
Suggesting that proteome-level chemical traits may reflect the selective pressures exerted by specific environments. Alejandro Uribe¹, Josep Ramoneda² and Malu Calle¹

¹University of Vic-UCC, Centre de es estudios avanzados de Blanes- CSIC.

Abstract

Motivation: Drought is one of the most severe environmental stresses affecting terrestrial ecosystems, generating conditions that have a major influence in the composition of the microbial communities of soil. While much work has focused on genetic and physiological mechanisms of microorganism stress tolerance, less attention has been paid to the chemical properties of microbial proteomes as potential signatures of adaptation. Previous studies have demonstrated systematic links between environmental factors such as salinity and pH and proteome composition, suggesting that proteome-level chemical traits may reflect the selective pressures exerted by specific environments. Similarly, we hypothesize that bacterial responses to drought are also associated with distinct proteome chemical properties.

Results: Using a curated dataset of 1,606 predicted proteomes mapped to bacterial taxa previously classified as drought-sensitive or drought-responsive from eight soil-moisture manipulation studies, we tested whether proteome chemical properties were associated with drought adaptation. Average proteome traits, including isoelectric point (pI), carbon oxidation state (Zc), and hydration number (nH₂Og), showed consistent differences between the two groups. We found that these proteomic traits displayed a significant phylogenetic signal, such that taxa that were phylogenetically closer tended to have more similar properties. Nevertheless, pairwise comparisons of closely related sensitive and responsive organisms within the same phylum revealed a clear difference in their extracellular protein chemistry, particularly in the pI of extracellular proteins. Overall, the results demonstrate that proteome chemistry provides a measurable signal distinguishing drought-responsive from drought-sensitive bacteria, likely underlying its adaptive role to cope with water limitation.

Contact: alejandro.uribe@uvic.cat

Supplementary information: Supplementary data are available at https://github.com/AlejandroUL/AlejandroUribe_TFM

1 Introduction

Anthropogenic activities are responsible for the fast global warming the world is experiencing, gases like CO₂, CH₄, N₂O, and so on are inducing climate change, generating an increasing of the temperatures at higher rates than before. During 2011–2020, global land surface temperature increased by 1.59 °C (1.34–1.83 °C) relative to 1850–1900 [1]. Rising temperatures increase evapotranspiration, this added to changes in rain patterns is leading to a decreasing soil moisture. [2] causing an increase of 74% on average of areas in drought during 2018–2022 compared with 1981–2017 [3].

Soil bacteria carry multiple ecosystem functions that are highly sensitive to drought, they drive the disintegration of organic matter and the mineralization of nutrients that sustain plant growth, regulating both nitrogen and phosphorus

availability through processes such as N mineralization/nitrification and phosphate solubilization [4, 5]. They also build and maintain soil structure by producing extracellular polymeric substances that forms aggregates with filamentous fungi enhancing water retention and erosion resistance [6]. Beyond nutrient supply and structure, soil microbiomes protect plants by generating communities that limit pathogen pressure, a function that can be disrupted by extreme moisture deficits [7, 8]. Drought also alters greenhouse gases flux by restricting microbial respiration and denitrification, often reducing N₂O emissions and shifting CO₂ and CH₄ dynamics depending on moisture. [3; 9]. It was also observed that repeated droughts can create ecological memory in soil microbiomes, changing community composition and multifunctionality and thereby modifying ecosystem responses to later drought episodes [10, 11]. These bacteria have generated several strategies to cope with osmotic and water limitation caused by drought, some

taxa survive these adverse conditions by forming resistant spores or entering in dormancy state, as observed in *Bacillus*, actinobacteria, and certain fungi that remain viable until favorable conditions return [12,13]. Others rely on the synthesis of compatible solutes or xeroprotectants, such as trehalose, ectoine, hydroxyectoine, and polyols, which stabilize proteins and membranes and help maintain osmotic balance under water depletion [14]. These adaptive mechanisms, ranging from physiological adjustments to structural defenses shape the composition of microbial communities under drought and ultimately influence their functional capacity in soil ecosystems.

While physiological and structural adaptations such as sporulation or osmolyte production have been widely studied, there is growing evidence that the inherent chemical properties of microbial proteomes may also play an important role in environmental adaptation. For example, Dick and Tan (2023) [15] demonstrated that the carbon oxidation state (Zc) of microbial community proteomes consistently mirrors environmental redox gradients across hydrothermal systems, stratified lakes, and shale gas wells, indicating that protein chemistry encodes geochemical conditions. Similarly, Cabello-Yeves and Rodriguez-Valera (2019) [16] showed that transitions between marine and freshwater environments involve extensive shifts in amino acid composition and isoelectric points (pI) of bacterial proteomes, particularly in extracellular proteins, reflecting the strong selective pressure exerted by salinity. Taken together, these observations align with established principles of protein biophysics: Protein function depends on a tightly bound hydration shell and on electrostatic interactions that are sensitive to water activity and ionic strength; reductions in hydration or shifts in ion composition alter folding dynamics and can impair activity [17, 18]. Proteins are least soluble near their isoelectric point, so adaptations that shift pI and surface charge can preserve net charge and electrostatic repulsion at the ambient pH/ionic regime, limiting aggregation effects that are especially relevant for extracellular proteins directly exposed to soil porewater chemistry [19]. The selection for highly acidic, strongly hydrated proteomes in halophiles exemplifies this principle, where negative protein surfaces retain water and remain soluble in high-salt environments [20]. Building on this framework, we believe that bacterial adaptation to drought may also involve distinct proteome-level chemical properties, and that this traits could help distinguish drought-responsive from drought-sensitive taxa.

2 Methods

2.1 Predicted proteomes and metadata used in this study

We assembled a comparative dataset of 1,606 bacterial proteomes to test whether proteome chemical properties differ

between drought-responsive and drought-sensitive taxa. Response labels and metadata were taken from a previously curated classification developed by our research group and were not modified here. That classification was built from eight independent studies in which soil moisture was experimentally manipulated and its influence on soil bacterial community composition was assessed. For each study, raw 16S rRNA gene reads were downloaded from the NCBI Sequence Read Archive, quality-filtered, and processed to infer amplicon sequence variants (ASVs). Differential abundance was then examined in order to identify taxa that increased or decreased in relative abundance under drought relative to controls, obtaining in this way the labels known as drought-responsive or drought-sensitive, respectively. ASVs were aligned to a reference genome database (Genome Taxonomy Database, GTDB vR220 [21]) to obtain representative genomes of those taxa. Phylogenetic context was provided by the GTDB reference tree (release R220). For each mapped accession, the predicted protein sequences from the corresponding GTDB genome were retrieved from the GTDB proteome database (R220) [2].

2.2 Extracellular and intracellular partitioning of proteins

We predicted secretory signal peptides (SP) with SignalP 6.0 using python version 3.9.21 [23], which classifies sequences into five SP types: SP (Sec/SPI), LIPO (Sec/SPII), PILIN (Sec/SPIII), TAT (Tat/SPI), and TATLIPO (Tat/SPII); sequences without a signal peptide are labeled OTHER. In this study, we defined the extracellular set strictly as proteins predicted with SP (Sec/SPI), i.e., substrates of the Sec pathway cleaved by Signal Peptidase I, a well-studied export route. Proteins predicted as LIPO, PILIN, TAT, or TATLIPO were not included in the extracellular set because these classes frequently correspond to membrane-anchored lipoproteins, pilus subunits, or proteins that remain periplasmic/surface-attached rather than freely secreted [24-27]. All proteins labeled OTHER were treated as intracellular.

2.3 Chemical metrics of the proteins

All computations were performed in R (v 4.3.2) and Python (v 3.13.1). Each protein isoelectric point (pI), was calculated in Python using ProtParam from Biopython (v 1.85), here the method proposed by Bjellqvist is applied in order to calculate this metric [28,29].

Carbon oxidation state (Zc) and hydration number (nH:Og) were computed in R with the canprot package (v 2.0.0) [26]. Canprot works calculating the average proteome properties from the overall aminoacid composition of it, and in order to obtain this the ProtParam routine from biopython was used. For Zc Canprot uses calculations based on the elemental formula of proteins CcHhNnOoSs as seen in equation 1:

$$Z_c = \frac{-h+3n+2o+2s}{c} (1)$$

While nH₂Og accounts for the theoretical number of water molecules involved in the reaction that produces certain protein normalized by its mass [31,32].

For each proteome and for each subcellular partition (global, extracellular, intracellular), we summarized metrics as the arithmetic mean across proteins. All metrics were computed independently per proteome.

2.4 Closest-relative pairing

Common evolutionary history creates common sequence patterns that can create spurious correlations between microbial traits and genome-derived features. To minimize broad phylogenetic effects, we paired each drought-sensitive proteome with the most closely related drought-responsive proteome using the phylogenetic tree obtained from previous research in the group and then we compared their chemical properties. We first read the GTDB-based bacterial tree and computed cophenetic pairwise distances (sum of branch lengths between tips) using the ape package (v 5.8-1) from R.

From the complete distance matrix, we built a submatrix with rows restricted to sensitive genomes and columns to responsive genomes. For each sensitive genome, we identified the single responsive genome with minimum cophenetic distance and recorded that pair together with the tree distance. This produced one closest-relative pair per sensitive genome.

For each pair, we retrieved the precomputed per-proteome metrics (pI, nH₂Og, Z_c) for both members, assembled them side by side, and calculated both absolute differences and ratios (Sensitive/Responsive).

2.5 Statistical analyses

All statistical analyses were performed in R (v4.3.2). In order to find the most suitable method for the significance, the distribution of the analyzed data was assessed using Shapiro-Wilk test. Based on that result the decision whether to make the one-sample one-sided t-test or Wilcoxon rank sums test was taken on the normality acceptance or rejection.

For drought-responsive vs drought sensitive comparison between phyla seen in the boxplots, two-sided Wilcoxon signed-rank tests was used due to the non-existence of normally distributed pI values within this data. In this comparison BH correction was used due to multiple hypothesis testing.

To test the differences in the magnitude of pI peaks between DR and DS extracellular proteomes, we ran Wilcoxon Rank-Sum Test test on the proteins within a range of ± 0.5 Ip points around the peak.

Finally, to determine whether phylogenetic relatedness was associated with proteome properties, we performed a Mantel test. Phylogenetic distances between genomes were obtained from the reference tree, and Euclidean distances of proteomes pI were calculated for the same set of genomes. The Mantel statistic is the Pearson correlation between the two distance matrices. Significance was assessed utilizing 999 of the distance matrix labels that were randomly permuted, as carried out through R package vegan's (v2.7-1) mantel function.

Statistical significance was evaluated at $\alpha = 0.05$.

2.6 Visualization

Figures were generated in R using ggplot2 (v 3.5.2). Boxplots summarize per-proteome distributions by group and by phylum. Density plots were made just for pI values using this metric from each individual protein for extracellular and whole proteomes. Phylogenetic trees was rendered using ITOL [33] and annotated with response labels and pI.

2.7 Reproducibility and availability

All scripts used to calculate chemical metrics, partition proteomes, perform phylogenetic pairing, and generate figures are available as supplementary material together with intermediate data tables required to reproduce the analyses in the following Github file:

https://github.com/AlejandroUL/AlejandroUribe_TFM

3 Results

3.1 Data Structure

Among 1,606 proteomes, 1,048 were classified as drought-responsive (DR) and 558 as drought-sensitive (DS). The DR set comprised 5,416,113 proteins and the DS set 2,565,670 (7,981,783 in total). To quantify class imbalance arising from unequal numbers of proteomes and proteins

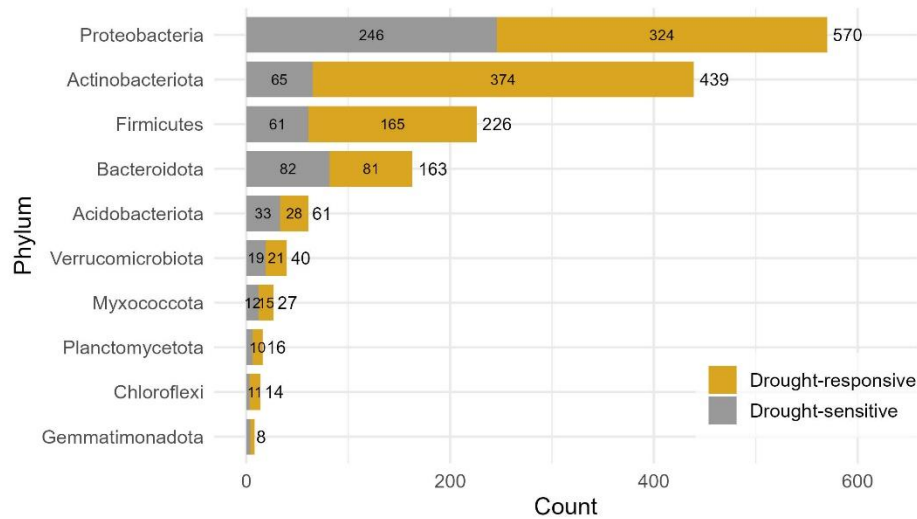


Figure 1 Number of proteomes per phylum stacked by drought response.

(by proteomes: 65.3% DR, 34.7% DS; by proteins: 67.9% DR, 32.1% DS), we computed Shannon entropy H using a base-2 logarithm. This metric quantifies the balance of a distribution, for two classes, H ranges from 0 (all observations in one class) to 1 bit (perfect balance) [34]. We obtained $H=0.932$ for proteome counts and $H=0.906$ for protein counts, ~93% and ~91% of the two-class maximum, respectively. These values indicate a moderate skew toward DR rather than extreme imbalance.

As seen in figure 1, The top 10 phyla (out of 26) account for 1,564 of the 1,606 proteomes. Proteobacteria (570) and Actinobacteriota (439) dominate the sampling, together comprising circa 65% of these proteomes; Firmicutes contribute 226, and Bacteroidota 163. Within-phyla response tallies are uneven: Actinobacteriota are strongly enriched in drought-responsive (DR) taxa (374 of 439, 85%), as are Firmicutes (165 of 226, 73%) and Chloroflexi (11 of 14, 79%). Proteobacteria show a milder DR skew (324 of 570, 57%), whereas Bacteroidota are essentially balanced (81 DR vs. 82 drought-sensitive, DS) and Acidobacteriota lean DS (28 DR vs. 33 DS). Several smaller phyla are DR-biased but with low counts (for example, Gemmatimonadota, 8 of 8 DR).

This phylum-level pattern mirrors widely reported drought responses in soils, where Gram-positive groups, particularly Actinobacteria and many Firmicutes, tend to increase in relative abundance, while several Gram-negative lineages decline. Multiple field and mesocosm studies have documented Actinobacteria and Firmicutes enrichment under water limitation or during dry years, often attributed to traits such as thicker cell walls and spore formation that enhance survival under desiccation and osmotic stress [35–36]. This concordance with prior literature supports interpreting the observed DR over-representation in Actinobacteriota and Firmicutes as biologically meaningful rather

than a purely sampling artifact, while reinforcing the need for phylogeny-aware comparisons in downstream analyses.

3.2 Overall pI patterns

We began with the a priori expectation that proteins most directly exposed to the environment would show the clearest signature of drought response. On that basis we repeated all comparisons at two levels: whole proteomes and the extracellular subset. Intracellular-only analyses were also performed but are deferred to the Supplementary Material because they closely resembled the whole proteome results and did not add additional signal.

Considering all proteins, pI differs significantly between DR and DS bacteria (Fig. 2A, $p = 1.01 \times 10^{-32}$) having a similar result for intracellular proteins (supplementary material Fig 1), this difference is seen as well when restricting to extracellular proteins (Fig. 2B, $p = 2.56 \times 10^{-12}$).

The density plots clarify where these differences arise. For whole proteomes (Fig. 2C), DR and DS largely overlap and both show a multimodal structure. The main acidic mode for both groups lies near $pI \approx 5.2$, where most proteins concentrate. In the basic range the order flips: DS shows a higher mode around $pI \approx 9.5$ than DR, indicating a relative enrichment of DS proteins at higher pI values.

For extracellular proteins (Fig. 2D), the dominant modes shift toward the basic side for both groups, and the tall acidic peak seen in Fig. 2C is reduced. The acidic modes are no longer aligned: DR higher peak falls near $pI \approx 5$, whereas the highest acidic mode for DS is closer to $pI \approx 6$. This offset is consistent with DR having a more acidic distribution than DS. The separation in the basic range is also stronger than in the whole-proteome plot, reinforcing that

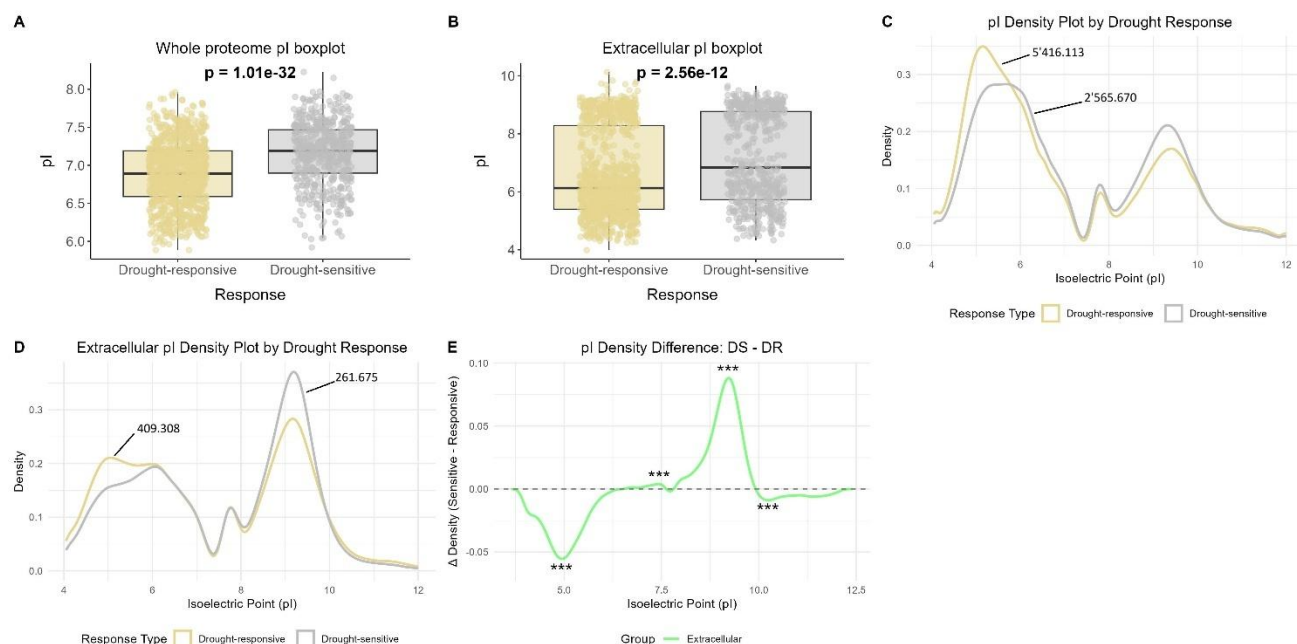


Figure 2 pI patterns for extracellular proteins and whole proteomes. * $p < 0.05$; ** $p < 0.001$, *** $p < 0.0001$.

extracellular proteins may accentuate the groups difference. To assess which peaks are driving this contrast between DS and DR, Fig. 2E shows the density difference Δ Density = Sensitive – Responsive for the extracellular subset. Here DS is significantly depleted around pI 4.5–5.5, significantly enriched around pI 7–8; 8.5–9.5, and depleted again in 9.5–10.5. Together, these patterns support that DR and DS differ in pI distributions.

Although DR and DS differ significantly in pI at the whole-proteome level and also for extracellular proteins, the dataset is not perfectly balanced across different taxa. Several phyla are more heavily represented than others, and some show a skew toward one response type. For example, Actinobacteriota includes many more DR than DS genomes, which could amplify group differences simply due to taxa composition. To address this potential phylogenetic bias, the next analyses stratify comparisons by phylum and then pair sensitive genomes with their closest responsive relatives within the same clade.

3.3 Phylogenetic relationship with isoelectric point

To address the phylogenetic bias noted above, we stratified the analysis by phylum. Panels A and B from figure 3 display pI distributions for the extracellular subset and for whole proteomes, respectively. In the extracellular fraction (A), the separation between response groups varies by lineage: DR medians exceed DS in Actinobacteriota, Chloroflexi, Bacteroidota, Acidobacteriota, and Firmicutes, whereas DS medians exceed DR in Planctomycetota, Myxococcota, Verrucomicrobiota, Gemmatimonadota, and

Proteobacteria. In several phyla the extracellular distributions are wider than their whole-proteome counterparts, indicating greater dispersion for secreted proteins. For whole proteomes (Fig. 3B), differences are generally smaller, and many phyla show DS with slightly higher medians than DR, consistent with the modest global shift reported earlier.

Panels C and D from figure 3 summarize these patterns with the median ratio of Sensitive to Responsive pI by phylum. For whole proteomes (Fig. 3C), the ratios cluster above 1 across most phyla, and a one-sample Wilcoxon test against 1 ($H_1 > 1$) is significant ($p = 0.031$), indicating a consistent tendency for DS to have higher pI than DR when averaged at the proteome level. In contrast, for extracellular proteins (Fig. 3D) the ratios scatter around 1 with mixed directions across phyla and the corresponding test (one sample on sided t-test) is not significant ($p = 0.323$) for $H_1 > 1$. This attenuation of the extracellular signal once phyla are separated supports the presence of a phylogenetic imprint on the global contrasts.

Panel E in Fig. 3 quantifies this imprint by relating pairwise pI distances to phylogenetic distances for all proteome pairs in the dataset. pI differences increase with evolutionary distance, as shown by the positive Mantel correlation ($r = 0.264$, $p = 0.001$). Together, these results indicate that much of the group-level separation is structured by taxonomic composition.

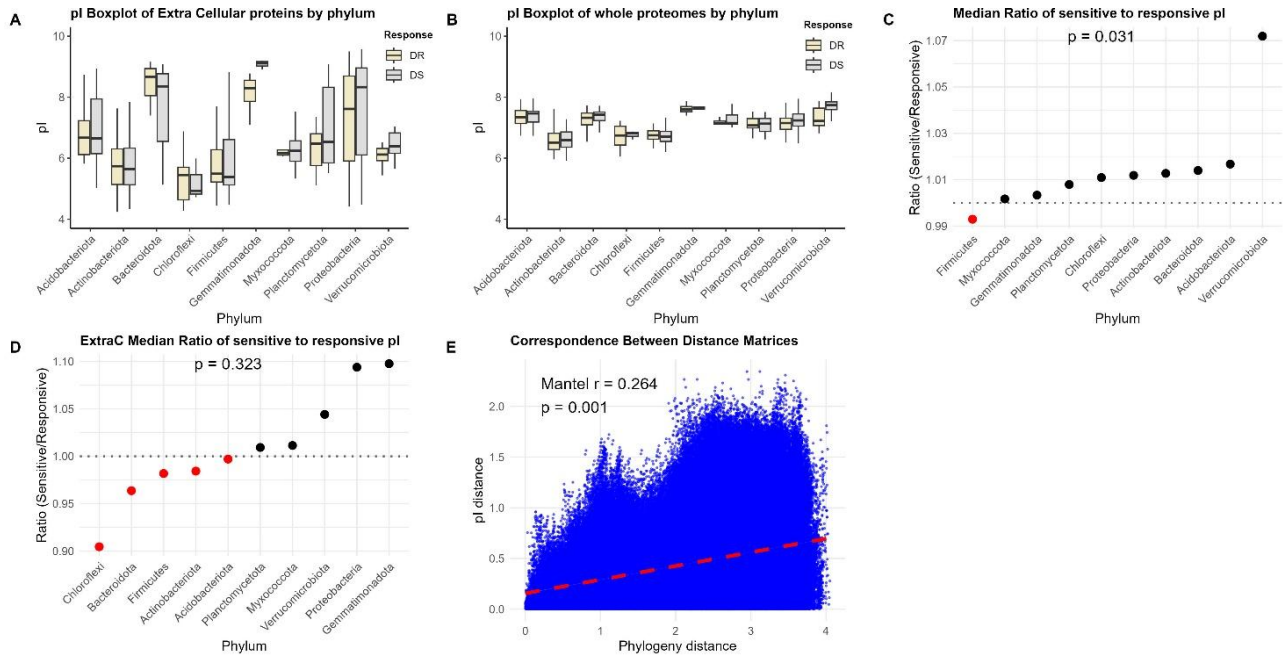


Figure 3 pI patterns stratified by phylum. * p adj < 0.05; ** p adj < 0.001, *** p adj < 0.0001.

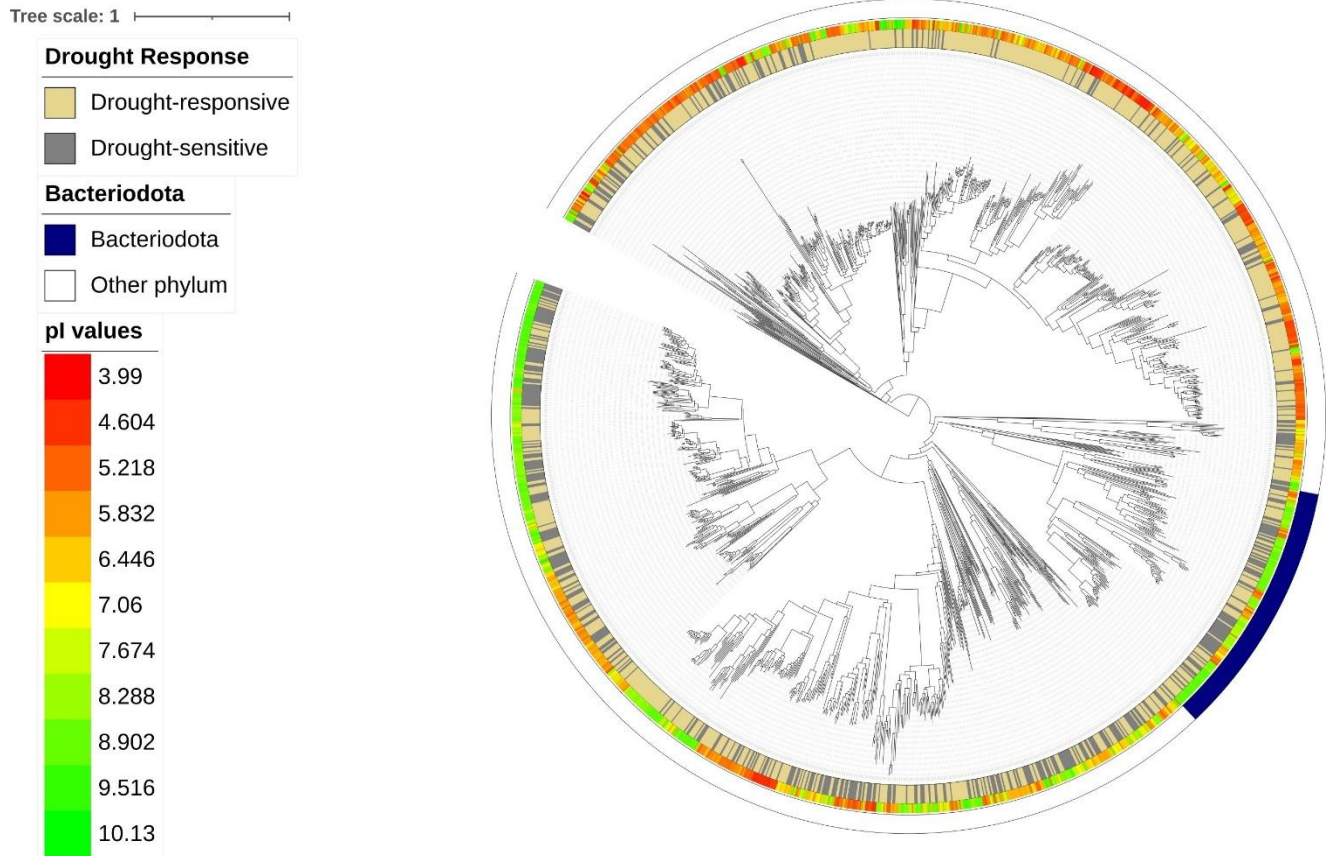


Figure 4 Phylogenetic tree with extracellular proteins pI and Drought response for every proteome

In the phylum-stratified boxplots, the across-phyla variation in pI is much more pronounced for the extracellular subset (Fig. 3A) than for whole proteomes (Fig. 3B). Even

when DR and DS are similar within a given phylum, the offsets between phyla are larger and the spread is wider for extracellular proteins. This means the overall DR–DS

contrast can be shaped strongly by which phyla are most represented, rather than by response type alone. To make this lineage effect explicit, we mapped only extracellular pI values onto a phylogenetic tree (Fig. 4), where stretches of similar color within clades highlight the phylogenetic imprint on extracellular proteome chemistry.

In the ITOL phylogenetic tree, contiguous segments of the outer ring display nearly uniform color, indicating similar extracellular pI values within clades. These blocks of color coincide with phylogeny rather than appearing randomly, consistent with a strong lineage effect on extracellular proteome chemistry. This observation, together with the within-phylum median pattern, motivated a more detailed exploration in order to minimize as much as possible the phylogenetic imprint observed in Figures 3E and 4. We therefore paired each DS genome with its closest DR

relative within the same phylum to test whether the differences in chemical properties were more remarkable.

3.4 Closest-relative ratios clarify extracellular pI differences

To account for the phylogenetic bias noted above, we paired each DS genome with its closest DR relative (i.e. the genome with minimal phylogenetic distance) within the same phylum and computed the pI ratio = Sensitive/Responsive. In the scatterplots, the y-axis is this ratio (values > 1 indicate higher pI in the sensitive proteome) and the x-axis are the phylums (Figure 5).

The overall trend is that sensitive extracellular proteomes have higher pI than their closest responsive relatives, and this is seen consistently across most phyla in panel A. By contrast, the whole-proteome ratios in panel B are less clear: several phyla, including Proteobacteria,

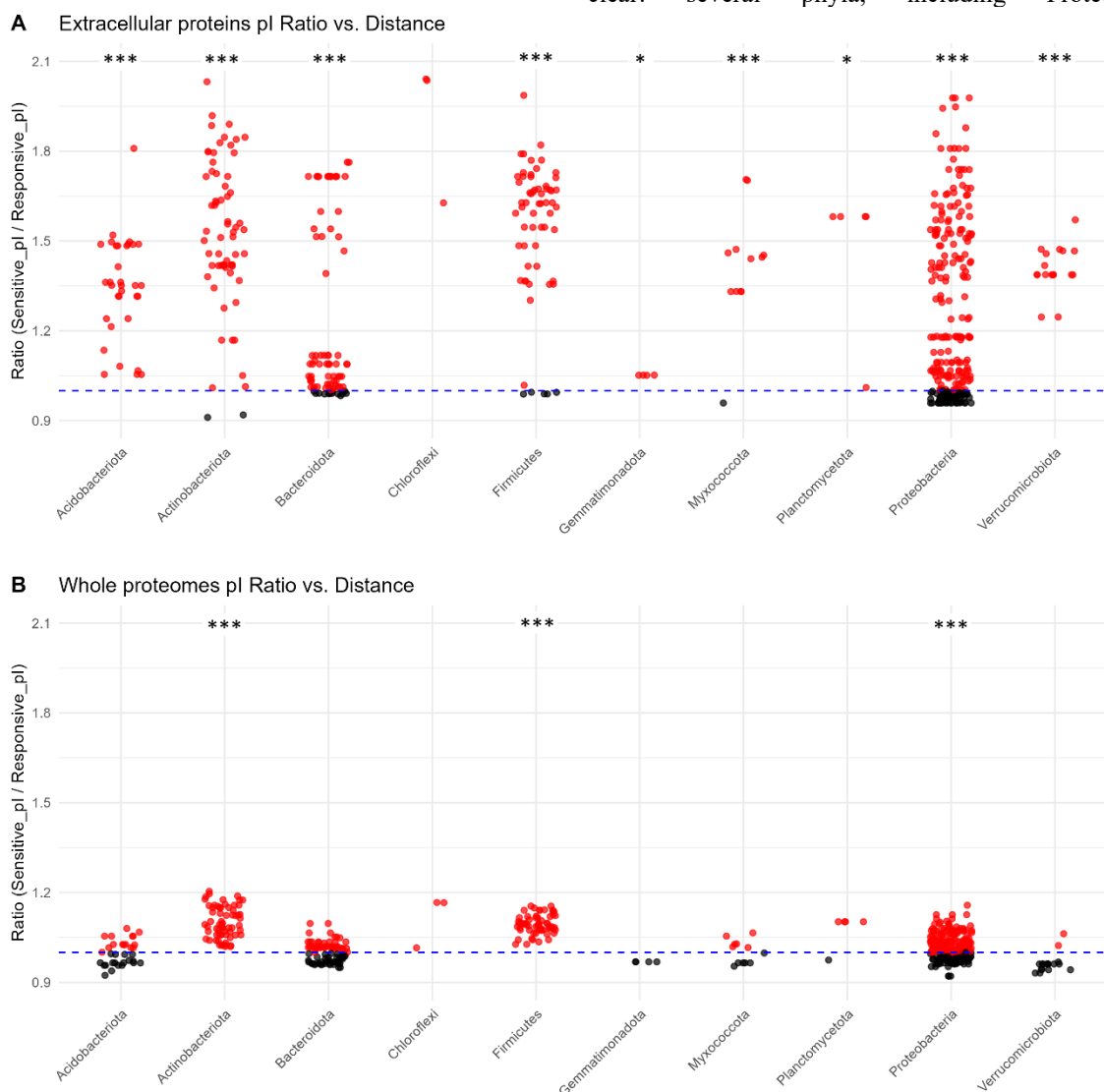


Figure 5 Closest relative ratio plot. Note the horizontal bands that appear in some phyla. These are expected given that multiple genomes can share the same nearest opposite-response neighbor, and closely related neighbors often have very similar pI (see Bacteroidota phylum in Fig. 4)

Bacteroidota, Acidobacteriota, and Myxococcota, cluster close to 1 with mixed directions, and in Gemmatimonadota and Verrucomicrobiota the tendency even reverses in some pairs (Figure 5). This contrast between panels A and B indicates that the signal is strongest and most consistent in the extracellular fraction, whereas inclusion of intracellular proteins masks it.

4 Discussion and conclusion

In this research we observed that DR bacterial proteomes tend to have a more acidic pI than those that are DS. Nevertheless, a strong phylogenetic bias was always present during the analyses, where similar taxa present similar values of pI. We showed that once this phylogenetic imprint was reduced, clear differences in the extracellular pI of drought-responsive and drought sensitive bacterias within the same phylum emerge. This suggests that proteome pI is not just a by-product of ancestry, but a trait that is linked to drought adaptation.

DR bacteria were found to have a more acidic extracellular proteome than DS taxa and several environmental features

during drought provide a probable reason for this pattern. Soil drying induces profound physicochemical changes that directly challenge protein stability [37]. This process increases ionic strength by several folds while simultaneously lowering water activity, conditions that are known to destabilize proteins through multiple mechanisms. These include the screening of electrostatic interactions, promotion of complex dissociation, and enhanced aggregation as molecular crowding intensifies [37]. Furthermore, pH changes during soil drying reduce this value but it remains circa 6-7, which stays above the isoelectric point (pI) of acidic extracellular proteins, thereby preserving their negative net charge throughout the stress period. Our results therefore align with extensive evidence that shows how negatively charged, low-pI proteins maintain solubility, hydration shells, and structural stability in high ionic strength environments [38-39].

The presence of a strong phylogenetic signal in pI seen in this study raises further questions. Why do entire phyla show consistent biases in proteome acidity? One possible explanation is that major bacterial groups have evolved in different ecological niches with varying baseline physicochemical conditions. As a result, their proteomes may have developed broad charge biases [40]. Actinobacteriota, for instance, have extracellular proteomes that are consistently more acidic compared to other phyla. This suggests that ecological specializations at the lineage level can leave lasting marks on proteome chemistry, which then interact with stress responses like drought.

The nature of the dataset has some limitations, first the imbalance presented in the majority of taxa may affect the comparisons made, furthermore, in silico predicted proteomes were used rather than experimental measurements, this in turn does not have into account post-translational modifications or context-dependent folding.

Looking forward, this investigation opens a promising line of research. Experimental validation of proteome charge effects under controlled desiccation would help test whether acidic proteomes indeed confer higher resilience. In addition to this, metaproteomics experiments over natural environments could reveal bacterial communities pI and its relationship with moisture.

References

- [1] IPCC, 2021: Summary for Policymakers. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 3–32, doi:10.1017/9781009157896.001.
- [2] Berg, A. and Sheffield, J. (2018) 'Climate change and drought: The Soil Moisture Perspective', *Current Climate Change Reports*, 4(2), pp. 180–191. doi:10.1007/s40641-018-0095-0.
- [3] Gebrechorkos, S.H. et al. (2025) 'Warming accelerates global drought severity', *Nature*, 642(8068), pp. 628–635. doi:10.1038/s41586-025-09047-2.
- [4] Deng, L. et al. (2021) 'Drought effects on soil carbon and nitrogen dynamics in global natural ecosystems', *Earth-Science Reviews*, 214, p. 103501. doi:10.1016/j.earscirev.2020.103501.
- [5] Pang, F. et al. (2024) 'Soil phosphorus transformation and plant uptake driven by phosphate-solubilizing microorganisms', *Frontiers in Microbiology*, 15. doi:10.3389/fmicb.2024.1383813.
- [6] Costa, O.Y., Raaijmakers, J.M. and Kuramae, E.E. (2018) 'Microbial extracellular polymeric substances: Ecological function and impact on soil aggregation', *Frontiers in Microbiology*, 9. doi:10.3389/fmicb.2018.01636.
- [7] Spooren, J. et al. (2024) 'Plant-driven assembly of disease-suppressive soil microbiomes', *Annual Review of Phytopathology*, 62(1), pp. 1–30. doi:10.1146/annurev-phyto-021622-100127.
- [8] Meisner, A. and de Boer, W. (2018) 'Strategies to maintain natural biocontrol of soil-borne crop diseases during severe drought and rainfall events', *Frontiers in Microbiology*, 9. doi:10.3389/fmicb.2018.02279.
- [9] Gillespie, L.M. et al. (2024) 'Drought effects on soil greenhouse gas fluxes in a boreal and a temperate forest', *Biogeochemistry*, 167(2), pp. 155–175. doi:10.1007/s10533-024-01126-2.
- [10] Canarini, A. et al. (2021) 'Ecological memory of recurrent drought modifies soil processes via changes in soil microbial community', *Nature Communications*, 12(1). doi:10.1038/s41467-021-25675-4.
- [11] Leyrer, V. et al. (2025) 'Drought impacts on plant–soil carbon allocation—integrating future mean climatic conditions', *Global Change Biology*, 31(2). doi:10.1111/gcb.70070.
- [12] De Silva, S., Kariyawasam Hetti Gamage, L. and Thapa, V.R. (2025) 'Impact of drought on soil microbial communities', *Microorganisms*, 13(7), p. 1625. doi:10.3390/microorganisms13071625.
- [13] Nicholson, W.L. et al. (2000) 'Resistance of *Bacillus* endospores to extreme terrestrial and extraterrestrial environments', *Microbiology and Molecular Biology Reviews*, 64(3), pp. 548–572. doi:10.1128/mmbr.64.3.548-572.2000.
- [14] Csonka, L.N. (1989) 'Physiological and genetic responses of bacteria to osmotic stress', *Microbiological Reviews*, 53(1), pp. 121–147. doi:10.1128/mr.53.1.121-147.1989.
- [15] Dick, J.M. and Tan, J. (2022) 'Chemical links between redox conditions and estimated community proteomes from 16S rRNA and reference protein sequences', *Microbial Ecology*, 85(4), pp. 1338–1355. doi:10.1007/s00248-022-01988-9.
- [16] Cabello-Yeves, P.J. and Rodriguez-Valera, F. (2019) 'Marine-freshwater prokaryotic transitions require extensive changes in the predicted proteome',

- Microbiome, 7(1). doi:10.1186/s40168-019-0731-5.
- [17] Bellissent-Funel, M.-C. et al. (2016) 'Water determines the structure and dynamics of proteins', *Chemical Reviews*, 116(13), pp. 7673–7697. doi:10.1021/acs.chemrev.5b00664.
- [18] Nakagawa, H. and Tamada, T. (2021) 'Hydration and its hydrogen bonding state on a protein surface in the crystalline state as revealed by molecular dynamics simulation', *Frontiers in Chemistry*, 9. doi:10.3389/fchem.2021.738077.
- [19] Lautenbach, V., Hosseinpour, S. and Peukert, W. (2021) 'Isoelectric point of proteins at hydrophobic interfaces', *Frontiers in Chemistry*, 9. doi:10.3389/fchem.2021.712978.
- [20] DasSarma, S. and DasSarma, P. (2015) 'Halophiles and their enzymes: Negativity put to good use', *Current Opinion in Microbiology*, 25, pp. 120–126. doi:10.1016/j.mib.2015.05.009.
- [21] Genome taxonomy database (2025) GTDB. Available at: <https://gtdb.ecogenomic.org/> (Accessed: 10 September 2025).
- [22] GTDB Data (2025) GTDB Data - /releases/release220/220.0/genomic_files_reps/. Available at: https://data.gtdb.ecogenomic.org/releases/release220/220.0/genomic_files_reps/2utm (Accessed: 10 September 2025).
- [23] Teufel, F. et al. (2022) 'SIGNALP 6.0 predicts all five types of signal peptides using protein language models', *Nature Biotechnology*, 40(7), pp. 1023–1025. doi:10.1038/s41587-021-01156-3.
- [24] Driessen, A.J.M. and Nouwen, N. (2008) 'Protein translocation across the bacterial cytoplasmic membrane', *Annual Review of Biochemistry*, 77(1), pp. 643–667. doi:10.1146/annurev.biochem.77.061606.160747.
- [25] Kostakioti, M. et al. (2005) 'Mechanisms of protein export across the bacterial outer membrane', *Journal of Bacteriology*, 187(13), pp. 4306–4314. doi:10.1128/jb.187.13.4306-4314.2005.
- [26] Okuda, S. and Tokuda, H. (2011) 'Lipoprotein sorting in bacteria', *Annual Review of Microbiology*, 65(1), pp. 239–259. doi:10.1146/annurev-micro-090110-102859.
- [27] Giltner, C.L., Nguyen, Y. and Burrows, L.L. (2012) 'Type IV pilin proteins: Versatile molecular modules', *Microbiology and Molecular Biology Reviews*, 76(4), pp. 740–772. doi:10.1128/mmb.00035-12.
- [28] Dick, J. (2024) Canprot: Chemical Analysis of Proteins, The Comprehensive R Archive Network. Available at: <https://cran.r-project.org/web/packages/canprot/index.html> (Accessed: 10 September 2025).
- [29] Bjellqvist, B. et al. (1993) 'The focusing positions of polypeptides in immobilized ph gradients can be predicted from their amino acid sequences', *ELECTROPHORESIS*, 14(1), pp. 1023–1031. doi:10.1002/elps.11501401163.
- [30] Bio.SeqUtils.IsoelectricPoint module¶ (no date) Bio.SeqUtils.IsoelectricPoint module - Biopython 1.76 documentation. Available at: <https://biopython.org/docs/1.76/api/Bio.SeqUtils.IsoelectricPoint.html> (Accessed: 10 September 2025).
- [31] Dick, J.M. (2014) 'Average oxidation state of carbon in proteins', *Journal of The Royal Society Interface*, 11(100), p. 20131095. doi:10.1098/rsif.2013.1095.
- [32] Dick, J. (2024a) Canprot: Chemical Analysis of Proteins, The Comprehensive R Archive Network. Available at: <https://cran.r-project.org/web/packages/canprot/index.html> (Accessed: 10 September 2025).
- [33] Letunic, I. and Bork, P. (2024) 'Interactive tree of life (itol) V6: Recent updates to the phylogenetic tree display and annotation tool', *Nucleic Acids Research*, 52(W1). doi:10.1093/nar/gkae268.
- [34] Shannon, C.E. (1948) 'A Mathematical Theory of Communication', *The Bell System Technical Journal*, 27, pp. 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x.
- [35] Naylor, D. and Coleman-Derr, D. (2018) 'Drought stress and root-associated bacterial communities', *Frontiers in Plant Science*, 8. doi:10.3389/fpls.2017.02223.
- [36] Pérez Castro, S. et al. (2019) 'Soil microbial responses to drought and exotic plants shift carbon metabolism', *The ISME Journal*, 13(7), pp. 1776–1787. doi:10.1038/s41396-019-0389-9.
- [37] Romero-Perez, P.S. et al. (2023) 'When phased without water: Biophysics of cellular desiccation, from biomolecules to condensates', *Chemical Reviews*, 123(14), pp. 9010–9035. doi:10.1021/acs.chemrev.2c00659.
- [38] Kiraga, J. et al. (2007) 'The relationships between the isoelectric point and: Length of proteins, taxonomy and ecology of organisms', *BMC Genomics*, 8(1). doi:10.1186/1471-2164-8-163.
- [39] Romero-Pérez, P.S. et al. (2025) Protein surface chemistry encodes an adaptive tolerance to desiccation [Preprint]. doi:10.2139/ssrn.5119488.
- [40] Smole, Z. et al. (2011) 'Proteome sequence features carry signatures of the environmental niche of prokaryotes', *BMC Evolutionary Biology*, 11(1). doi:10.1186/1471-2148-11-26.