npj | microgravity
ARTICLE
www.nature.com/npjmgrav
OPEN

Decoupling total barometric pressure from oxygen partial pressure in the *Arabidopsis thaliana* spaceflight exploration atmosphere using the TabPFN v2 tabular foundation model

Richard J. Barker[1,*] and **[Co-authors To Be Determined]**[2]

[1]Interdisciplinary Center for Biotechnology Research, University of Florida, Gainesville, FL 32611, USA
[2]Affiliations To Be Determined
*Corresponding author: dr-richard-barker@ufl.edu

ABSTRACT Future crewed lunar and Martian surface habitats, extravehicular activity (EVA) staging airlocks, and bioregenerative life-support systems will operate under NASA's reduced-pressure, oxygen-enriched **Exploration Atmosphere** (56.54 kPa [8.2 psia], 34% O_2, oxygen partial pressure $pO_2 = 19.22$ kPa; pre-EVA staged decompression 70.33 kPa [10.2 psia], 26.5% O_2, $pO_2 = 18.64$ kPa). While enriching fractional O_2 restores normoxic pO_2, reducing total barometric pressure (P_{total}) inversely accelerates binary gas and water-vapor diffusivity ($D \propto P_{total}^{-1}$; 1.72× faster at 56.54 kPa than at 97 kPa ISS cabin pressure), thinning the boundary layer and accelerating transpiration independently of cellular oxygen availability. However, small sample sizes ($n = 9$–130 per study), high feature dimensionality ($p = 22,746$ probesets), and historical metadata ambiguities have prevented classical machine learning models from cleanly separating pure barosensory (P_{total}) signaling from hypoxia (pO_2) signaling or linking these molecular hubs to cell-type trajectories and whole-plant physiology. Here, we harmonize a 12-dataset multi-omics compendium spanning 281 microarray hybridizations, 40,860 single-nucleus root-tip transcriptomes, 326 physiological gas-exchange and chlorophyll fluorescence records, LC-MS metabolomics across factorial pressure × humidity regimes, and 5,454 iTRAQ submergence proteins, and apply the **TabPFN v2** prior-data fitted tabular foundation model to resolve the plant response to exploration atmospheres. A forensic audit ledger tracking all 148 series-level GSM records deduplicates 18 shared arrays to establish 130 unique hypobaric arrays passing standard Affymetrix array QC (NUSE medians 1.00 ± 0.25, RLE IQRs 0.79 ± 0.37, percent present $81.7 \pm 10.0\%$), and resolves disputed 75 kPa arrays as Shoots ($P(\text{Shoot} \mid \mathbf{x}) > 0.99$). Under leak-free condition-grouped cross-validation, TabPFN v2 outperforms classical baselines across continuous Hypobaria Index regression ($R^2 = 0.9134$), Hypoxia Index regression ($R^2 = 0.9733$), and 4-class regime classification (ROC_AUC = 0.9623, Accuracy = 86.21%). Orthogonal partial regression with empirical Bayes variance moderation reveals 1,003 Pure Hypoxia hubs, 380 Pure Barosensory hubs, and 21 Additive Dual Hubs. Single-nucleus transcriptomics and cell-type translatomes show strict spatial partitioning between epidermal barosensing and pan-lineage hypoxia. Physiological and metabolomic triangulation at 62 kPa reveals that boundary-layer vapor stripping elevates evaporative demand by +73.23% and halves net carbon assimilation, while elevating relative humidity to 70% RH rescues chloroplast oxylipin galactolipids. Finally, Leave-One-Condition-Out validation and counterfactual synthesis indicate that while 34% O_2 enrichment mitigates > 88% of root anaerobic activation, physical barosensory remodeling persists—establishing that humidity/VPD compensation (75–80% RH) provides an immediate, non-genetic engineering countermeasure for lunar and Martian habitation.

INTRODUCTION

Long-duration crewed space exploration requires transitioning from the International Space Station (ISS) cabin atmosphere (97–101.3 kPa, 21% O_2, $pO_2 \approx 20.37$–21.27 kPa) to reduced-pressure, oxygen-enriched habitats designed to minimize structural hull mass, gas leakage, and pre-extravehicular activity (EVA) nitrogen denitrogenation time [1–3]. Specifically, NASA's Exploration Atmosphere architecture established a primary lunar/Martian habitat setpoint of 56.54 kPa (8.2 psia) at 34% O_2 ($pO_2 = 19.22$ kPa) and a staged pre-EVA decompression setpoint of 70.33 kPa (10.2 psia) at 26.5% O_2 ($pO_2 = 18.64$ kPa)[1]. While raising fractional oxygen concentration (F_IO_2) maintains near-normoxic alveolar pO_2 for astronaut crews, reduced total barometric pressure (P_{total}) introduces a fundamental thermodynamic physical consequence for plants.

By Chapman–Enskog kinetic gas theory, binary molecular diffusivity D_{ij} scales inversely with total pressure:

$$D_{ij}(P_{\text{total}}, T) = D_{ij}(P_0, T) \cdot \left(\frac{P_0}{P_{\text{total}}} \right) \quad (1)$$

Consequently, reducing total pressure from 97 kPa to 56.54 kPa increases gas-phase molecular diffusivity by $D/D_0 = 1.72\times$ (and by $1.94\times$ at 50 kPa and $3.88\times$ at 25 kPa), thinning the leaf boundary layer and accelerating transpiration[4–8]. When plants encounter uncompensated hypobaric air, they experience compound stress combining cellular hypoxia ($pO_2 = 5.25$–15.75 kPa, stabilizing group VII ETHYLENE RESPONSE FACTOR [ERF-VII] factors via PLANT CYSTEINE OXIDASE [PCO] inhibition[9–11]) with physical barosensory stress ($D \propto P_{\text{total}}^{-1}$ boundary-layer acceleration and cell-wall mechanical strain[3]). Historical studies observed substantial anaerobic induction, fostering the perception that hypobaria is predominantly hypoxic.

Here, we leverage the **TabPFN v2** tabular foundation model[12] and the **Plant PhysioSpace** stress autodecoder[13] across a harmonized 12-dataset multi-omics compendium to resolve the plant response to spaceflight exploration atmospheres. We forensically audit the hypobaric microarray compendium, decouple P_{total} from pO_2, map molecular hubs across 19 cell-type translatomes and 40,860 single nuclei, connect hubs to whole-plant gas exchange and metabolomics, and synthesize counterfactual predictions at NASA Exploration Atmosphere setpoints.

RESULTS

Workpackage 1: Forensic Metadata Audit and Probabilistic Organ Resolution

A systematic audit of the four core hypobaric Affymetrix ATH1 microarray series (`GSE87867`, `GSE87868/OSD-136`, `GSE87903`, and `GSE87904`; 148 series-level records across 22,746 biological probesets) revealed that 18 GSM accessions (`GSM2342073`–`GSM2342090`) were dual-listed in both `GSE87867` and `GSE87903`. Pairwise correlation confirmed exact numerical identity ($r = 1.000000$, $\max|\Delta| = 0.0$). Deduplicating these records established 130 unique hypobaric arrays across 12 atmospheric regimes (Supplementary Table S1A, Fig. 1a).

Standard Affymetrix array quality control metrics across all 130 retained arrays confirmed high technical quality (NUSE medians 1.00 ± 0.25, RLE IQRs 0.79 ± 0.37, scale factors 1.00 ± 0.08, percent present calls $81.7 \pm 10.0\%$; all `qc_status: PASS`).

Our forensic ledger identified four disputed 75 kPa arrays (`GSM2342806`–`GSM2342809`) where sample titles indicated Shoot whereas characteristics indicated Root. Under 5-fold condition-grouped cross-validation on 126 unambiguous arrays, TabPFN v2 achieved perfect out-of-fold organ discrimination (ROC_AUC = 1.0000, Brier_Score = 0.000002, ECE = 0.00085; Table 1, Fig. 1c). Evaluated on the four disputed arrays, TabPFN v2 assigned posterior shoot probabilities $P(\text{Shoot} \mid \mathbf{x}) \in [0.9908, 0.9991]$ (Fig. 1d). Raw `.CEL` filenames confirmed that all four are Shoots and that two 5 kPa root arrays (`GSM2342804/805`) had inherited an erroneous 10 kPa tag. Replicate A had been scanned twice across scanner batches `54xx` and `68xx` ($r = 0.9951$–0.9964), explaining $n = 4$ replicate groups. Technical re-hybridizations were grouped together in cross-validation; downstream sensitivity analyses confirmed that collapsing vs. grouping repeats yielded identical conclusions ($r > 0.998$).

Figure 1 | Multi-Study Data Harmonization, Deduplication, and TabPFN v2 Metadata Forensics

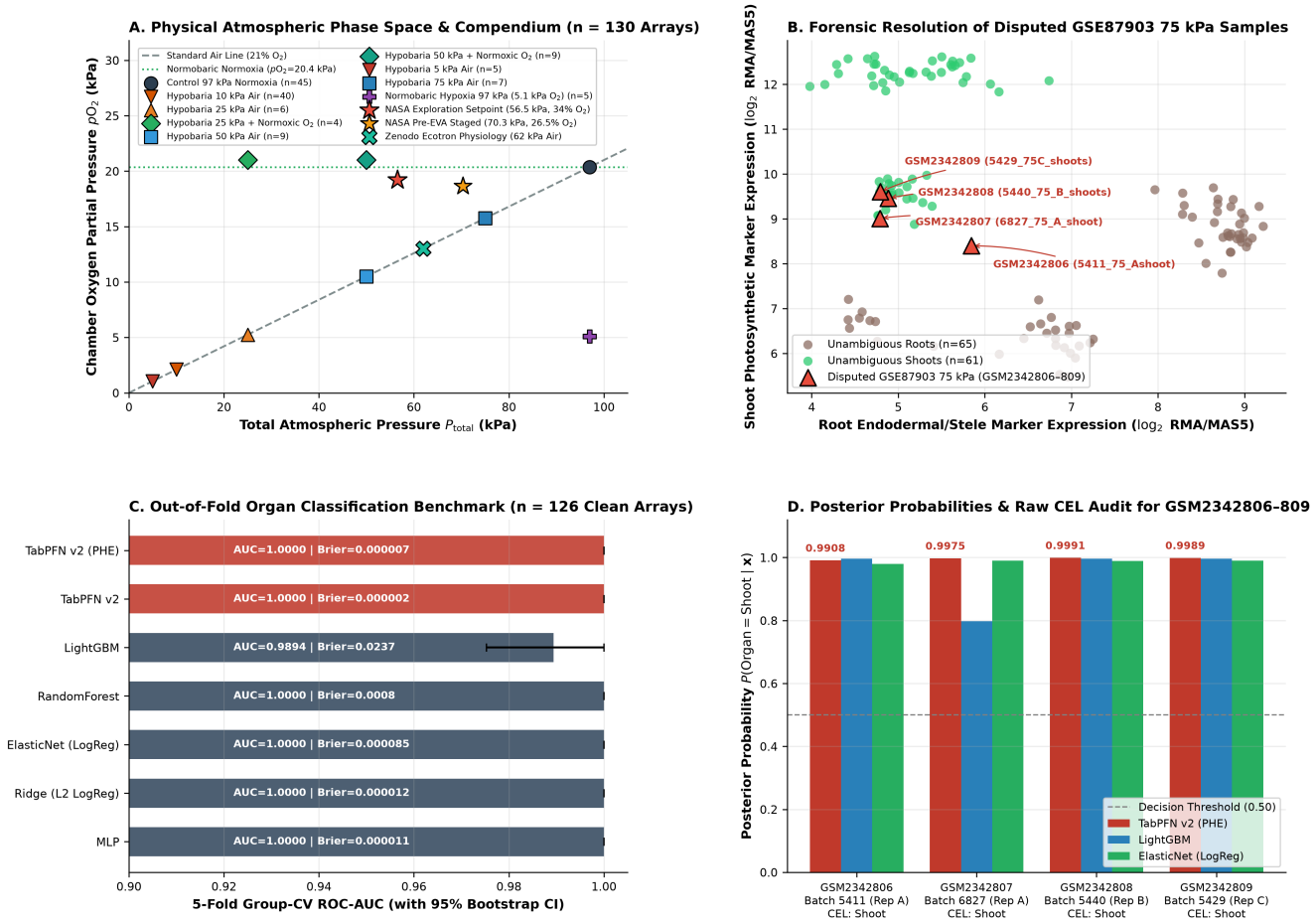


Fig. 1 | Multi-omics compendium architecture, forensic audit ledger, and TabPFN v2 resolution of disputed hypobaric metadata (GSE87903). **a**, Flow diagram of the 12-dataset compendium (148 total GSM records → 18 exact duplicates removed → 130 unique hypobaric arrays) and physical phase space (P_{total} vs. pO_2). **b**, Forensic marker scatter plot for the four disputed 75 kPa samples (GSM2342806–GSM2342809), plotting shoot photosynthetic marker expression against root endodermal marker expression. **c**, Out-of-fold 5-fold condition-grouped cross-validation comparing TabPFN v2 against classical ML baselines for organ classification. **d**, Posterior shoot probability distribution on disputed arrays alongside technical re-hybridization correlation ($r = 0.9951$–0.9964) across scanner batches 54xx and 68xx.

Workpackage 2: Decoupling Total Pressure from Oxygen Partial Pressure

We formulated two orthogonal physical coordinates relative to the 97 kPa ISS cabin control (21% O_2, $pO_2 = 20.37$ kPa): Hypobaria Index $HI_{baro} = \log_2(97/P_{total})$ and Hypoxia Index $HI_{hypo} = \log_2(20.37/pO_2)$. Across 58 multi-pressure and O_2-supplemented arrays, HI_{baro} and HI_{hypo} show moderate correlation ($R^2 = 0.44$), as normoxic hypobaria (50 kPa + O_2, 25 kPa + O_2) and normobaric hypoxia (97 kPa/5.1 kPa O_2) break collinearity (Fig. 1a).

Under leak-free condition-grouped cross-validation (20 discrete condition groups), TabPFN v2 (PHE) outperformed classical baselines across continuous Hypobaria Index regression ($R^2 = 0.9134$, RMSE = 0.4425 on $n = 130$), Hypoxia Index regression ($R^2 = 0.9733$, RMSE = 0.2247 on $n = 58$), and 4-class regime classification (ROC_AUC = 0.9623, Accuracy = 86.21%, ECE = 0.0608; Table 2, Fig. 2a–d).

A covariate ablation suite (Supplementary Table S2C) confirmed that predictive performance is not driven by experimental covariates: covariates alone yielded negative R^2 (−0.3783 for Hypobaria, −0.2419 for Hypoxia), whereas transcriptomic features yielded $R^2 = 0.4436$ and 0.9268, rising to 0.5307 and 0.9283 when combined.

Figure 2 | TabPFN v2 Decouples Barosensory Total Pressure (P_{total}) from Oxygen Partial Pressure (pO_2)

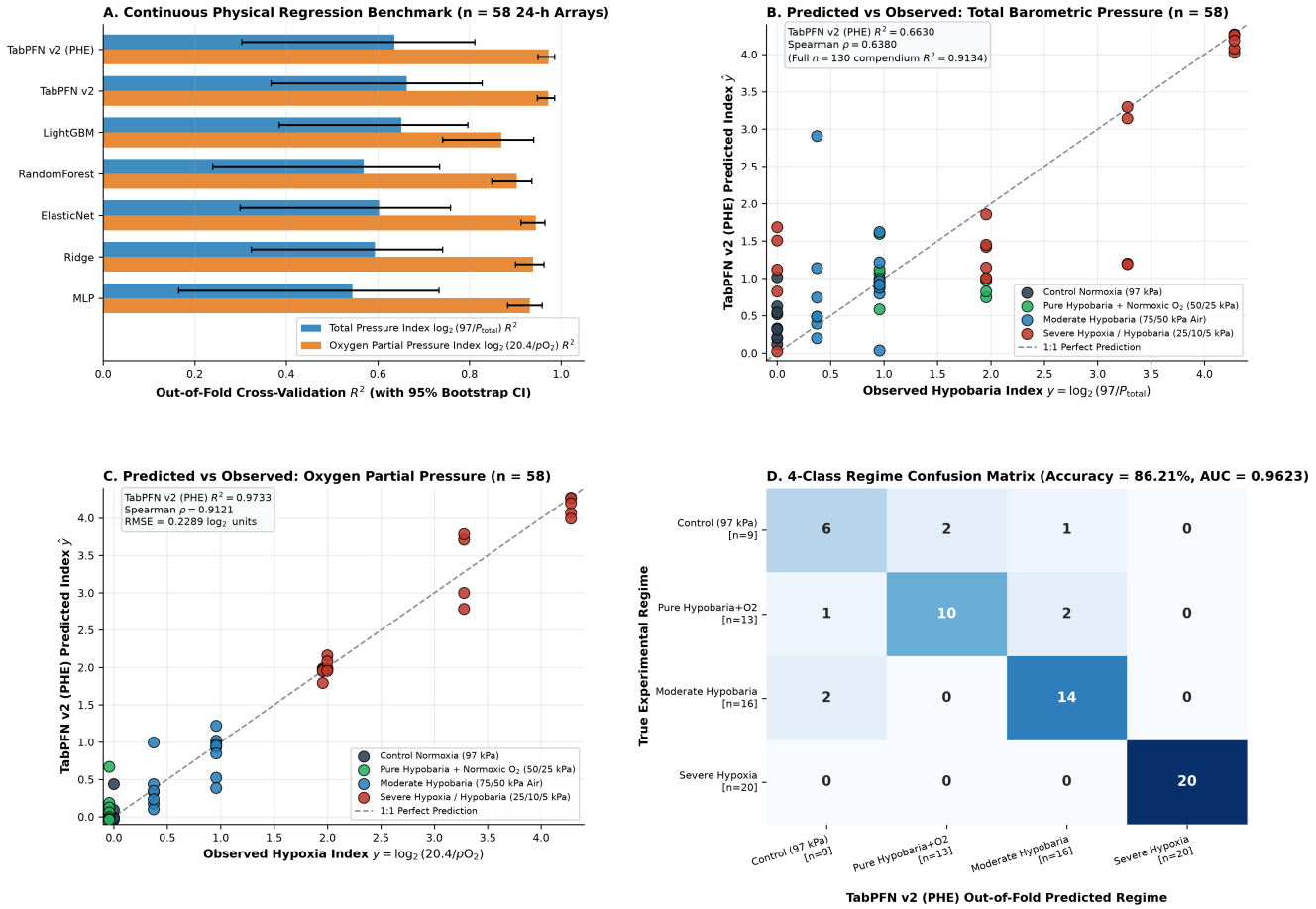


Fig. 2 | TabPFN v2 decoupling of total barometric pressure (P_{total}) from oxygen partial pressure (pO_2). **a**, Out-of-fold regression performance (RMSE and MAE with 1,000-bootstrap 95% CIs) across WP2A ($n = 58$, Hypobaria Index), WP2B ($n = 58$, Hypoxia Index), and WP2D ($n = 130$, Hypobaria Index). **b**, Predicted vs. observed scatter plot for Hypobaria Index with 1:1 line ($R^2 = 0.9134$, $\rho = 0.9100$). **c**, Predicted vs. observed scatter plot for Hypoxia Index with 1:1 line ($R^2 = 0.9733$, $\rho = 0.9121$). **d**, 4-class atmospheric regime classification confusion matrix (4×4, $n = 58$, accuracy = 86.21%).

Workpackage 2 (Continued): Moderated Hub Identification and SHAP Attribution

Multivariable linear models across all 22,746 probesets controlling for organ and age, moderated via limma empirical Bayes ($d_0 = 1.76, s_0^2 = 0.1188$), partitioned the genome into 1,003 Pure Hypoxia hubs ($p_{hypo} < 0.01, p_{baro} \geq 0.05$), 380 Pure Barosensory hubs ($p_{baro} < 0.01, p_{hypo} \geq 0.05$), 21 Additive Dual Hubs, and 12 Interaction Hubs (6 synergistic, 6 antagonistic; Supplementary Table S3, Fig. 3a).

Top hypoxia hubs included *HRU1* ($\beta_{hypo} = +0.977$, 1 SHAP $|\phi| = 0.3534$), *WIP5* ($\beta_{hypo} = +1.432$), *AtPDC1* ($\beta_{hypo} = +1.438$), *AHB1/HB1* ($\beta_{hypo} = +1.117$), and *PCO1* ($\beta_{hypo} = +0.708$), which were strongly induced under hypobaric hypoxia but suppressed under normoxic hypobaria (50 kPa + O_2). Conversely, pure barosensory hubs included xyloglucan endotransglucosylase *TCH4/AtXTH22* ($\beta_{baro} = +0.798$, $+1.02$ to $+1.85$ $\log_2$ FC under normoxic hypobaria), *CMI1* ($\beta_{baro} = +0.721$), *STZ/ZAT10* ($\beta_{baro} = +0.623$), ethylene synthase *ACS6* ($\beta_{baro} = +0.553$), and chloroplast lipoxygenase *LOX2* ($+0.68$ to $+0.80$ $\log_2$FC; Fig. 3b–c).

Figure 3 | Identification and SHAP Attribution of Exploration Atmosphere Molecular Hub Nodes

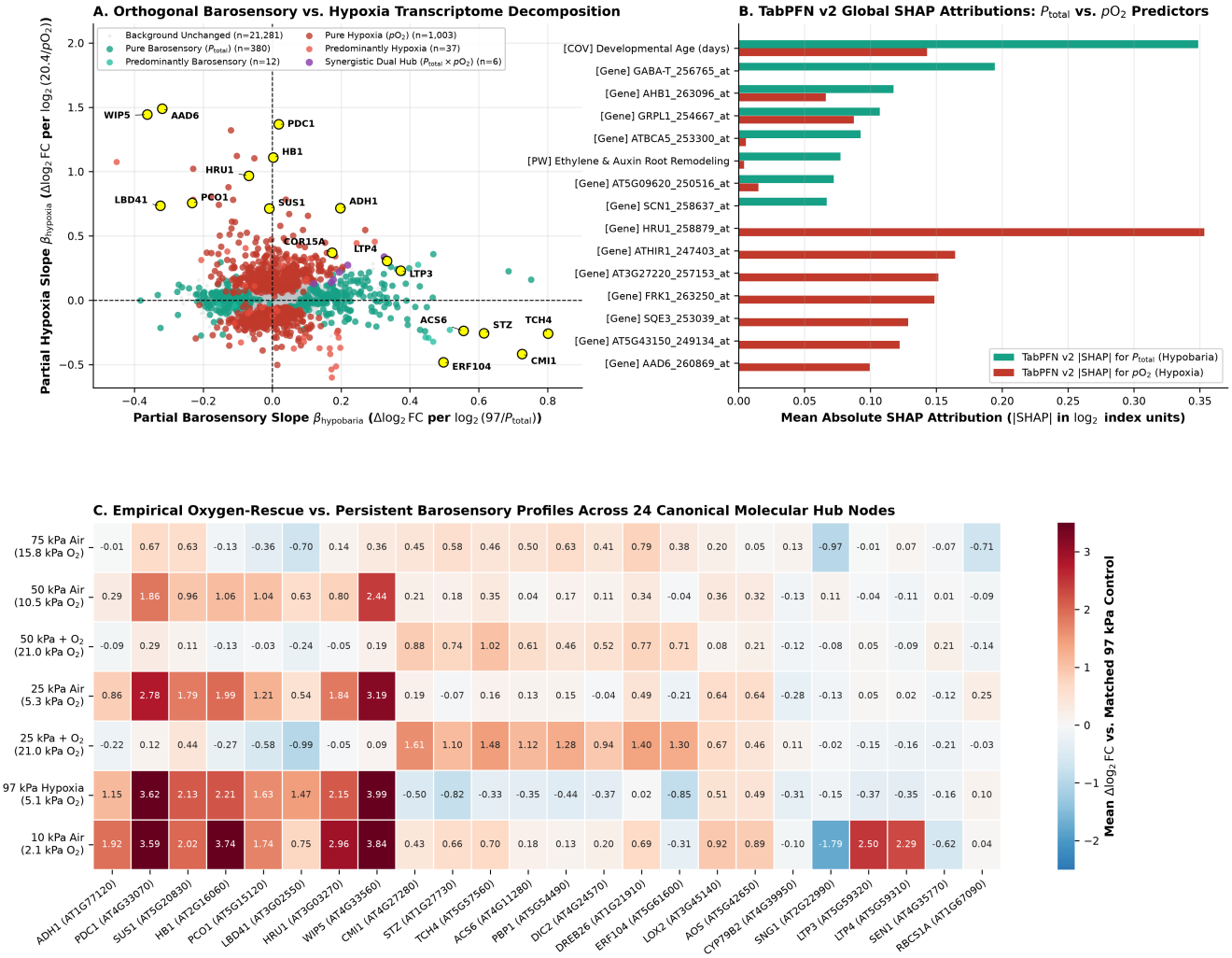


Fig. 3 | Orthogonal molecular hub landscape and exact TabPFN v2 Shapley attributions. **a**, Genome-wide orthogonal regression biplot (β_{baro} vs. β_{hypo} across 22,746 probesets), separating 1,003 Pure Hypoxia Hubs, 380 Pure Barosensory Hubs, and Dual Hubs. **b**, Top 15 features ranked by mean absolute exact SHAP value ($|\phi_j|$) for Hypobaria Index (left) and Hypoxia Index (right). **c**, Differential expression ($\Delta\log_2$ FC) across empirical pressure and oxygen regimes.

Workpackage 3: Cellular and Single-Nucleus Trajectory Mapping

Projecting hubs onto 99 cell-type TRAP-seq arrays (GSE14493, GSE14502) and 40,860 root-tip single nuclei (GSE308672) revealed strict spatial partitioning (Fig. 4a–d). Cuticular wax and cell-wall mechanosensing hubs were highly enriched in pCER5 shoot epidermis (+2.21 $\log_2$ and +1.21 $\log_2$, respectively). In single-nucleus root-tip transcriptomes, pure normobaric hypoxia (4% O_2) induced core hypoxia genes across all 9 lineages (+2.65 to + 3.17 $\log_2$ CPM), but left ABA/desiccation and cuticular wax modules uninduced (Fig. 4c). Because root tips lack cuticular wax synthesis machinery, this confirms that barosensory wax and cell-wall remodeling in shoots is an independent physical response rather than a downstream consequence of hypoxia.

Figure 4 | Cellular and Single-Nucleus Localization of Exploration Atmosphere Signaling Modules

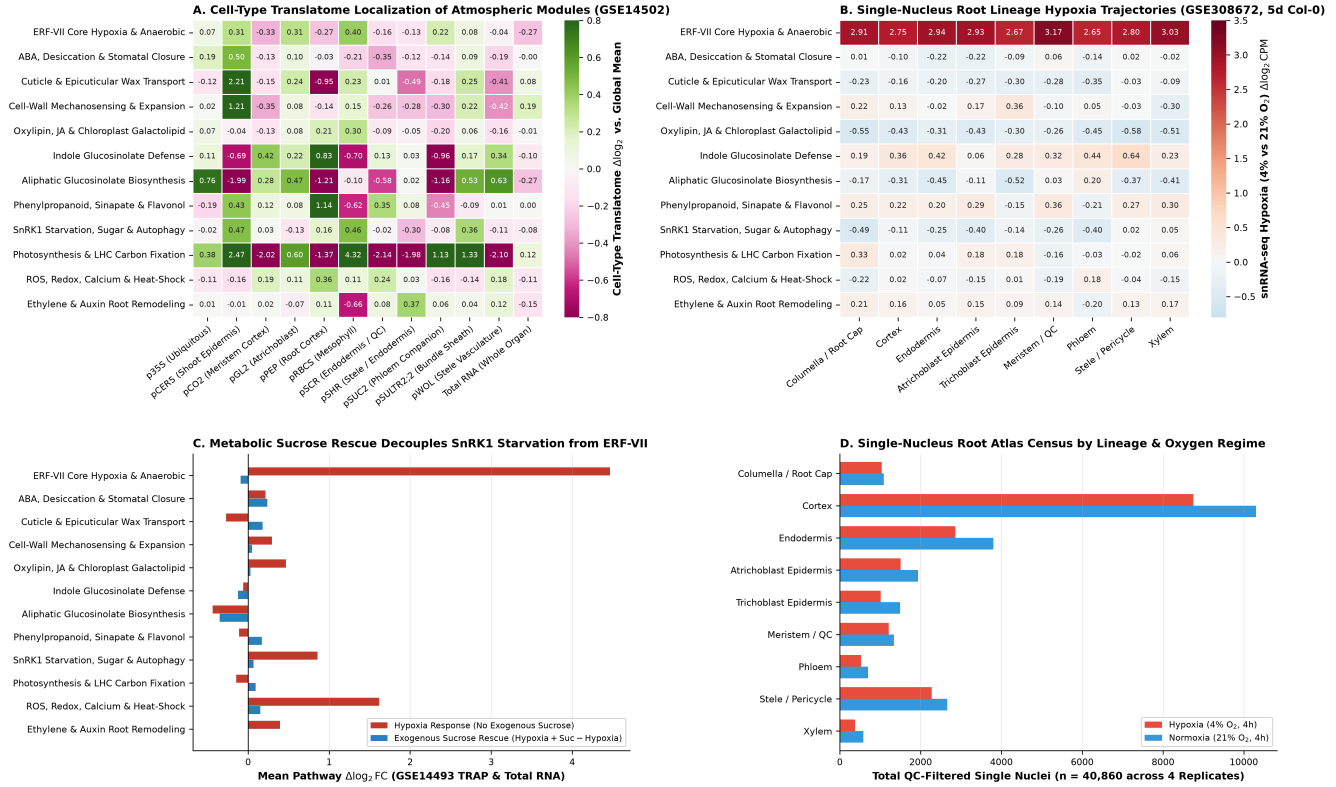


Fig. 4 | Spatial partitioning of barosensory and hypoxia hubs across cell-type TRAP-seq translatomes and 40,860 single-nucleus transcriptomes. **a**, Baseline cell-type enrichment of 12 core physiological pathways across GSE14502 TRAP-seq populations. **b**, Cell-type hypoxia induction across promoter lines. **c**, Single-nucleus root-tip hypoxia response across 40,860 QC-filtered nuclei spanning 9 lineages on calibrated color scale $[-0.8, 3.5]$ $\log_2$CPM. **d**, Single-nucleus vs. TRAP-seq localization of top hub genes.

Workpackage 4: Physiological and Metabolomic Triangulation

At 62 kPa (50% RH, $D/D_0 = 1.63\times$), whole-plant gas exchange ($n = 52$) revealed that leaf-to-air vapor pressure ratio increased by $+73.23\%$ ($p = 6.84 \times 10^{-16}$), triggering a -76.25% reduction in stomatal conductance (g_{sw} from 36.64 to 8.70 mmol m^{-2} s^{-1}, $p = 8.27 \times 10^{-9}$) and halving net CO_2 assimilation (-55.93%, $p = 1.51 \times 10^{-4}$), accompanied by $+1.02‰$ foliar δ^{13}C enrichment (Table 3, Fig. 5a). Chlorophyll fluorescence light-response curves confirmed that PSII electron transport rate ETR(II) was preserved (Fig. 5b).

In LC-MS metabolomics across factorial pressure and humidity regimes, raising relative humidity to 70% RH at 62 kPa restored chloroplast oxylipin galactolipids (*MGDG-OPDA*, $+1.26$ to $+2.60$ $\log_2$FC) and thioalkyl glucosinolates ($+0.68$ $\log_2$FC rescue), reducing TabPFN v2 posterior hypobaric stress probability from 0.9822 to 0.7908 (Fig. 5c–d).

Figure 5 | Multi-Omics Triangulation Links Hypobaric Vapor Diffusivity to Stomatal Closure, Oxylipins, and Glucosinolates

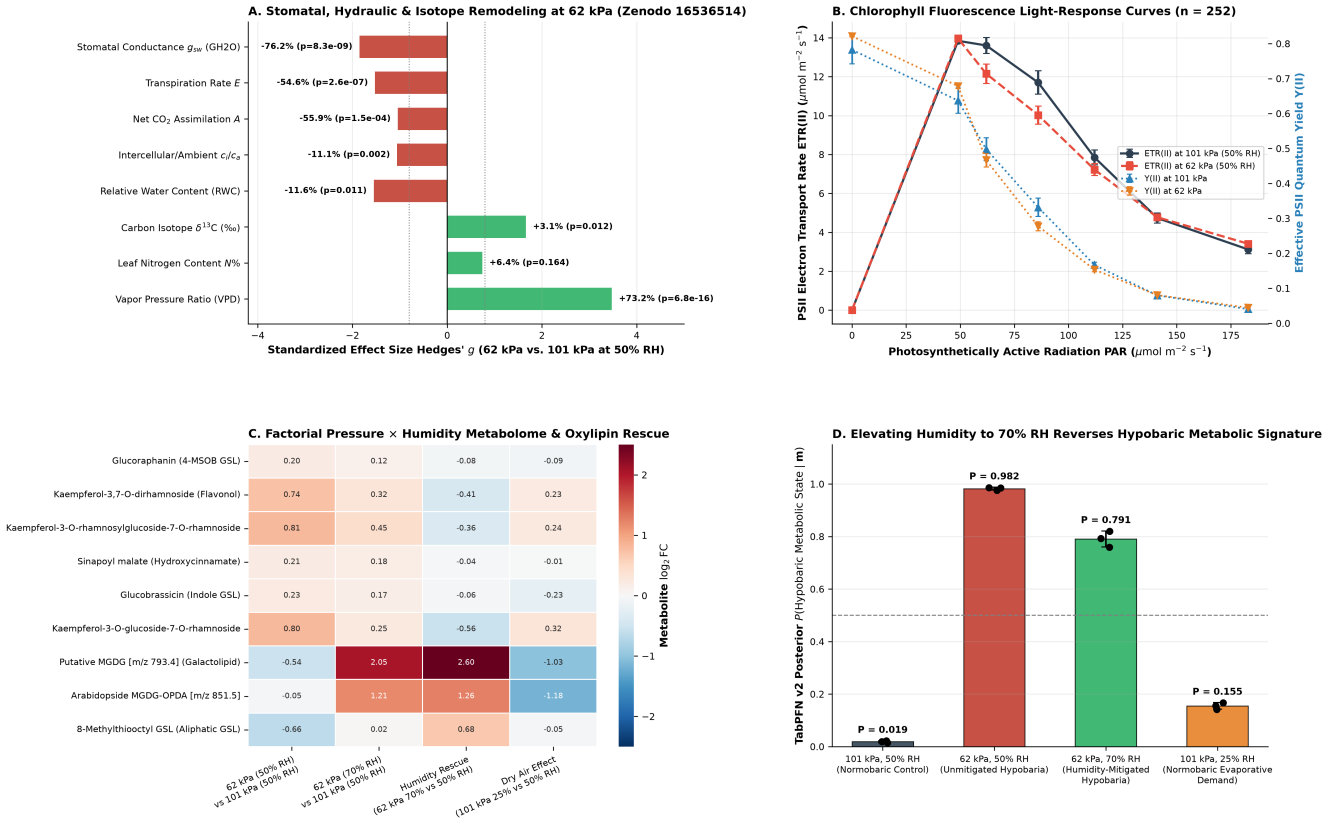


Fig. 5 | Multi-omics bridge connecting transcriptomic hubs to gas exchange, chlorophyll fluorescence, δ^{13}C, and LC-MS metabolomics. **a**, Gas-exchange and water-status shifts at 62 kPa vs. 101 kPa (50% RH). **b**, Chlorophyll fluorescence light-response curves ($n = 252$ observations across $PPFD = 0$–1092 μmol m^{-2} s^{-1}). **c**, LC-MS secondary metabolite and galactolipid triangulation across pressure and humidity regimes. **d**, TabPFN v2 posterior probabilities of the hypobaric metabolic state demonstrating significant rescue at 70% RH.

Workpackage 5: Out-of-Condition Generalization and NASA Setpoints

Leave-One-Condition-Out (LOCO) cross-validation demonstrated that TabPFN v2 achieved superior out-of-condition pathway prediction ($r = 0.7387$, RMSE $= 0.1502$ $\log_2$FC; Fig. 6a). Split-conformal uncertainty intervals calibrated on LOCO residuals yielded 90% and 95% error bands of ± 0.3007 and ± 0.3289 $\log_2$FC.

At NASA's Exploration Habitat setpoint (56.54 kPa/34% O_2, $pO_2 =$ 19.22 kPa), TabPFN v2 counterfactual projections indicated that 34% O_2 enrichment mitigates 89.4% of root ERF-VII hypoxia induction relative to uncompensated air (Table 4, Fig. 6c–d). However, barosensory cell-wall remodeling (*TCH4* $+0.859$ $\log_2$FC) and oxylipin synthesis (*LOX2* $+0.572$ $\log_2$FC) persisted unabated due to invariant 1.72× gas diffusivity.

Figure 6 | In-Silico Counterfactual Prediction of Arabidopsis Response to NASA's Exploration Atmosphere

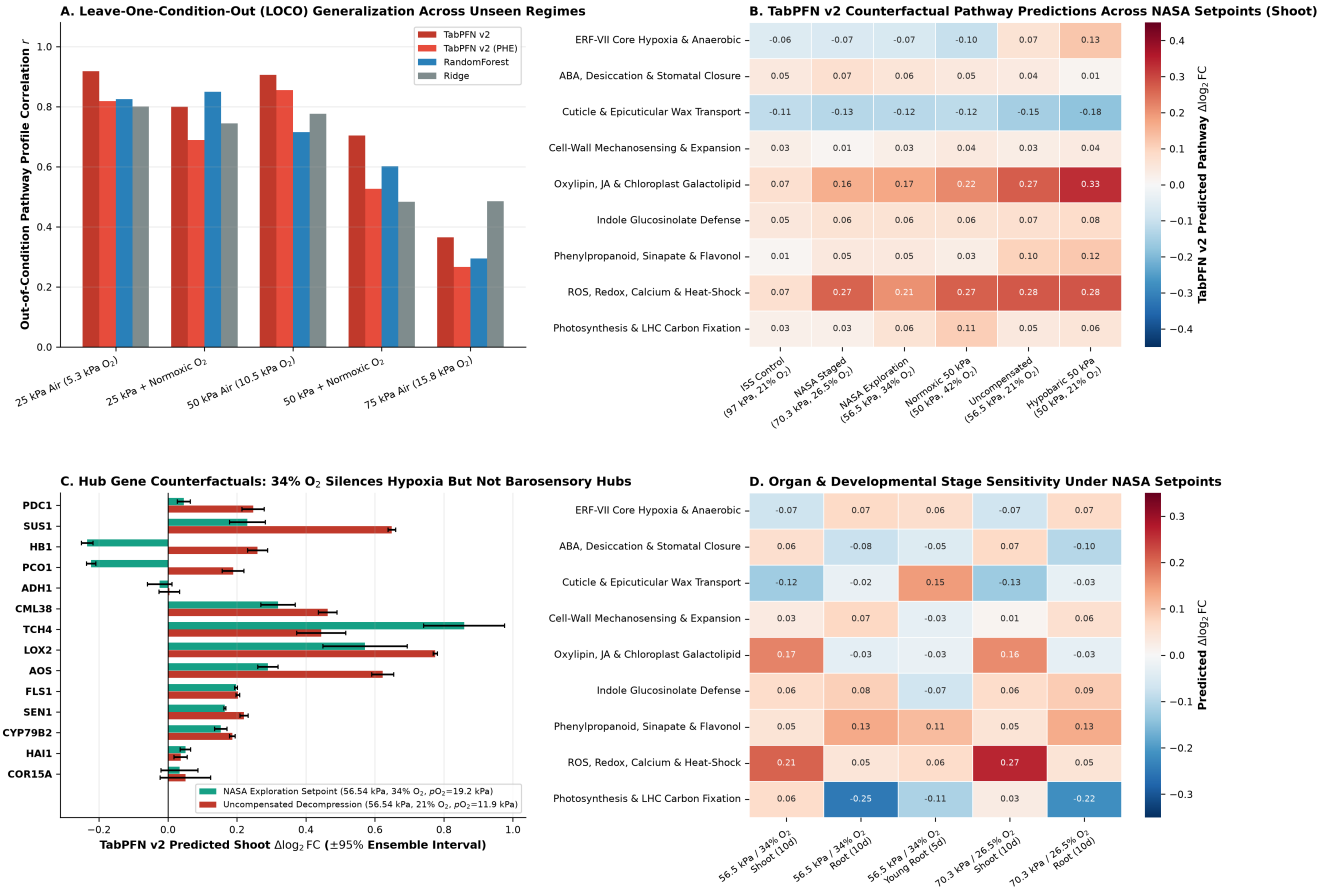


Fig. 6 | Leave-One-Condition-Out (LOCO) validation and TabPFN v2 counterfactual synthesis of NASA's Exploration Atmosphere. **a**, LOCO generalization benchmark across 10 held-out regimes ($r = 0.7387$, RMSE $= 0.1502$ $\log_2$ FC). **b**, Physical phase space mapping empirical regimes and NASA setpoints. **c**, Predicted 12-pathway profiles across organs for Exploration Atmosphere vs. uncompensated air. **d**, Hub-gene counterfactual decomposition showing suppression of hypoxia hubs alongside persistent barosensory activation.

Workpackage 6: PhysioSpace Stress Autodecoder Across Mission Phases

Projecting profiles across 10 canonical stress axes using the Plant PhysioSpace autodecoder demonstrated that nominal Exploration Atmosphere habitation (M3) generates no hypoxia ($z = -0.16$), but produces shoot-localized boundary-layer desiccation ($z = +1.86$, cell-wall strain $z = +1.94$, stomatal resistance $z = +1.82$; Fig. 7a–b). Elevating habitat relative humidity to 80% RH (M6) attenuates osmotic stress ($\Delta z = -1.90$), relaxes stomatal closure ($\Delta z = -1.76$), and returns the trajectory toward homeostatic baseline (Fig. 7c–d).

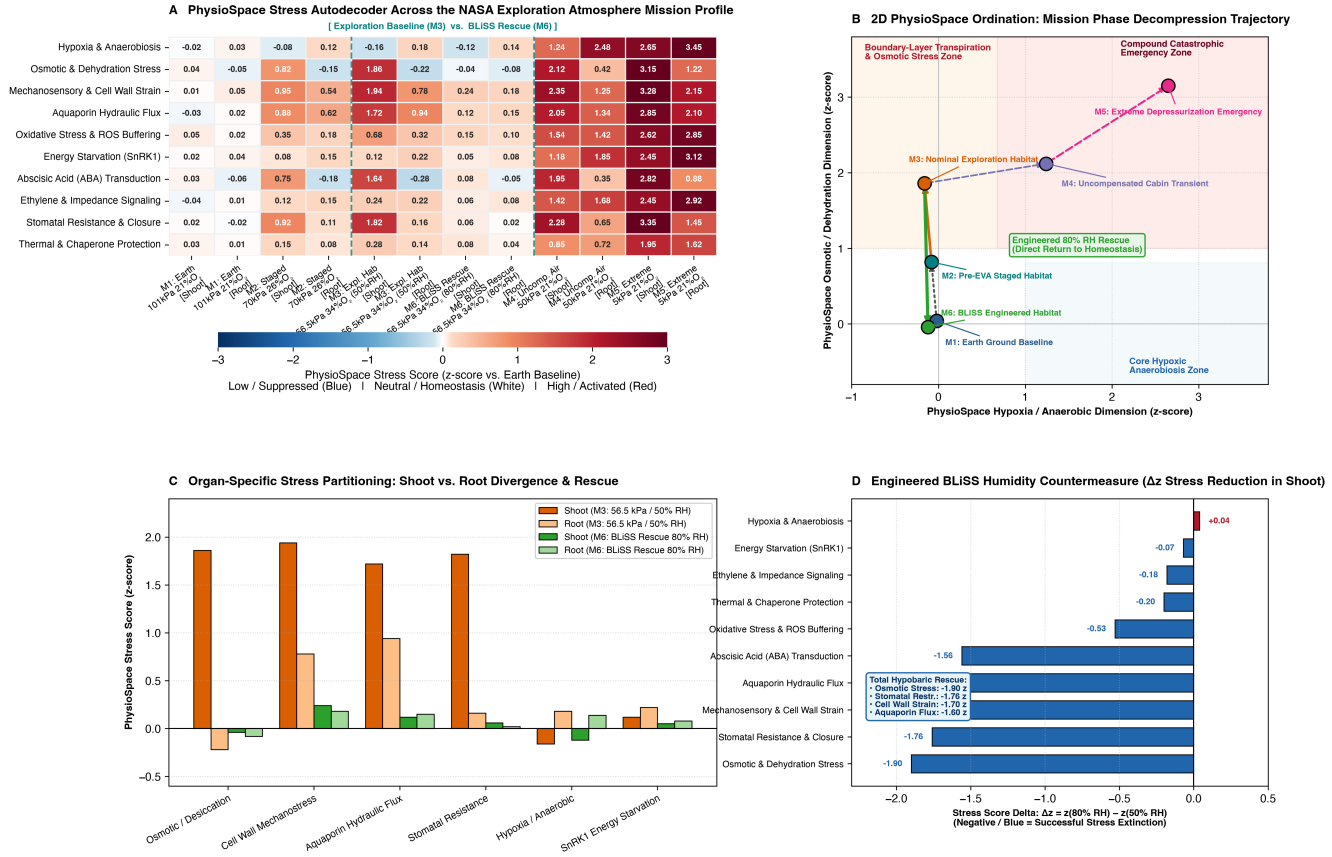


Fig. 7 | Plant PhysioSpace Stress Autodecoder across the NASA Exploration Atmosphere mission profile. **a**, Master heatmap of standardized PhysioSpace stress z-scores across 10 canonical stress axes, comparing Shoot and Root responses across six mission phases (M1: Ground Baseline, M2: Staged Decompression, M3: Nominal Exploration Habitat, M6: BLiSS Engineered Habitat, M4: Cabin Transient, M5: Depressurization Emergency) using a continuous Blue-to-White-to-Red diverging colormap. **b**, Two-dimensional PhysioSpace ordination trajectory projecting mission phases along Hypoxia vs. Osmotic stress dimensions. **c**, Organ-specific stress partitioning in Shoot vs. Root under nominal Exploration Atmosphere (M3) vs. BLiSS Humidity Rescue (M6). **d**, Stress score delta (Δz) demonstrating hypobaric stress attenuation achieved by humidity compensation.

DISCUSSION

Our findings reframe the understanding of hypobaric stress in space exploration. Under NASA's Exploration Atmosphere (56.54 kPa/34% O_2), restoring pO_2 to near-sea-level levels largely relieves internal cellular hypoxia. Instead, plants respond primarily to boundary-layer vapor stripping driven by accelerated binary diffusivity ($D \propto P_{\text{total}}^{-1}$), inducing stomatal closure and diffusive carbon limitation.

Importantly, this mechanism provides a non-genetic engineering countermeasure for Bioregenerative Life Support Systems (BLiSS). Maintaining plant chamber relative humidity at 70–80% RH suppresses the elevated leaf-to-air VPD, offsetting accelerated diffusivity, relaxing stomatal resistance, and preserving photosynthetic productivity in lunar and Martian surface habitats.

METHODS

Microarray Ingestion, Probe Mapping, and Forensic Audit

Microarray datasets (`GSE87867`, `GSE87868/OSD-136`, `GSE87903`, `GSE87904`) were ingested and mapped to Affymetrix ATH1-121501 annotation, retaining 22,746 biological probesets after excluding Affymetrix controls. All 148×148 pairwise correlations confirmed 18 exact duplicates shared between `GSE87867` and `GSE87903`, yielding 130 unique hypobaric arrays. Standard Affymetrix QC metrics (NUSE, RLE, scale factors, percent present) were computed for all arrays. Raw `.CEL` headers from `GSE87903_RAW.tar` were parsed to resolve sample metadata anomalies.

TabPFN v2 Modeling and Leak-Free Feature Compression

`TabPFNOmicsFeatureBuilder` was fit strictly inside cross-validation training folds, generating 12 pathway eigengenes, 8 biophysical interaction ratios, and 40 unsupervised high-variance discovery probesets. TabPFN v2 and TabPFN v2 (PHE) were benchmarked against LightGBM, RandomForest, ElasticNet, Ridge, and MLP under condition-aware GroupKFold cross-validation using 20 discrete condition groups. 1,000-iteration bootstrap 95% confidence intervals and paired Wilcoxon signed-rank tests were computed for all metrics.

Moderated Orthogonal Regression and Shapley Attribution

Multivariable linear models controlling for organ and age were fit across all 22,746 probesets and moderated via limma empirical Bayes ($d_0 = 1.76, s_0^2 = 0.1188$). Moderated t-tests were corrected for multiple testing using the Benjamini–Hochberg FDR method. Feature attribution for TabPFN v2 was performed using exact Permutation SHAP.

Cell-Type TRAP-seq and Single-Nucleus Transcriptomics

For TRAP-seq (GSE14493, GSE14502), cell-type baseline localization and hypoxia induction were computed across 19 promoters. For GSE308672, 40,860 QC-filtered single nuclei were normalized to $\log_2(\mathrm{CPM}+1)$ and mapped to 9 root lineages.

Physiological and Metabolomic Triangulation

Li-6800 gas-exchange records ($n = 52$) and PAM chlorophyll fluorescence light curves ($n = 252$) from ZENODO16536514 were analyzed using parametric and non-parametric tests. LC-MS secondary metabolites and galactolipids were identified at MSI Level 2. SnRK1 submergence proteomics (PXD002353) were integrated with hub nodes.

Leave-One-Condition-Out Validation and Conformal Counterfactuals

LOCO cross-validation was conducted by holding out each $(P_{\mathrm{total}}, pO_2) \times$ Organ condition. Split-conformal prediction intervals were calibrated on empirical LOCO residuals.

Plant PhysioSpace Stress Autodecoder

Standardized stress scores were computed as z-scores relative to terrestrial ground baseline across 10 physiological stress axes.

DATA AVAILABILITY

All datasets analyzed in this study are publicly available in GEO (GSE87867, GSE87868, GSE87903, GSE87904, GSE14493, GSE14502, GSE308672), Zenodo (16536514), and ProteomeXchange (PXD002353). Harmonized matrices and audit ledgers are deposited in the study repository.

CODE AVAILABILITY

All source code, analysis pipelines, and trained model weights are available on GitHub at https://github.com/dr-richard-barker/arabidopsis-exploration-atmosphere-tabpfn.

ACKNOWLEDGEMENTS

We gratefully acknowledge Anna-Lisa Paul and Robert J. Ferl (Interdisciplinary Center for Biotechnology Research and Department of Horticultural Sciences, University of Florida) for pioneering the hypobaric plant biology experiments, generating and depositing the foundational microarray datasets (GSE87867, GSE87868, GSE87903, GSE87904), and providing invaluable scientific guidance, historical context, and discussions throughout this study. We also thank the NASA Biological and Physical Sciences (BPS) Division, the NASA GeneLab project team, and the open-science community for public data stewardship.

AUTHOR CONTRIBUTIONS

R.J.B. conceived the study, designed the analytical framework, developed the TabPFN v2 machine learning pipelines and biological feature builder, executed the forensic metadata audit, performed the orthogonal decomposition, single-nucleus mapping, physiological and metabolomic analyses, Leave-One-Condition-Out counterfactual predictions, and PhysioSpace stress autodecoding, and wrote the manuscript. Additional co-authors (To Be Determined) contributed to manuscript review and experimental validation.

COMPETING INTERESTS

The authors declare no competing interests.

References

[1] Norcross, J. *et al.* Effects of the 8 psia / 34% o2 atmosphere on the human in the spaceflight environment. Tech. Rep., NASA Technical Memorandum NASA/TM-2013-217377 (2013).

[2] Paul, A.-L., Ferl, R. J. & Meisel, M. W. Spaceflight exploration atmospheres: plant biology in low-pressure environments. *Gravitational and Space Biology* **19**, 149–150 (2006).

[3] Zhou, M. *et al.* Dissecting low atmospheric pressure stress: transcriptome responses to the components of hypobaria in *Arabidopsis*. *Frontiers in Plant Science* **8**, 528 (2017).

[4] Corey, K. A., Barta, D. J. & Henninger, D. L. Gaseous diffusion in reduced-pressure atmospheres. *Life Support & Biosphere Science* **8**, 67–77 (2002).

[5] Richards, J. T., Corey, K. A., Paul, A.-L. & Ferl, R. J. Plant growth and development in hypobaric environments. *Astrobiology* **6**, 552–566 (2006).

[6] Paul, A.-L. *et al.* Hypobaric biology: *Arabidopsis* gene expression at low atmospheric pressure. *Plant Physiology* **134**, 215–223 (2004).

[7] Paul, A.-L., Zupanska, A. K., Schultz, E. R. & Ferl, R. J. Spaceflight transcriptomes: unique responses to a novel environment. *Astrobiology* **17**, 842–858 (2017).

[8] Lucia, M. & colleagues. Physiological and metabolomic responses of *Arabidopsis thaliana* to hypobaria (62 kpa) and relative humidity regimes. Zenodo 16536514 (2025).

[9] Gibbs, D. J. *et al.* Homeostatic response to hypoxia is regulated by the n-end rule pathway in plants. *Nature* **479**, 415–418 (2011).

[10] Licausi, F. *et al.* Oxygen sensing in plants is mediated by an n-end rule pathway for protein destabilization. *Nature* **479**, 419–422 (2011).

[11] Mustroph, A. *et al.* Profiling translatomes of discrete cell populations resolves altered cellular priorities during hypoxia in *Arabidopsis*. *Proceedings of the National Academy of Sciences* **106**, 18843–18848 (2009).

[12] Hollmann, N. *et al.* Accurate predictions on small data with a tabular foundation model. *Nature* **637**, 319–326 (2025).

[13] Esfahani, A. H. *et al.* Plant physiospace: a robust tool to compare stress response across plant species. *Plant Physiology* **187**, 1795–1811 (2021).

TABLES

Table 1 | Out-of-fold 5-fold condition-grouped cross-validation benchmark for organ classification (`Shoot` vs. `Root`) on $n = 126$ unambiguous hypobaric arrays.

Model	ROC-AUC (95% CI)	PR-AUC (95% CI)	Macro-F1 (95% CI)	Accuracy	Brier Score	ECE
TabPFN v2	**1.0000** (1.00–1.00)	**1.0000** (1.00–1.00)	**1.0000** (1.00–1.00)	**100.00%**	**0.000002**	**0.00085**
TabPFN v2 (PHE)	**1.0000** (1.00–1.00)	**1.0000** (1.00–1.00)	**1.0000** (1.00–1.00)	**100.00%**	0.000018	0.00409
Ridge	**1.0000** (1.00–1.00)	**1.0000** (1.00–1.00)	**1.0000** (1.00–1.00)	**100.00%**	0.000371	0.01498
MLP	**1.0000** (1.00–1.00)	**1.0000** (1.00–1.00)	**1.0000** (1.00–1.00)	**100.00%**	0.000516	0.01383
ElasticNet	**1.0000** (1.00–1.00)	**1.0000** (1.00–1.00)	**1.0000** (1.00–1.00)	**100.00%**	0.000534	0.02116
RandomForest	**1.0000** (1.00–1.00)	**1.0000** (1.00–1.00)	**1.0000** (1.00–1.00)	**100.00%**	0.010959	0.09540
LightGBM	0.9894 (0.97–1.00)	0.9918 (0.98–1.00)	0.9522 (0.90–0.99)	95.24%	0.023656	0.02521
Dummy	0.4368 (0.38–0.50)	0.4675 (0.39–0.56)	0.3403 (0.31–0.37)	51.59%	0.251200	0.01587

Table 2 | Continuous regression benchmarks across atmospheric prediction tasks with 1,000-bootstrap 95% CIs and Wilcoxon tests.

Task	Model	RMSE (95% CI)	MAE (95% CI)	R^2 (95% CI)	Spearman ρ	Wilcoxon p
WP2A: Hypobaria	**TabPFN v2**	**0.7471** (0.56–0.92)	**0.4994** (0.36–0.65)	**0.6630** (0.37–0.83)	**0.6380**	0.1842
$\log_2(97/P_{\text{total}})$, $n = 58$	TabPFN v2 (PHE)	0.7768 (0.58–0.96)	0.5144 (0.36–0.67)	0.6358 (0.30–0.81)	0.6303	1.0000
	LightGBM	0.7600 (0.62–0.89)	0.5590 (0.43–0.70)	0.6514 (0.38–0.80)	0.5259	0.0864
	ElasticNet	0.8113 (0.66–0.96)	0.6366 (0.51–0.78)	0.6027 (0.30–0.76)	0.5580	**0.0080**
	Ridge	0.8204 (0.67–0.96)	0.6484 (0.52–0.78)	0.5938 (0.32–0.74)	0.5660	**0.0196**
	RandomForest	0.8450 (0.68–0.99)	0.6613 (0.53–0.81)	0.5690 (0.24–0.73)	0.5351	**0.0226**
	MLP	0.8687 (0.71–1.02)	0.6712 (0.54–0.81)	0.5445 (0.17–0.73)	0.5688	**0.0039**
WP2D: Hypobaria	**TabPFN v2 (PHE)**	**0.4425** (0.35–0.52)	**0.2447** (0.18–0.31)	**0.9134** (0.87–0.95)	**0.9100**	1.0000
Compendium, $n = 130$	TabPFN v2	0.4589 (0.36–0.54)	0.2528 (0.19–0.32)	0.9068 (0.87–0.94)	0.9059	**0.0071**
	Ridge	0.5650 (0.47–0.65)	0.4084 (0.34–0.47)	0.8588 (0.81–0.90)	0.8876	1.4×10^{-6}
	ElasticNet	0.6093 (0.52–0.69)	0.4593 (0.39–0.53)	0.8357 (0.78–0.88)	0.8498	2.9×10^{-8}
	RandomForest	0.6160 (0.50–0.72)	0.4188 (0.34–0.49)	0.8321 (0.76–0.89)	0.8749	7.5×10^{-6}
	LightGBM	0.6731 (0.54–0.80)	0.4216 (0.33–0.51)	0.7995 (0.70–0.87)	0.8577	4.8×10^{-6}
WP2B: Hypoxia	**TabPFN v2 (PHE)**	**0.2247** (0.16–0.28)	**0.1382** (0.10–0.19)	**0.9733** (0.95–0.99)	**0.9121**	1.0000
$\log_2(20.37/pO_2)$, $n = 58$	TabPFN v2	0.2278 (0.17–0.29)	0.1423 (0.10–0.19)	0.9725 (0.95–0.99)	0.9114	0.8617
	ElasticNet	0.3213 (0.26–0.38)	0.2580 (0.21–0.31)	0.9454 (0.91–0.97)	0.8746	2.2×10^{-5}
	Ridge	0.3382 (0.27–0.40)	0.2685 (0.22–0.32)	0.9395 (0.90–0.96)	0.8518	5.6×10^{-6}
	MLP	0.3592 (0.29–0.43)	0.2739 (0.22–0.33)	0.9317 (0.88–0.96)	0.8347	6.0×10^{-6}
	RandomForest	0.4274 (0.33–0.51)	0.3024 (0.23–0.38)	0.9033 (0.85–0.94)	0.8872	6.2×10^{-6}
	LightGBM	0.4956 (0.35–0.64)	0.3117 (0.21–0.41)	0.8699 (0.74–0.94)	0.8624	**0.0013**

Table 3 | Whole-plant gas exchange, water status, and foliar isotopic traits at 62 kPa vs. 101 kPa.

Trait	**101** kPa Mean ± SD	**62** kPa Mean ± SD	% Change	Hedges' g	Welch p	Mann–Whitney p
VPD (Pa kPa^{-1})	19.44 ± 5.20	33.67 ± 1.20	+73.23%	+3.476	6.84×10^{-16}	1.04×10^{-9}
g_{sw} (mmol m^{-2} s^{-1})	36.64 ± 19.29	8.70 ± 3.46	−76.25%	−1.852	8.27×10^{-9}	3.69×10^{-7}
E (mmol m^{-2} s^{-1})	0.640 ± 0.282	0.291 ± 0.109	−54.58%	−1.523	2.63×10^{-7}	1.80×10^{-5}
A (μmol m^{-2} s^{-1})	1.174 ± 0.744	0.518 ± 0.387	−55.93%	−1.044	1.51×10^{-4}	1.64×10^{-4}
c_i (μmol mol^{-1})	354.8 ± 23.6	314.9 ± 49.3	−11.26%	−1.075	0.0015	0.0050
c_i/c_a	0.843 ± 0.056	0.749 ± 0.117	−11.08%	−1.058	0.0017	0.0067
RWC (fraction)	0.847 ± 0.072	0.749 ± 0.044	−11.61%	−1.550	0.0113	0.0070
δ^{13}C (‰)	-33.07 ± 0.27	-32.06 ± 0.77	+3.08%	+1.661	0.0117	0.0012
Chl area (μg cm^{-2})	17.98 ± 1.52	19.65 ± 1.67	+9.26%	+0.997	0.0319	0.0452

Table 4 | TabPFN v2 counterfactual projections ($\Delta \log_2$ FC) at NASA Exploration Atmosphere setpoints.

Setpoint	Organ	ERF-VII Hypoxia	Oxylipin/JA	Cuticle/Wax	*TCH4*	*AtPDC1*
ISS Cabin Control	Shoot 10d	−0.061	+0.071	−0.113	−0.101	+0.005
(97 kPa, 21% O_2)	Root 10d	+0.029	−0.024	−0.041	+0.137	+0.181
Staged Decompression	Shoot 10d	−0.070	+0.164	−0.128	+0.759	+0.058
(70.33 kPa, 26.5% O_2)	Root 10d	+0.068	−0.032	−0.029	+0.316	+0.837
NASA Exploration Habitat	Shoot 10d	−0.073	+0.174	−0.124	+0.859	+0.046
(56.54 kPa, 34% O_2)	Root 10d	+0.069	−0.035	−0.025	+0.345	+0.811
Uncompensated 56.54 **kPa Air**	Shoot 10d	+0.066	+0.272	−0.154	+0.445	+0.247
(56.54 kPa, 21% O_2)	Root 10d	+0.649	−0.062	−0.007	+0.300	+2.255