

---

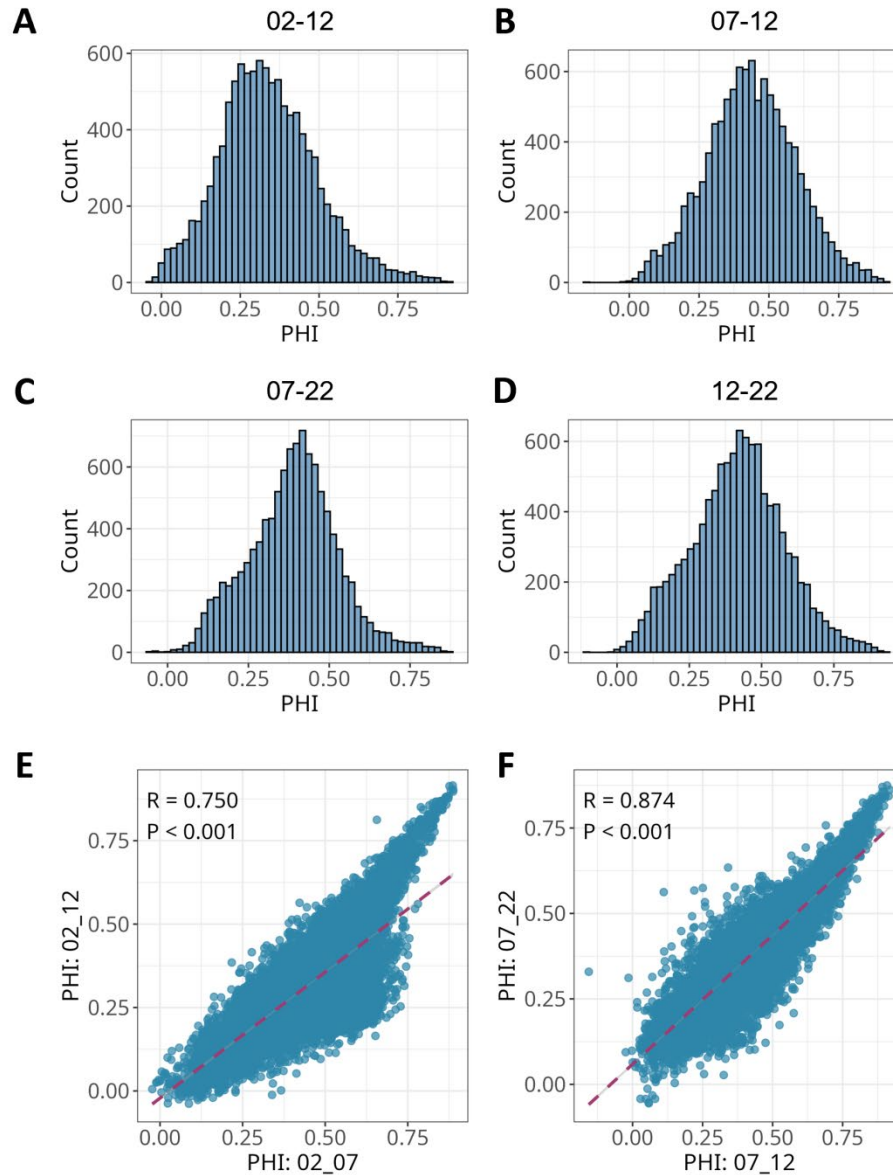
# **Temporal dynamics of the human blood proteome in ageing and disease**

---

In the format provided by the  
authors and unedited

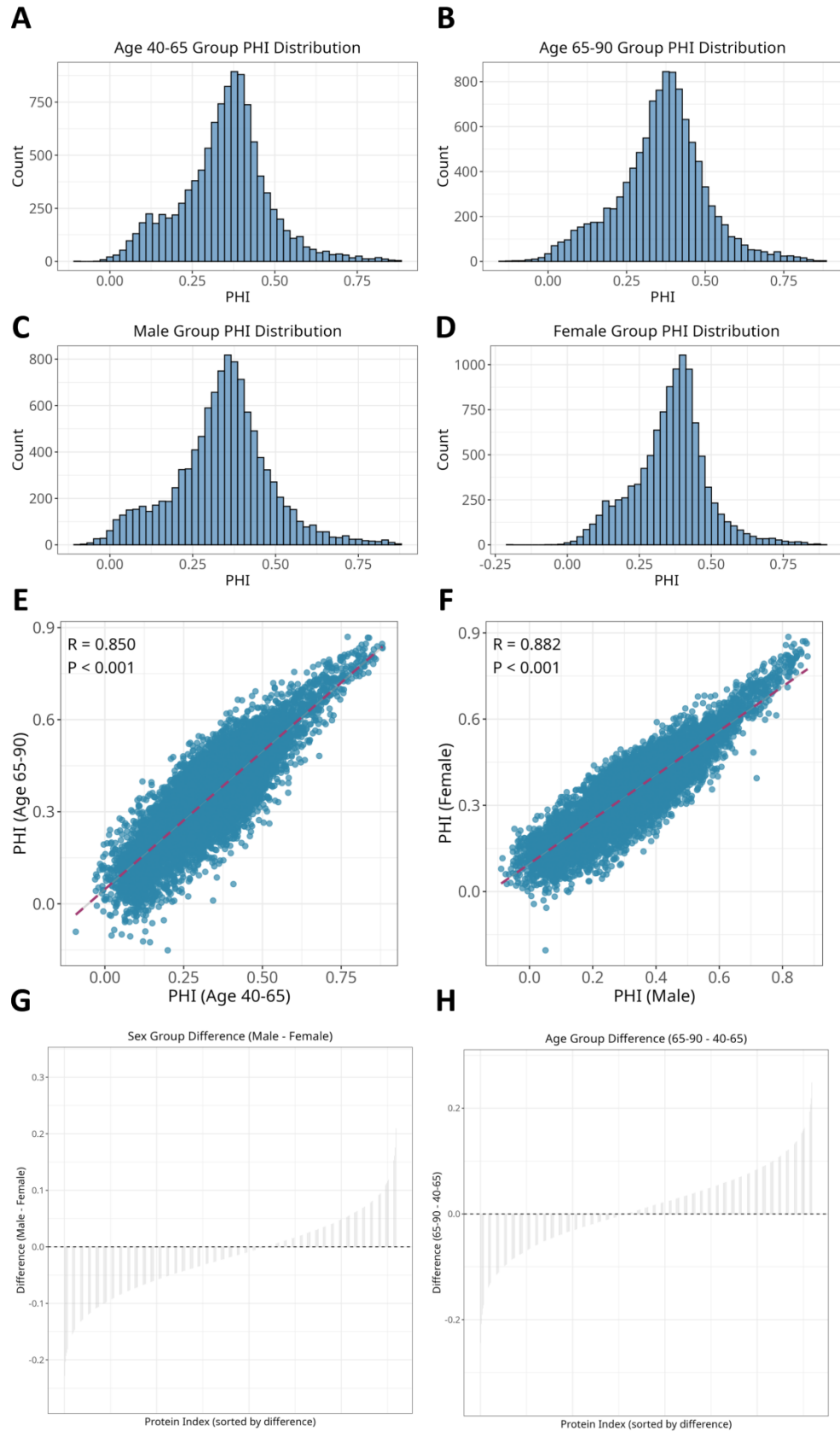
---

## Supplementary Figures



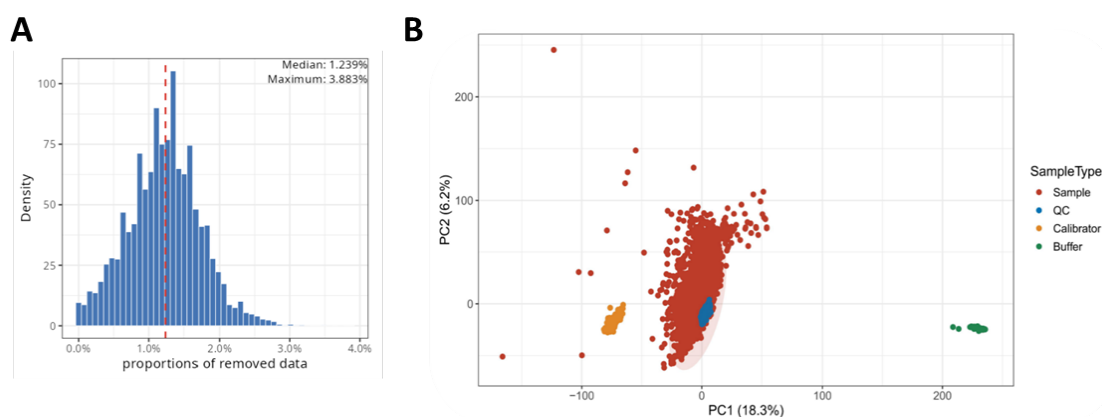
**Figure S1. Temporal analysis of PHI distributions and their longitudinal correlations across varying time intervals.** (A-D) Histograms (50 bins) show the distribution of protein-level PHI for pairwise visit intervals: 2002–2012 (A), 2007–2012 (B), 2007–2022 (C), and 2012–2022 (D). Each panel includes  $n = 10,776$  proteins; y-axis, protein count; x-axis, PHI. (E-F) Scatter plots assess the temporal consistency of protein-specific PHI (one point per protein;  $n = 10,776$ ). Plot (E) correlates PHI from 2002–2007 (x-axis) with PHI from 2007–

2012 (y-axis). Pearson  $R = 0.750$ , two-sided  $P < 1 \times 10^{-300}$ , the p-value below the machine precision limit. Plot (F) correlates PHI from 2007-2022 (x-axis) with PHI from 2002-2012 (y-axis). Pearson  $R = 0.874$ , two-sided  $P < 1 \times 10^{-300}$ , the p-value below the machine precision limit.

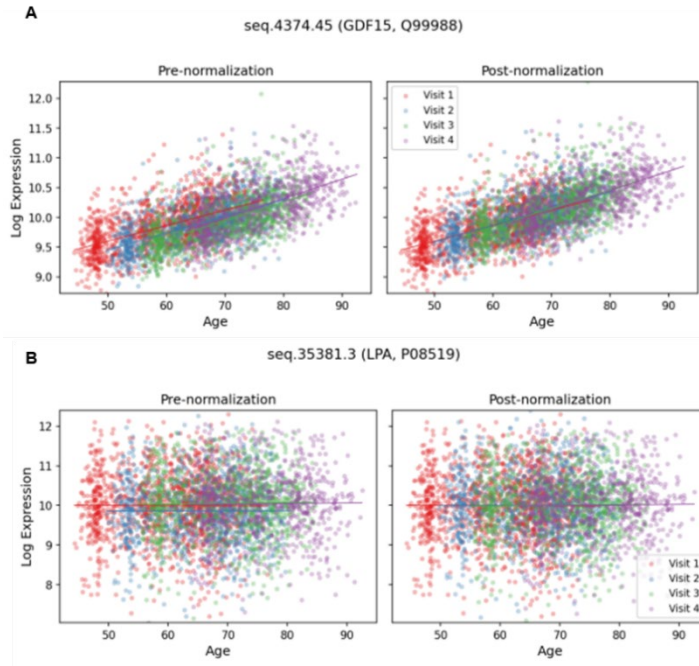


**Figure S2. Subgroup distributions and inter-group correlations of PHI.** (A-D) Histograms (50 bins;  $n = 10,776$  proteins per panel) show PHI distributions in the 40–65 years (A), 65–90

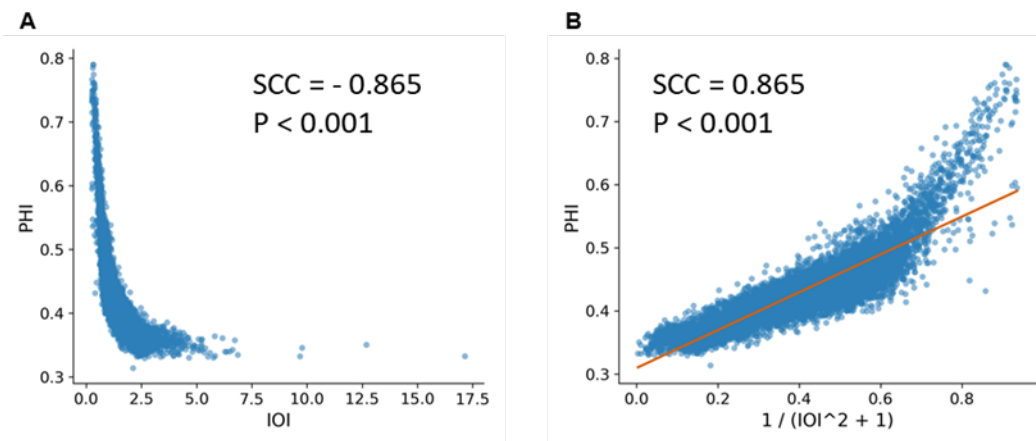
years (B), male (C), and female (D) subgroups. (E, F) Scatter plots evaluating the correlation of protein-wise PHI values between demographic groups. Panel (E) shows a strong positive correlation between PHI from the 40-65 years age group (x-axis) and the 65-90 years age group (y-axis). Pearson  $R = 0.850$ , two-sided  $P < 1 \times 10^{-300}$ , the p-value below the machine precision limit. Panel (F) shows a similarly strong positive correlation between PHI from the male group (x-axis) and the female group (y-axis). Pearson  $R = 0.882$ , two-sided  $P < 1 \times 10^{-300}$ , the p-value below the machine precision limit. (G, H) Magnitude of protein-specific PHI differences between demographic groups. Proteins are ranked along the x-axis by the absolute difference in PHI, with the y-axis showing the corresponding difference value. Panel (A) presents PHI differences between male and female groups, while panel (B) presents differences between the two age groups (65-90 and 40-65 years).



**Figure S3.** Distribution of outlier removal proportions and principal component analysis of SomaScan assay samples. **A.** For each of the 10,776 protein markers at each longitudinal visit, outliers were defined as values exceeding  $\pm 3$  standard deviations from the mean. The histogram shows the distribution of the per-protein proportion of removed data points (median = 1.2%; maximum = 3.9%).  $n = 10,776$  proteins. **B.** Principal component analysis (PCA) of study samples (samples from  $n = 1,298$  participants), calibrator controls, technical QC replicates, and buffer blanks.

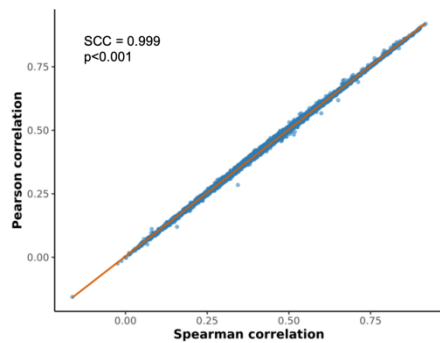


**Figure S4.** Visit-specific batch correction. For each protein, visit-specific offsets were removed by fitting  $\log(\text{expr} + 1) \sim \text{age} + \text{Dataset}$  and subtracting the estimated visit (Dataset) effects. Example proteins GDF15 (A) and LPA (B) are shown. Scatter plots display log-transformed protein levels versus age before (left) and after (right) correction; points are coloured by the four visits (2002, 2007, 2012, 2022). Solid lines, visit-stratified linear regression. Panels are illustrative/descriptive. Sample points from  $n = 1,298$  participants.



**Figure S5.** Comparison of the PHI and the index of individuality (IOI). (A) Scatter plot of the 20-year PHI (2002–2022) against raw IOI for all 10,776 assayed protein markers. Association summarized by Spearman rank correlation (SCC) between protein-level vectors:  $\text{SCC} = -0.865$ , two-sided  $P < 1 \times 10^{-300}$ , the p-value below the machine precision limit. (B) Scatter plot of 2002–2022 PHI versus transformed IOI,  $1/(1 + \text{IOI}^2)$ , with IOI estimated using all four longitudinal

visits; orange line, ordinary least-squares linear regression across proteins.  $SCC = 0.865$ , two-sided  $P < 1 \times 10^{-300}$ , the p-value below the machine precision limit.



**Figure S6.** Concordance between Pearson- and Spearman-based PHI estimates. Each point represents one protein ( $n = 10,776$ ). Pearson- and Spearman-based PHI estimates were highly concordant across proteins ( $SCC=0.999$ ,  $R^2 = 0.999$ , two-sided  $p<0.001$ , the p-value below the machine precision limit), indicating that the results were robust to the choice of correlation metric.

## Captions of Supplementary Tables 1-20

**Table S1.** Proteome-wide long-term pairwise stability and correlation coefficients. Calculated across distinct longitudinal follow-up visit combinations (e.g., 2002–2007, 2002–2012, 2002–2022) for all 10,776 protein markers using Pearson and Spearman correlation frameworks to evaluate interval-invariant intrinsic stability.

**Table S2.** Demographic stratification of protein stability and sample sizes. Detailed per-protein baseline characteristics, correlation coefficients, and effective sample sizes calculated within independent demographic subgroups, including biological sex (male and female) and chronological age blocks (40–65 years and 65–90 years).

**Table S3.** Macro-level non-parametric hypothesis tests for demographic stratification. Global Wilcoxon signed-rank test statistics, shift parameters, and exact asymptotic P-values evaluating systemic distribution differences in proteomic stability between sex and age cohorts.

**Table S4.** Ranked sentinel proteins exhibiting maximum demographic stability deviations. Top 50 circulating proteins displaying the most pronounced, statistically significant shifts in PHI between males versus females, and mid-life adulthood versus advanced age cohorts.

**Table S5.** Basic descriptors and distribution parameters of PHI. Comprehensive map listing raw and adjusted mean PHI, standard error, and internal variance parameters for every evaluated aptamer target across the human blood circulating proteome.

**Table S6.** Descriptive statistics of PHI values discretized by interval bins. Sample sizes, standard deviations, medians, and distribution ranges of calculated PHI values stratified across four mathematical intervals (0-0.25, 0.25-0.50, 0.50-0.75, 0.75-1.00).

**Table S7.** Pairwise post-hoc significance testing across PHI discretization bins. Matrix of statistical test statistics and adjustment parameters evaluating whether individual-specific stability distributions significantly differ when shifting between low, intermediate, and high PHI categories.

**Table S8.** Regression coefficient atlas mapping high-stability targets to absolute genomic beta values. Master look-up table pairing calculated protein PHI indices with absolute regression effect sizes (Log\_BETA) from downstream multi-variable association models.

**Table S9.** Distribution profiles of association effect sizes across structural PHI clusters. Quantile metrics and descriptive summaries showing the systematic, monotonic increase of regression beta values and effect sizes as proteins are grouped into progressively higher PHI stability bins.

**Table S10.** Multiple comparison significance tests for effect sizes across PHI quantiles. Statistical matrices mapping differences, confidence intervals, and multiple-testing adjusted P-values for downstream regression effect sizes across the four PHI stability tiers.

**Table S11.** Environmental and lifestyle exposure score framework matrix. Mapping of individual protein markers to their raw, adjusted, and standardized Exposure Z-Scores derived from the HEAP analytical framework.

**Table S12.** Distribution profiles of exposomic contribution variances across PHI tiers. Stratified summaries detailing how persistent lifestyle and environmental exposure components scale across the four distinct intervals of PHI.

**Table S13.** Post-hoc pairwise tests of exposomic contribution variances across PHI tiers. Two-sample testing metrics proving that intermediate-to-high PHI categories exhibit significantly greater exposomic imprints than the lowest stability tier.

**Table S14.** Proteome-wide cellular and subcellular localization classifications. Annotation matrix categorizing all 10,776 profiled blood protein targets into distinct cellular compartments (Intracellular, Extracellular Membrane, or Soluble Secreted) derived from UniProtKB cross-references.

**Table S15.** Non-parametric group differences across subcellular proteomic categories. Statistical tests assessing the distribution shifts of PHI values across cellular compartments, confirming significant homeostatic buffering variations between intracellular leakage products and active secretory factors.

**Table S16.** Structural composition of highly stable, liver-enriched circulating factors. Subcellular localization distributions, target counts, and relative percentages restricted to high-stability proteins ( $\text{PHI} > 0.75$ ) annotated as transcriptionally or translationally enriched in hepatic tissue.

**Table S17.** Structural composition of highly stable, brain-enriched neuroendocrine signals. Subcellular localization distributions, target counts, and relative percentages restricted to high-

stability proteins (PHI > 0.75) exhibiting enriched expression profiles within the central nervous system.

**Table S18.** Structural composition of highly stable, immune-system surface programs. Subcellular localization distributions, target counts, and relative percentages restricted to high-stability proteins (PHI > 0.75) enriched within immune cell types, highlighting the structural role of stable cell-surface receptors.

**Table S19.** Homeostatic failures and longitudinal trajectory deviations across three negative outcomes. Atlas of Cox proportional hazards models, Hazard Ratios (HR), Likelihood Ratio Tests, and BH-FDR adjusted P-values and the proportional hazards assumption test using global Schoenfeld residual test for prioritized proteins across three outcomes, including all-cause mortality, CHD, and HF, where individual longitudinal baseline deviations serve as independent predictors.

**Table S20.** SomaScan 11K high-throughput assay technical metadata and quality control matrix. Comprehensive technical registry for all experimental targets, including SOMAmer Probe IDs, Sequence Identifiers (SeqId), Target Full Names, UniProt Accession Numbers, Entrez Gene IDs, sample assay dilution factors, plate calibration references, and formal Row/Column Quality Control pass/fail alerts.

