimum of one unique peptide per protein were used. These filters were applied for each comparison individually. For each comparison, differential expression analysis was carried out using the MSqRob algorithm (Goeminne et al. 2020), followed by Benjamini–Hochberg correction for multiple-testing. A q-value (corrected *P*-value) of 5% was set as a threshold for the analysis. Normalization was performed using the vsn followed by modebetween_protein algorithms at MS-DAP, as recommended in the original publication (Koopmans et al. 2023). A $\log_2$-fold change threshold of 0 was set. The results from the differential expression analysis were exported to RStudio for follow-up analysis. Pathway enrichment analysis based on arCOG annotations and GO terms was done in R using the clusterProfiler package (Yu et al. 2012, Wu et al. 2021, Xu et al. 2024). Significance p-values were corrected using the Benjamini–Hochberg method. Clustering analysis was done using the pheatmap package using Euclidean distances and ward.D2 clustering method. Figures were generated using the ggplot2 package with data imported using the diann-rpackage (Demichev et al. 2020).

Results

Complete protein extraction

Several methods are reported in literature for protein extraction from *H. volcanii* (Kirkland et al. 2006, Kaminski et al. 2012, Schiller et al. 2022, 2024). We investigated which method would yield for our strains and experimental conditions the largest amount of protein extracted. We compared protein band intensities and distribution on SDS-PAGE for equal extract volumes obtained with MCE (Wessel 1984), acid guanidium thiocyanate–phenol–chloroform extraction (AGCP) (Chomczynski 1987) and cell lysis with SDS (see Methods). Cell extracts were prepared from cultures in mid-exponential growth phase (4×10^8 cells/ml), and volumes equivalent to 5×10^7 cells were loaded on a gel separated by SDS-PAGE (Supporting Fig. 1), using three biological replicates for each method. The number, size distribution and reproducibility of the protein bands on the SDS-PAGE indicated that the MCE method extracted a larger and more comprehensive proteome in comparison to the other two methods. This was also confirmed by UV-280 absorption measurement, in which the MCE protocol yielded higher protein amounts (Supporting Fig. 1). Thus, the MCE method was employed for all subsequent experiments in this study.

Experimental design

We then grew cells in synthetic medium with glucose as sole carbon and energy source, as more metabolic pathways are active in synthetic media in comparison to complex media (Portnoy et al. 2011). This growth condition also enables the linking of proteomics results to the results from previous experiments (Nagel et al. 2019, Üresin et al. 2023, 2024). Cells were grown to mid-exponential growth phase to guarantee steady-state levels of all proteins, ensuring high reproducibility (Clark 2021).

We investigated how reproducible and similar *H. volcanii* proteomes are expressed in different experiments under these identical growth conditions, to assess how many independent biological replicates would be needed to obtain robust data for the 19 deletion strains. To this end, we grew wild type *H. volcanii* in five independent experiments in identical conditions. We then conducted a statistical analysis, pairing the data in two "mock" groups, each comprised of different wild type samples, and comparing them based on the proteome profiles and the expression levels of each protein. To do so, we employed our downstream differential expression analysis processing pipeline (see Methods section). We observed only minimal differences between the two mock groups (Supporting Fig. 2). We thus decided to use biological duplicates for screening differentially expressed proteins in the zinc finger μ-proteins lacking strains.

Largest proteome detected for *H. volcanii*

We employed a dia-PASEF workflow optimized for sensitive and robust protein quantitation, using gradients of 25 min per sample. Our workflow led to the largest proteome of *H. volcanii* detected to date, with a total of 3476 proteins detected across all samples accounting for 82% of the entire annotated proteome of this organism (Supporting Table 2, Fig. 2A-C). From these, 27% are membrane proteins. Given that the fraction of membrane proteins in our dataset is not far from the fraction of membrane proteins in the annotated proteome (20% vs. 24%), we conclude that protein solubilisation is efficient in our workflow. As expected, however, comparing sequence coverage of membrane and cytosolic proteins we observed significantly lower coverage for predicted membrane proteins than soluble proteins (23.91% vs. 57.42%).

To enable a direct comparison of our dataset to the ArcPP database, we reprocessed our data following their stringent recommendations ($\leq$ 0.5% FDR, $\leq$ 1% PEP, and $\geq$ 2 PSMs per protein). Following these standards, we confidently identify 3072 proteins, 223 of them for the first time. From these 3072 proteins, 76 are μ-proteins, from which 20 were not previously reported in ArcPP.

Also comparing to the ArcPP database, we do not detect 993 annotated proteins. From these, only 182 were previously identified by other studies while 811 were also missed by all other studies listed in ArcPP. The *H. volcanii* proteome contains large groups of proteins that are expressed under specific conditions only, such as viral defense, stationary phase, biofilm formation-related, and stress response proteins (Dantuluri et al. 2016, McMillan et al. 2018, Jevtić et al. 2019, Schulze et al. 2020, 2021). Our dataset was acquired in mid-exponential phase conditions, so we expected to miss proteins specific for other conditions.

Notably, of the 993 undetected proteins in this study, approximately one third (323 proteins) are known or predicted to be membrane proteins. Furthermore, 520 (∼52%) of them are described as hypothetical proteins, a higher proportion compared to 1047 (∼25.8%) in the total ArcPP reported proteome.

Analysis of the mutants

Due to the high data completeness and reproducibility enabled by our dia-PASEF workflow, most of the detected proteins could be used for differential expression analysis. At 1% FDR, we detected on average ∼2600 proteins in each mutant, with ∼2300 passing the filters for quantitation: detectable in both replicates with at least one unique peptide (Supporting Table 3). For each pairwise comparison, we evaluated statistical significance using the MSqRob algorithm, finding a total of 137 up- or down-regulated proteins in at least one deletion strain (Supporting Fig. 3).

In addition, we analyzed if we could detect proteins that were highly abundant in certain deletion strains and completely absent in the wild type (thus precluding statistical tests based on intensities in both samples), as well as proteins highly abundant in the wild type but absent in the deletion strains. To do so, we initially looked for proteins detected in mutant strains with at least three unique peptides, but undetected in the wild type. Given the high data completeness of the wild type, we only found a small number of proteins that could be detected in the deletion strains but not in the wild types (Supporting Table 4). While some proteins are commonly found across multiple deletion strains, e.g. HVO_0423 and CoxC1 (HVO_1138), no clear association to the phenotypes could be made. Reciprocally, we also looked for proteins detected with at least three unique peptides in the wild type but absent in the deletion strains. In each deletion strain, 30–114 proteins were missing in comparison to the wild type. Intensity-wise, these missing values are distributed across different orders of magnitude (Supporting Fig. 4), indicating that while some might be absent simply due to low abundance and experimental variability, others might be linked to physiological changes.

Our systematic analysis also enabled the clustering of different strains lacking zinc-finger μ-proteins based on their proteome remodelling (Fig. 2D). By comparing the expression of 1800 proteins consistently identified in all twenty strains (wild type included), we clustered them in five different groups. From these groups, three have distinct features. Δ2983A is clustered alone, and showed differential expression for cytosolic proteins involved in the metabolism of small molecules. For Δ*2142*, Δ*2523*, and Δ*2901* we see a differential abundance in proteins involved in glycosylation and in the S-layer structure. These include changes in archaeal glycosylation components and in RdfA (HVO_2174) and Sph3 (HVO_2175), which are related to rod/disk structuring (Schiller et al. 2024).

To systematically analyze the physiological effects of the μ-protein deletions on *H. volcanii*, we first conducted multiple gene onthology analyses based on arCOG categories and GO-terms. For 16 out of the 19 mutants, we observed enrichments in arCOG categories (Supporting Table 5). We have also detected differentially enriched GO-terms in all of the mutant strains (Supporting Table 6), with a recognizable overlap between GO-terms, arCOG categories and observed phenotypes (Table 1).

Seven mutants (Δ*0546*, Δ*0767*, Δ*0885*, Δ*2142*, Δ*2523*, Δ*2901*, and Δ*A0457A*) show a depletion of proteins involved in cell motility, while four mutants (Δ*1352*, Δ*1533*, Δ*2400*, and Δ*A0254A*) exhibit enriched proteins for this category based on arCOG terms. Furthermore, we also identified motility-related GO-terms differ-

Table 1 Differentially regulated GO and arCOG pathways for the studied strains.

Deletion strain	Observed phenotypes*	Differential arCOG categories	Differential motility-linked GO terms
Δ*0546*	Swarming deficit No rod-shaped cells	Cell motility (-)	Chemotaxis (-) Archaeal-type flagellum (-)
Δ*0695*	Swarming deficit No rod-shaped cells		
Δ*0767*	Swarming deficit No rod-shaped cells Growth deficit in low salt Growth deficit in glycerol	Cell motility (-), carbohydrate transport and metabolism (+)	Chemotaxis (-) Archaeal-type flagellum (-) archaeal or bacterial-type flagellum-dependent cell motility (-)
Δ*0885*	Swarming deficit**	Cell motility (-), inorganic ion transport and metabolism (+)	Chemotaxis (-) Archaeal-type flagellum (-) archaeal or bacterial-type flagellum-dependent cell motility (-)
Δ*1118*		Cell wall/membrane/envelope biogenesis, PTMs, protein turnover, chaperones (-)	
Δ*1352*	Swarming improvement More rod-shaped cells	Cell motility (+)	Chemotaxis (+)
Δ*1359*	Growth deficit in glycerol	Amino acid transport and metabolism, signal transduction mechanisms (+)	Chemotaxis (+)
Δ*1533*	Swarming improvement More rod-shaped cells	Cell motility (+)	Chemotaxis (+)
Δ*1677*			
Δ*2142*	Growth deficit in glycerol Swarming deficit Increased biofilm formation	Cell motility (-)	Chemotaxis (-) Archaeal-type flagellum (-) archaeal or bacterial-type flagellum-dependent cell motility (-)
Δ*2400*	Growth improvement in glycerol	Cell motility (+)	Chemotaxis (+)
Δ*2523*	Growth deficit in glycerol Swarming deficit Increased biofilm formation	Cell motility, defense mechanisms (-), carbohydrate transport and metabolism (+)	Chemotaxis (-) Archaeal-type flagellum (-) archaeal or bacterial-type flagellum-dependent cell motility (-)
Δ*2805A*		Cell wall/membrane/envelope biogenesis, defense mechanisms (-)	
Δ*2901*	Swarming deficit Increased biofilm formation	Cell motility (-), cell wall/membrane/envelope biogenesis (+)	Chemotaxis (-) Archaeal-type flagellum (-) archaeal or bacterial-type flagellum-dependent cell motility (-)
Δ*2982*	Swarming deficit More rod-shaped cells		Chemotaxis (-)
Δ*2983A*	Swarming improvement Reduced biofilm formation Growth deficit in low salt Growth improvement in glycerol	Secondary metabolites biosynthesis, transport and catabolism (+)	
Δ*A0089*	Growth deficit in glycerol	Cell wall/membrane/envelope biogenesis, PTMs, protein turnover, chaperones (-)	Chemotaxis (+)
Δ*A0254A*		Cell motility (+)	Chemotaxis (+)
Δ*A0457A*	***	Cell motility, signal transduction mechanisms (-)	

(-) for depletion and (+) for enrichment is given based on arCOG classes from ArcPP and on UniProt GO terms. *data from Üresin et al., in preparation, Nagel et al. 2019. ** Preliminary data (not available) *** Unstudied phenotype.

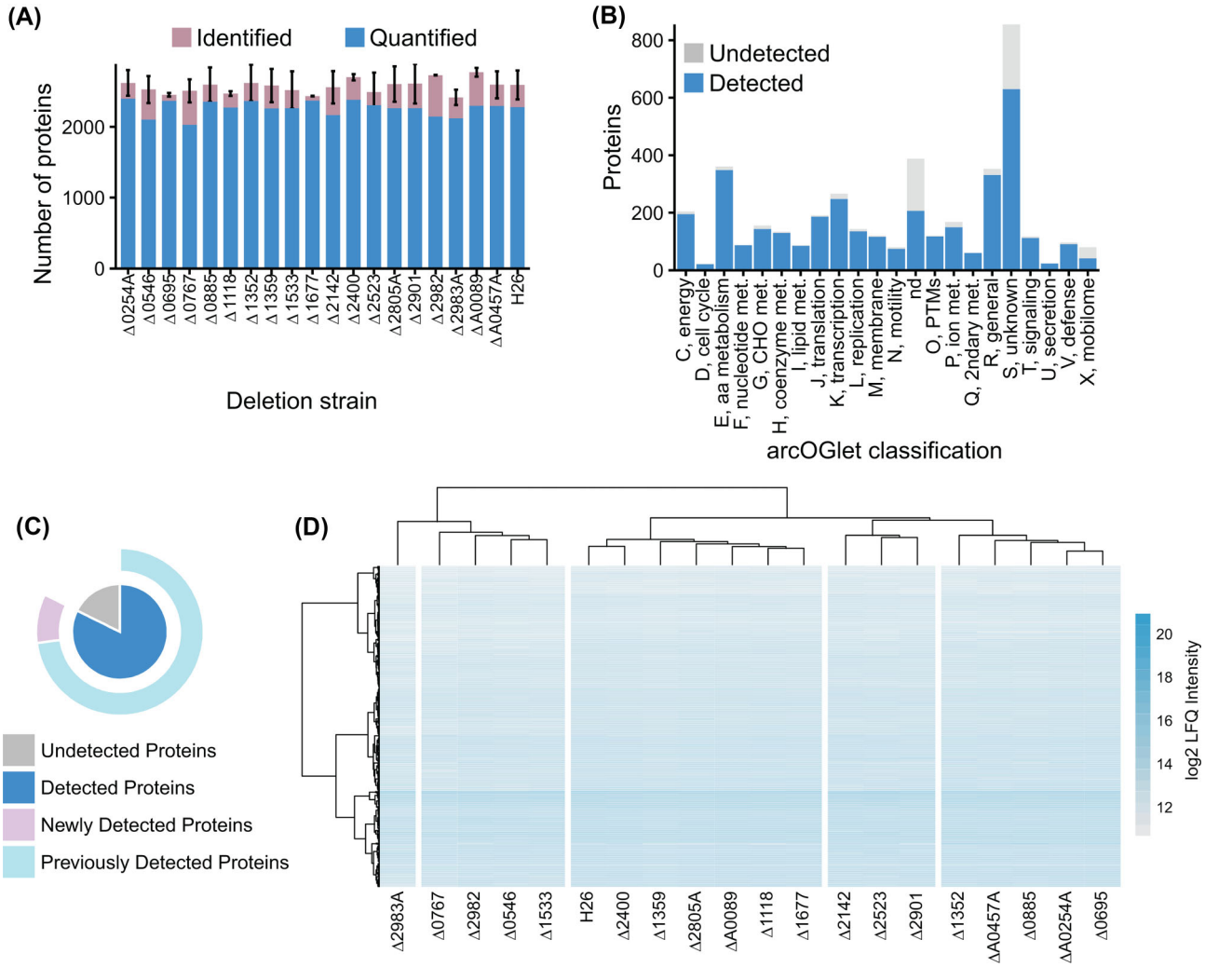

Figure 2 Detection of the largest MS-based proteome of *H. volcanii*. (A) Number of identified proteins in each deletion strain ($n = 2$) and in the wild type ($n = 5$). Bars are represented as mean ± sd. (B) Coverage of the detected proteome in this study is almost complete for all arCOG annotations, with the exception of unclassified proteins and proteins with unknown function. Annotations are J, translation, ribosomal structure and biogenesis; U, intracellular trafficking, secretion, and vesicular transport; F, nucleotide transport and metabolism; H, coenzyme transport and metabolism; O, post-translational modification, protein turnover, chaperones; Q, secondary metabolites biosynthesis, transport and catabolism; I, lipid transport and metabolism; V, defense mechanisms; E, amino acid transport and metabolism; C, energy production and conversion; L, replication, recombination and repair; T, signal transduction mechanisms; R, general function prediction only; D, cell cycle control, cell division, chromosome partitioning; P, inorganic ion transport and metabolism; K, transcription; M, cell wall/membrane/envelope biogenesis; G, carbohydrate transport and metabolism; S, function unknown; N, cell motility; X, mobilome. (C) Representation of the proteome coverage based on the number of proteins detected in our experiments compared to the proteins listed in the FASTA database, as well as the number of proteins we detected for the first time, in contrast to proteins that were previously reported in the ArcPP (D) Clustering analysis of the nineteen studied strains and the wild type based on proteins that were identified across all samples.

entially regulated in three additional strains: an up-regulation in *Δ1359* and *ΔA0089*, and down-regulation in *Δ2982*.

For 17 of these strains, we were able to compare the proteomics data to phenotypical findings. The phenotypes for seven of these strains were published in Üresin et al., (in preparation), and for 10 of them in Nagel et al., (2019). *Δ0885* and *ΔA0457A* are the only ones with no published phenotype to this date.

Our proteomics findings are consistent to the phenotypical observations for most of the strains (Fig. 3A and B). *Δ0546*, *Δ0695*, *Δ0767*, *Δ2142*, *Δ2523*, *Δ2901*, and *Δ2982* presented deficit in their swarming capabilities. Unpublished preliminary data also indicates that *Δ0885* might have swarming deficit. In the opposite direction, *Δ1352*, *Δ1533*, and *Δ2983A* showed increased motility phenotypes (Üresin et al., in preparation, planned for back-to-back publications). In addition, we observe that the strain *Δ2901* shows the highest difference in the expression of several motility-related proteins in comparison to the wild type (Fig. 3C). We observe for this mutant a significant depletion in CheR (HVO_1222), CheW (HVO_1225), ArlA1 (also referred to as FlgA1, HVO_1210), and ArlA2 (also referred to as FlgA2, HVO_1211), ranging from 2.9 to 7.8 $\log_2$-fold differences.

We found the archaellin proteins ArlA1 (FlgA1, HVO_1210) and ArlA2 (FlgA2, HVO_1211) to be a major component affected in the strains with motility defects. In six out of the eight strains

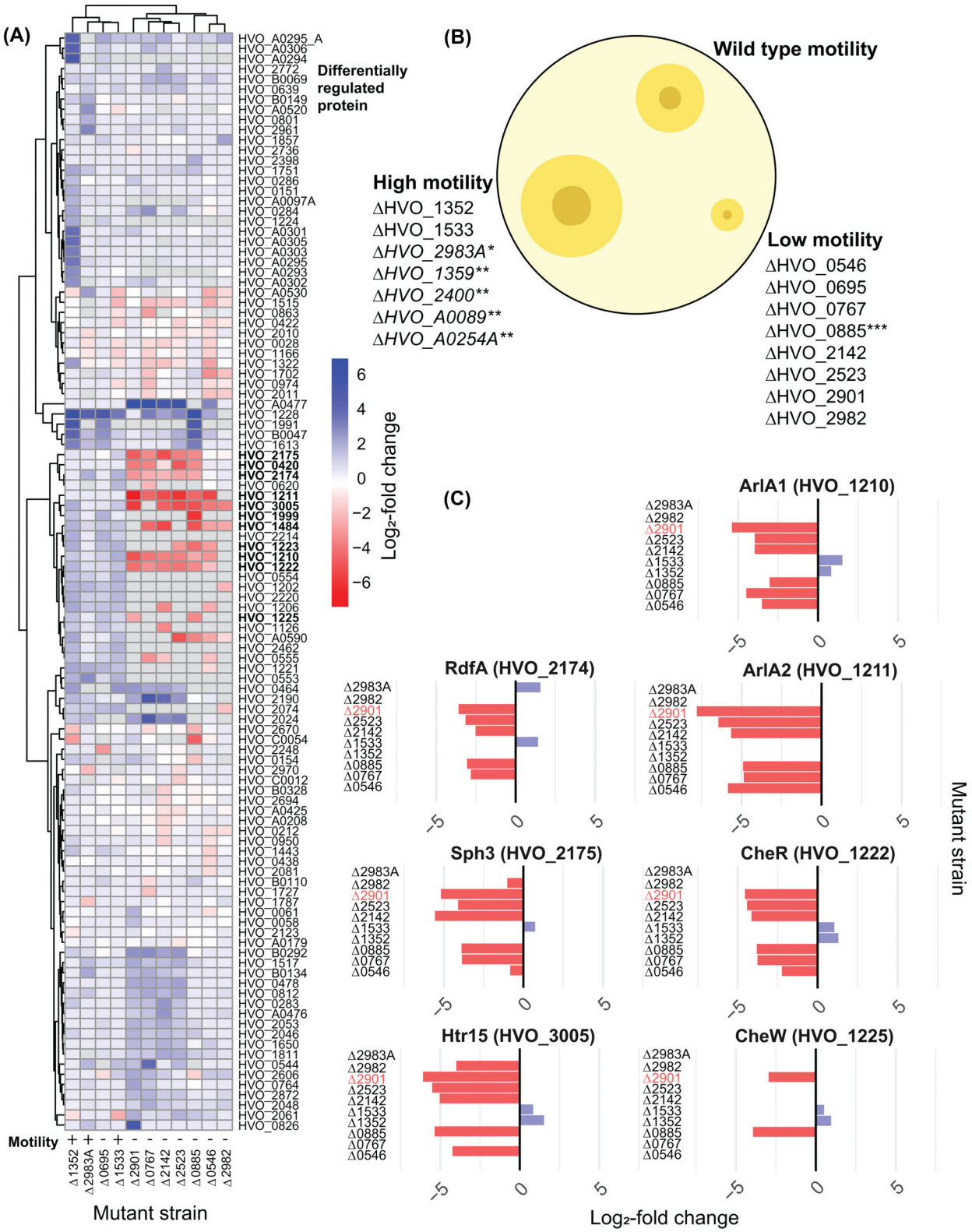

Figure 3 Proteomics data show differential abundance of proteins related to chemotaxis and flagellum structure, consistent with observed phenotypes. (A) Clustering analysis of proteins differentially regulated in at least one of the strains for which we observed phenotypical changes in cell motility. Color scale reflects the Log_2-fold change in comparison to the wild type. Grey fields refer are proteins not quantified in the respective sample. (B) Illustrative summary of the observed motility phenotypes based on cell biology assays (extracted from Üresin et al. (in preparation) and Nagel et al. 2019) and on the proteomics data. ∗strains which present the phenotype only on cell biology studies. ∗∗strains which show no phenotype but have enriched or depleted cell motility proteins. ∗∗∗ strain with unpublished preliminary phenotypical data. (C) Bar plots highlighting the fold-changes of proteins linked to cell motility in the strains with observed phenotypes. The Δ2901 strain was highlighted in red as it shows the highest overall fold changes in motility-related proteins.

with lower motility (including Δ*0885* due to the preliminary phenotype and Δ*0695* and Δ*2982* being the exceptions), ArlA1 and ArlA2 are consistently found depleted with a $\log_2$-fold change >3. This high degree of difference puts ArlA1 and ArlA2 among the 15 most de-enriched proteins in all of these six strains. The chemotaxis proteins CheW (HVO_1225), CheR (HVO_1222), and CheA (HVO_1223), along with the HemAT1 (HVO_1484) and Htr15 (HVO_3005) transducers are also commonly present as the most downregulated proteins in these strains. Furthermore, the Htr7 transducer (HVO_1999) is significantly de-enriched in the Δ*0885* mutant, and completely absent in the other mutants with swarming deficit. Htr7 is postulated to be involved in chemotaxis and shape transition signaling cascades (Schiller et al. 2024). The strong fold-changes observed for Δ*0885*, for which most motility related proteins are down-regulated by a factor of 5 $\log_2$-fold, support the preliminary swarming deficit phenotype.

Furthermore, a segmentation analysis of the differentially regulated proteins differentiates the strains with biologically observed increased and decreased swarming, forming distinct clusters (Fig. 3A). Δ*1352*, Δ*1533*, and Δ*2983A*, which show increased motility, cluster together and separate from the swarming deficit strains, with the exception of Δ*0695* (Supporting Fig. 4). Phenotypical data reported in Üresin et al. (in preparation), showed that, in contrast to the other swarming deficit mutants, Δ*0695* only has a minor phenotype, that recovers to levels similar to the wild-type over time. We hypothesize that the proteomic data reflects compensatory mechanisms used by *H. volcanii* to revert the phenotype.

The heatmap (Fig. 3A) for the remaining swarming deficit strains also shows strong differential abundance or the complete absence of proteins previously reported to be involved in cell motility, highlighted in bold (Schiller et al. 2024). The transducer BasT (HVO_0554) is completely absent in all of these strains and upregulated in the strains with improved swarming. MpcT (HVO_0420) is also down regulated in some motility defective strains *(*Δ*2901*, Δ*0885*, Δ*0767*, and Δ*2523*).

We then analyzed the list of proteins that were identified in the wild type but were missing in the deletion strains. When plotting their ranked intensities (Supporting Fig. 5), we observed that a small number of these proteins are highly abundant in the wild type. The absence of these proteins in the deletion strains suggests that there is likely a biological effect diminishing or nullifying their expression, and that it is not merely a missing value due to low abundance issue. Interestingly, for Δ*2523*, Δ*2901*, and Δ*2982*, we observed a complete absence of three high abundance proteins involved in motility. The HemAT transducer protein (HVO_1484) was not expressed in either replicate of the Δ*2523* and Δ*2901* mutants, and the archaellin proteins ArlA1 (HVO_1210) and ArlA2 (HVO_1211) were also not expressed in the Δ*2982* strain. Each of these proteins is also detected with a significant number of unique peptides for the matched protein: an average of 18 per sample for HemAT, ∼7 per sample for ArlA1 and ∼3 per sample for ArlA2.

The strains Δ*1352*, Δ*1533*, and Δ*2983A* showed *in vivo* improvement in their swarming. While the first two are consistent with our differential expression observations, i.e. arCOG and GO enrichments for cell motility and chemotaxis terms, we do not see such an enrichment for the latter. It seems that while the reduced motility we observed for several strains is related to depletion of proteins directly related to motility, the increased swarming in these strains might be an indirect effect or mediated by other proteins. Interestingly, no proteins are depleted with $\log_2$-fold changes ≤ -2 in all three strains and the only protein consistently enriched in all three strains with $\log_2$-fold changes ≥ 2 is the HVO_1228. HVO_1228 is annotated as "Fbr", a cytochrome-like protein on HaloWeb, and as a halocyanin/DUF5059 on Uniprot. The protein has not been implied in cell motility or chemotaxis before, and no metadata from other organisms are available that could explain its role in cell motility.

Another important observation we had was that a number of proteins were differentially expressed in several of the deletion mutants (Table 2). The protein Htr15 (HVO_3005) was differentially expressed in 10 out of 19 mutants and ArlA2 (HVO_1211) was identified in 6 mutants. Htr15 had previously only been identified in two other studies (Schulze et al. 2021, Knüppel et al. 2021). Both proteins are involved in cell motility, explaining their prevalent differential regulation in the studied strains (Legerme et al. 2016, Schiller et al. 2024) (Üresin et al., in preparation).

While some proteins are consistently found in multiple comparisons as they have been shown to be associated with the phenotypes, it is not unusual to identify stress response related proteins differentially regulated in gene deletion mutants. In the stress proteome of *H. volcanii* reported by (Jevtić et al. 2019), the proteins MpcT (HVO_0420) and Gap1 (HVO_0478) were found to be downregulated in all stress conditions tested. HVO_B0134, Htr38 (HVO_2220), and HVO_1613 were upregulated in the universal stress proteome. It is reasonable to consider that these proteins were differentially regulated in our deletion mutants due to stress caused by the respective gene deletion other than due to association with the phenotype (Kohram et al. 2024).

In addition, we verified a differential regulation of RdfA (HVO_2174) in five different strains for which we also observed swarming deficit. Schiller et al. (2024) annotated this protein as rod-determining-factor A (RdfA) due to its requirement for rod formation and motility (Schiller et al. 2024). The same holds true for Sph3 (HVO_2175) and MpcT (HVO_0420), proteins which were also previously associated with cell-shape (Schiller et al. 2024). The conserved uncharacterized protein HVO_A0590 is also downregulated with $\log_2$-fold changes larger than 3 in three mutants with swarming deficit *(*Δ*0885*, Δ*0546*, and Δ*2523*), and with a depletion of RdfA and Sph3. These data are in accordance with the observations from Chatterjee et al. in their study on quorum sensing and motility transitions (Chatterjee et al. 2025).

Discussion

Here we present a comprehensive analysis of the *H. volcanii* proteome, and its remodelling upon deletion of 19 zinc-finger μ-proteins. Using a sensitive and robust dia-PASEF workflow on a timsTOF instrument, we reproducibly detected over 3000 proteins in *H. volcanii* using stringent quality metrics. This considerably extends the number of proteins experimentally detected by mass spectrometry in this organism, adding to the substantial work previously included in the ArcPP database.

The successful quantitation of over 2000 proteins in each strain enables a detailed analysis of proteome remodeling in the zinc-finger μ-proteins deletion strains and how they correlate with the observed growth phenotypes and swarming behaviour.

We observed for all 19 studied strains, that the deletion of zinc finger μ-proteins influences the level of other proteins. While three

Table 2 Proteins are differentially expressed in multiple mutants. A number of proteins were found to be differentially expressed in multiple mutants. Presented here are proteins identified in three or more deletion strains, and their annotation from ArcPP.

Protein	Differentially expressed in how many mutants	Annotation
HVO_3005	10	(htr15b) transducer protein Htr15
HVO_2061	9	(agl6) low-salt glycan biosynthesis hexosyltransferase Agl6
HVO_1222	8	(cheR) protein-glutamate O-methyltransferase CheR
HVO_1223	8	(cheA) taxis sensor histidine kinase CheA
HVO_1484	7	(hemAT1) transducer protein HemAT
HVO_2175	6	(sph3) Smc-like protein Sph3
HVO_1211	6	(arlA2) archaellin A2
HVO_0420	5	(mpcT) transducer protein MpcT
HVO_2462	5	(htr37) transducer protein Htr37
HVO_2606	5	PQQ repeat protein
HVO_1221	5	Protein of unknown function (DUF439) superfamily
HVO_B0134	5	ABC-type transport system ATP-binding protein
HVO_1210	5	(arlA1) archaellin A1
HVO_1228	5	(hcpE) DUF5059 domain/halocyanin domain protein
HVO_0284	4	cupin 2 barrel domain protein
HVO_2220	4	(htr38) transducer protein Htr38
HVO_2174	4	conserved hypothetical protein
HVO_2670	4	FAD-dependent oxidoreductase (GlcD/DLD_GlcF/GlpC domain fusion protein)
HVO_0478	4	(gap1) glyceraldehyde-3-phosphate dehydrogenase (NAD(P)) (phosphorylating)
HVO_0812	3	(ppsA) phosphoenolpyruvate synthase
HVO_1126	3	(hemAT2) transducer protein HemAT
HVO_1613	3	dolichyl-phosphate hexosyltransferase
HVO_2872	3	conserved hypothetical protein
HVO_0422	3	NP_1176A family transcription regulator
HVO_2074	3	probable secreted glycoprotein
HVO_1650	3	(ppk1) polyphosphate kinase
HVO_1202	3	conserved hypothetical protein

of the studied mutants had very small effects on the proteome, the majority of them had larger effects, which included statistically significant changes in arCOG categories. Predominantly, changes in proteins related to cell motility were observed on the proteome, with seven mutants showing a downregulation of proteins linked to motility pathways, and four mutants showing an overexpression. This is consistent to what has been observed in the phenotypical evaluation of these strains.

The coherence between phenotypical and proteomic data for cell motility indicates that the roles of other zinc-finger μ-proteins might be inferred from differential expression analysis by MS-based proteomics. Six out of the 19 strains showed no motility phenotype but presented differential regulation of arCOG categories: Δ*1359* showed an enrichment in amino acid transport and metabolism and signal transduction, Δ*29083A* in secondary metabolite biosynthesis, and Δ0085 in inorganic ion transport and metabolism. Δ*A0089*, Δ*1118*, and Δ*2805A* had similar depletions: cell wall/membrane/envelope biogenesis and PTMs, protein turnover and chaperones. Clustering analysis reinforces the similarities between these three strains, but little difference to the wild type and to the two strains which showed no arCOG enrichment: Δ*1677* and Δ*2400*.

Interestingly, the Δ*1359*, which showed a growth deficit in glycerol as phenotype, had chemotaxis as an enriched term in GO analysis. This zinc-finger μ-protein was previously linked to motility as it seems to be involved in the transition from rod to disks (Schiller et al. 2024). A deletion of this protein could implicate in reduced disk formation and thus hypermotility (Duggin et al. 2015, Schiller et al. 2024).

Our data also corroborates previous findings that disk-deficient phenotypes seem to be hypermotile (Duggin et al. 2015, Schiller et al. 2024). *Haloferax volcanii* was initially described to have a pleomorphic cell shape and to be non-motile (Mullakhanbhai and Larsen 1975), and this remained general belief for decades. In 2010 researchers found that *H. volcanii* becomes motile under microaerobic conditions (Tripepi et al. 2010). Recently it has been reported that *H. volcanii* cells are rod-shaped during early growth (OD_{600} of 0.1). Rod-shaped cells are typically found at the motile borders of colonies in soft agar and are associated with increased motility (Duggin et al. 2015, Hackley et al. 2024). In late-log growth and in the centre of motility halos, *H. volcanii* change their morphology to flat, pleomorphic disks. Two important proteins for the formation of rod-shaped mobile cells are the rod-determining factor A protein, RdfA (HVO_2174) and Sph3 (HVO_2175). Both proteins are consistently largely de-enriched in the strains which presented swarming deficit. In contrast, we did not detect any differential abundance in the disk-determining factor A protein DdfA (HVO_2176). This is either due to the low number of peptides that

can, in theory, be generated from this short sequence protein (66 aa long), or perhaps it is involved in other mechanisms not influenced by the studied zinc-finger μ-proteins.

In conclusion, our study provides valuable insights into the proteomic landscape of *H. volcanii*, and lays the groundwork for future research into the functional roles of zinc finger μ-proteins and their involvement in cellular processes. Our approach also demonstrates the increase in proteome coverage as well as quantitative depth and accuracy of dia-PASEF. The implementation of our work in the ArcPP database further facilitates future projects by facilitating meta-analyses and *in silico* studies of the *H. volcanii* proteome.

Acknowledgements

The authors would like to acknowledge Imke Wüllenweber (Max-Planck Institute of Biophysics) and Mathias Hammelmann (Institute of Molecular Biosciences, Goethe University) for their excellent technical assistance.

Supplementary material

Supplementary material is available at *FEMSML Journal* online.

Conflicts of interest

None declared.

Funding

Financial support was provided by the Max Planck Society and by the Deutsche Forschungsgemeinschaft (DFG) in the framework of the Priority Program “Small Proteins in Prokaryotes: An unexplored world” (Grant No. 3542/1–1 to J.D.L. and So264/26 to J.S.).

Data availability statement

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al. 2022) partner repository with the dataset identifier PXD066008.

References

Ahrens CH, Wade JT, Champion MM *et al.* A practical guide to small protein discovery and characterization using mass spectrometry. *J Bacteriol* 2022;**204**:e0035321. https://doi.org/10.1128/JB.00353-21.

Allers T, Ngo H-P, Mevarech M *et al.* Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the leuB and trpA genes. *Appl Environ Microb* 2004;**70**:943–53. https://doi.org/10.1128/aem.70.2.943-953.2004.

Breuert S, Allers T, Spohn G *et al.* Regulated polyploidy in halophilic archaea. *PLoS One* 2006;**1**:e92. https://doi.org/10.1371/journal.pone.0000092.

Burton AT, Zeinert R, Storz G. Large roles of small proteins. *Annu Rev Microbiol* 2024;**78**:1–22. https://doi.org/10.1146/annurev-micro-112723-083001.

Cassidy L, Prasse D, Linke D *et al.* Combination of bottom-up 2D-LC-MS and semi-top-down GelFree-LC-MS enhances coverage of proteome and low molecular weight short open reading frame encoded peptides of the archaeon methanosarcina mazei. *J Proteome Res* 2016;**15**:3773–83. https://doi.org/10.1021/acs.jproteome.6b00569.

Chatterjee P, Consoli CE, Schiller H *et al.* Quorum Sensing Mediates Morphology and Motility Transitions in the Model Archaeon Haloferax Volcanii *bioRxiv* 2025. https://doi.org/10.1128/mbio.00906-25.

Chomczynski PSN. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 1987;**162**:156–9. https://doi.org/10.1016/0003-2697(87)90021-2.

Dambeck M, Soppa J. Characterization of a Haloferax volcanii member of the enolase superfamily: deletion mutant construction, expression analysis, and transcriptome comparison. *Arch Microbiol* 2008;**190**:341–53. https://doi.org/10.1007/s00203-008-0379-1.

Dantuluri S, Wu Y, Hepowit NL *et al.* Proteome targets of ubiquitin-like samp1ylation are associated with sulfur metabolism and oxidative stress in *Haloferax volcanii*. *Proteomics* 2016;**16**:1100–10. https://doi.org/10.1002/pmic.201500153.

Demichev V, Messner CB, Vernardis SI *et al.* DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods* 2020;**17**:41–44. https://doi.org/10.1038/s41592-019-0638-x.

Demichev V, Szyrwiel L, Yu F *et al.* dia-PASEF data analysis using FragPipe and DIA-NN for deep proteomics of low sample amounts. *Nat Commun* 2022;**13**. :3944https://doi.org/10.1038/s41467-022-31492-0.

Duggin IG, Aylett CHS, Walsh JC *et al.* CetZ tubulin-like proteins control archaeal cell shape. *Nature* 2015;**519**:362–5. https://doi.org/10.1038/nature13983.

Eom KS, Cheong JS, Lee SJ. Structural analyses of zinc finger domains for specific interactions with DNA. *J Microbiol Biotechnol* 2016;**26**:2019–29. https://doi.org/10.4014/jmb.1609.09021.

Finkel Y, Mizrahi O, Nachshon A *et al.* The coding capacity of SARS-CoV-2. *Nature* 2021;**589**:125–30. https://doi.org/10.1038/s41586-020-2739-1.

Franco PHC, Zeinert R, Meier-Credo J *et al.* Detection and quantitation of small proteins using mass spectrometry. *Mol Cell Proteomics* 2025;**24**:101052. https://doi.org/10.1016/j.mcpro.2025.101052.

Fuchs S, Kucklick M, Lehmann E *et al.* Towards the characterization of the hidden world of small proteins in *Staphylococcus aureus*, a proteogenomics approach. *PLoS Genet* 2021;**17**:e1009585. https://doi.org/10.1371/journal.pgen.1009585.

Goeminne LJE, Sticker A, Martens L *et al.* MSqRob takes the missing hurdle: uniting intensity- and count-based proteomics. *Anal Chem* 2020;**92**:6278–87. https://doi.org/10.1021/acs.analchem.9b04375.

Hackley RK, Hwang S, Herb JT *et al.* <scp>TbsP</scp>and <scp>TrmB</scp>jointly regulate gapII to influence cell development phenotypes in the archaeon *Haloferax volcanii*. *Mol Microbiol* 2024;**121**:742–66. https://doi.org/10.1111/mmi.15225.

Hadjeras L, Bartel J, Maier L-K *et al.* Revealing the small proteome of *Haloferax volcanii* by combining ribosome profil-

ing and small-protein optimised mass spectrometry. *microLife* 2023;**4**:uqad001. https://doi.org/10.1093/femsml/uqad001.

HaileMariam M, Eguez RV, Singh H *et al.* S-trap, an ultrafast sample-preparation approach for shotgun proteomics. *J Proteome Res* 2018;**17**:2917–24. https://doi.org/10.1021/acs.jproteome.8b00505.

Hammelmann M, Soppa J. Optimized generation of vectors for the construction of *Haloferax volcanii* deletion mutants. *J Microbiol Methods* 2008;**75**:201–4. https://doi.org/10.1016/j.mimet.2008.05.029.

Hemm MR, Weaver J, Storz G. *Escherichia coli* small proteome. *EcoSal Plus* 2020;**9**. https://doi.org/10.1128/ecosalplus.ESP-0031-2019.

Hobbs EC, Yin X, Paul BJ *et al.* Conserved small protein associates with the multidrug efflux pump AcrB and differentially affects antibiotic resistance. *Proc Natl Acad Sci USA* 2012;**109**:16696–701. https://doi.org/10.1073/pnas.1210093109.

Hoffman LM, Jensen CC, Beckerle MC. Phosphorylation of the small heat shock protein HspB1 regulates cytoskeletal recruitment and cell motility. *MBoC* 2022;**33**:ar100. https://doi.org/10.1091/mbc.e22-02-0057.

Jevtić Ž, Stoll B, Pfeiffer F *et al.* The response of *Haloferax volcanii* to salt and temperature stress: a proteome study by label-free mass spectrometry. *Proteomics* 2019;**19**:1800491. https://doi.org/10.1002/pmic.201800491.

Kaminski L, Guan Z, Abu-Qarn M *et al.* AglR is required for addition of the final mannose residue of the N-linked glycan decorating the *Haloferax volcanii* S-layer glycoprotein. *Biochimica et Biophysica Acta (BBA)—General Subjects* 2012;**1820**:1664–70. https://doi.org/10.1016/j.bbagen.2012.06.014.

Kirkland PA, Busby J, Stevens S, Jr *et al.* Trizol-based method for sample preparation and isoelectric focusing of halophilic proteins. *Anal Biochem* 2006;**351**:254–9. https://doi.org/10.1016/j.ab.2006.01.017.

Knüppel R, Trahan C, Kern M *et al.* Insights into synthesis and function of KsgA/Dim1-dependent rRNA modifications in archaea. *Nucleic Acids Res* 2021;**49**:1662–87. https://doi.org/10.1093/nar/gkaa1268.

Kohram M, Sanderson AE, Loui A *et al.* Nonlethal deleterious mutation–induced stress accelerates bacterial aging. *Proc Natl Acad Sci USA* 2024;**121**:e2316271121. https://doi.org/10.1073/pnas.2316271121.

Koopmans F, Li KW, Klaassen RV *et al.* MS-DAP platform for downstream data analysis of label-free proteomics uncovers optimal workflows in benchmark data sets and increased sensitivity in analysis of alzheimer's biomarker data. *J Proteome Res* 2023;**22**:374–86. https://doi.org/10.1021/acs.jproteome.2c00513.

Kushwaha AK, Dwivedi S, Mukherjee A *et al.* Plant microProteins: small but powerful modulators of plant development. *iScience* 2022;**25**:105400. https://doi.org/10.1016/j.isci.

Legerme G, Yang E, Esquivel R *et al.* Screening of a *Haloferax volcanii* transposon library reveals novel motility and adhesion mutants. *Life* 2016;**6**:41. https://doi.org/10.3390/life6040041.

Li X, Han M, Zhang H *et al.* Structures and biological functions of zinc finger proteins and their roles in hepatocellular carcinoma. *Biomark Res* 2022;**10**:2. https://doi.org/10.1186/s40364-021-00345-1.

Madigan M Sattley W Aiyer J Stahl D Buckley D *Brock Biology of Microorganisms*. 16th edition, Pearson Education, Inc.; 2021, 1128. 978-1-292-40506-3

Matthews JM, Sunde M. Zinc fingers–folds for many occasions. *IUBMB Life* 2002;**54**:351–5. https://doi.org/10.1080/15216540216035.

McMillan LJ, Hwang S, Farah RE *et al.* Multiplex quantitative SILAC for analysis of archaeal proteomes: a case study of oxidative stress responses. *Environ Microbiol* 2018;**20**:385–401. https://doi.org/10.1111/1462-2920.14014.

Meier F, Brunner A-D, Frank M *et al.* diaPASEF: parallel accumulation–serial fragmentation combined with data-independent acquisition. *Nat Methods* 2020;**17**:1229–36. https://doi.org/10.1038/s41592-020-00998-0.

Meier-Credo J, Heiniger B, Schori C *et al.* Detection of known and novel small proteins in pseudomonas stutzeri using a combination of bottom-up and digest-free proteomics and proteogenomics. *Anal Chem* 2023;**95**:11892–900. https://doi.org/10.1021/acs.analchem.3c00676.

Mullakhanbhai MF, Larsen H. *Halobacterium volcanii* spec. nov., a Dead Sea halobacterium with a moderate salt requirement. *Arch Microbiol* 1975;**104**:207–14. https://doi.org/10.1007/bf00447326.

Nagel C, Machulla A, Zahn S *et al.* Several one-domain zinc finger μ-proteins of *Haloferax Volcanii* are important for stress adaptation, biofilm formation, and swarming. *Genes* 2019;**10**:361. https://doi.org/10.3390/genes10050361.

Perez-Riverol Y, Bai J, Bandla C *et al.* The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res* 2022;**50**:D543–52. https://doi.org/10.1093/nar/gkab1038.

Pfeiffer F, Broicher A, Gillich T *et al.* Genome information management and integrated data analysis with HaloLex. *Arch Microbiol* 2008;**190**:281–99. https://doi.org/10.1007/s00203-008-0389-z.

Pohlschroder M, Schulze S, Pfeiffer F *et al. Haloferax volcanii*: a versatile model for studying archaeal biology. *J Bacteriol* 2025;**207**. https://doi.org/10.1128/jb.00062-25.

Portnoy VA, Bezdan D, Zengler K. Adaptive laboratory evolution–harnessing the power of biology for metabolic engineering. *Curr Opin Biotechnol* 2011;**22**:590–4. https://doi.org/10.1016/j.copbio.2011.03.007.

Sandmann C-L, Schulz JF, Ruiz-Orera J *et al.* Evolutionary origins and interactomes of human, young microproteins and small peptides translated from short open reading frames. *Mol Cell* 2023;**83**:994–1011.e18. https://doi.org/10.1016/j.molcel.2023.01.023.

Schiller H, Hong Y, Kouassi J *et al.* Identification of structural and regulatory cell-shape determinants in *Haloferax volcanii*. *Nat Commun* 2024;**15**:1414. https://doi.org/10.1038/s41467-024-45196-0.

Schiller H, Young C, Schulze S *et al.* Accessible and Insightful Scientific Learning Experiences Using the Microorganism Haloferax volcanii *Methods Mol Biol* 2022:531–45.https://doi.org/10.1007/978-1-0716-2445-6_34

Schulze S, Adams Z, Cerletti M *et al.* The archaeal proteome project advances knowledge about archaeal cell biology through comprehensive proteomics. *Nat Commun* 2020;**11**:3145. https://doi.org/10.1038/s41467-020-16784-7.

Schulze S, Pfeiffer F, Garcia BA *et al.* Comprehensive glycoproteomics shines new light on the complexity and extent of gly-

cosylation in archaea. *PLoS Biol* 2021;**19**:e3001277. https://doi.org/10.1371/journal.pbio.3001277.

Storz G, Wolf YI, Ramamurthi KS. Small proteins can no longer be ignored. *Annu Rev Biochem* 2014;**83**:753–77. https://doi.org/10.1146/annurev-biochem-070611-102400.

Tripepi M, Imam S, Pohlschro¨ Der M. *Haloferax volcanii* flagella are required for motility but are not involved in PibD-dependent surface adhesion. *J Bacteriol* 2010;**192**:3093–102. https://doi.org/10.1128/jb.00133-10.

Üresin D, Pyper DJ, Borst A *et al.* Characterization of the zinc finger μ-protein HVO_0758 from *Haloferax volcanii*: biological roles, zinc binding, and NMR solution structure. *Front Microbiol* 2023;**14**:1280972. https://doi.org/10.3389/fmicb.2023.1280972.

Üresin D, Schulte J, Morgner N *et al.* C(P)XCG proteins of *Haloferax volcanii* with predicted zinc finger domains: the majority bind zinc, but several do not. *Int J Mol Sci* 2024;**25**:7166. https://doi.org/10.3390/ijms25137166.

Wessel DFUI. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem* 1984;**138**:141–3. https://doi.org/10.1016/0003-2697(84)90782-6.

Wright Z, Seymour M, Paszczak K *et al.* The small protein MntS evolved from a signal peptide and acquired a novel function regulating manganese homeostasis in *Escherichia coli*. *Mol Microbiol* 2024;**121**:152–66. https://doi.org/10.1111/mmi.15206.

Wu T, Hu E, Xu S *et al.* clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *The Innovation* 2021;**2**:100141. https://doi.org/10.1016/j.xinn.2021.100141.

Xu S, Hu E, Cai Y *et al.* Using clusterProfiler to characterize multiomics data. *Nat Protoc* 2024;**19**:3292–320. https://doi.org/10.1038/s41596-024-01020-z.

Yu G, Wang L-G, Han Y *et al.* clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* 2012;**16**:284–7. https://doi.org/10.1089/omi.2011.0118.

Received: 19 July 2025. **Revised:** 3 February 2026. **Accepted:** 8 February 2026