PROTIFI™
Bringing precision omics to life!™

SimpliFi™: intuitive multi-omics understanding of complex data series using mutual information

Jim Palmeri, John P. Wilson, PhD

INTRODUCTION

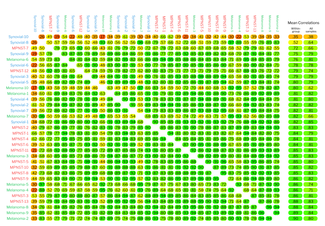
The rapid growth of omics data makes extracting actionable insight hard. The most informative biomolecule is rarely known in advance: protein, nucleic acid, lipid, metabolite or glycan, so analysis must interrogate many omics layers at once.

SimpliFi is an intuitive, assumption-free data-to-meaning engine. It derives statistics directly from the data's own structure using non-parametric methods, and uses resampling to attach a confidence interval to every p-value and fold-change. It handles non-Gaussian biological measurements, works across all omics types, runs on GPUs for results in minutes, and is shared by a single URL.

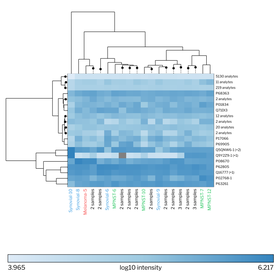
Responses across time or concentration take many shapes, so methods that assume one shape miss real biology. SimpliFi adds mutual information to find trends of any shape.

QUALITY CONTROL & VISUALIZATION

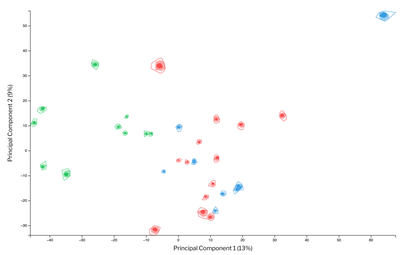
Before analysis, SimpliFi makes data quality visible and lets users explore interactively. A few of its built-in views:



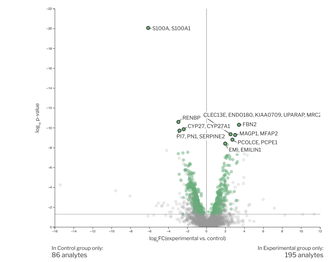
Sample correlation



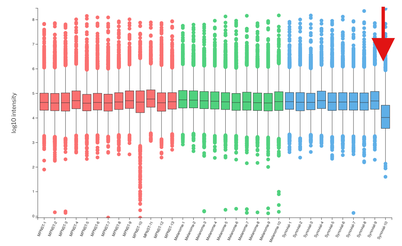
Hierarchical clustering



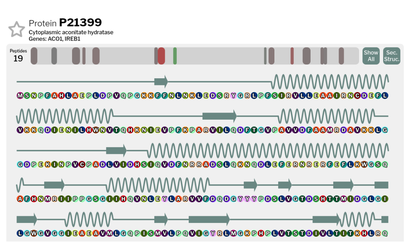
PCA and outlier detection



Volcano plots



Intensity distributions (outlier flagged)



Peptide-level views

NON-PARAMETRIC ANALYSIS

Biological measurements are rarely Gaussian, and their variability depends strongly on signal intensity: at low intensity, stochastic sampling gives broad variation, while at high intensity effects such as saturation appear. Parametric tests, including the t-test, assume this away and can be wrong by orders of magnitude.

SimpliFi instead derives significance from the data itself and reports every result with a confidence interval, so confidence scales with how strongly a feature is observed and with replicate agreement. This non-parametric foundation is what makes assumption-free trend detection possible, and it carries directly into the mutual-information analysis below.

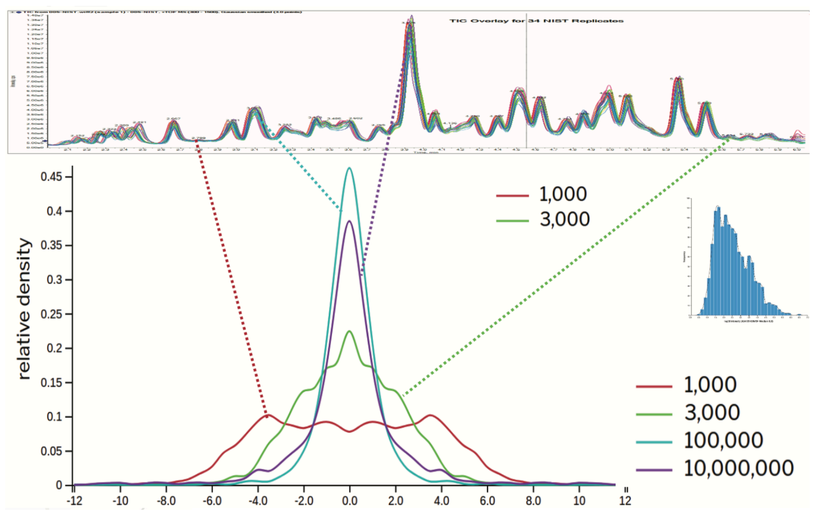


Fig. 1. Replicate variability is a function of observation intensity. The spread of repeated measurements is broad at low signal and narrows as intensity increases, so SimpliFi accounts for intensity rather than assuming constant variance.

MUTUAL INFORMATION

For an ordered factor such as time or concentration, SimpliFi quantifies how strongly each biomolecule depends on that factor using mutual information (MI). MI measures dependence rather than slope, so it assumes no linearity or specific shape and flags monotonic, U-shaped, inverted-U, delayed, multiphasic and cyclic responses alike. A directional rank correlation gives the sign of monotonic trends, with permutation-based significance for both. Because MI needs no model form, the same workflow applies to every omics modality at once, and significance is calibrated by 1000 sample-label permutations.

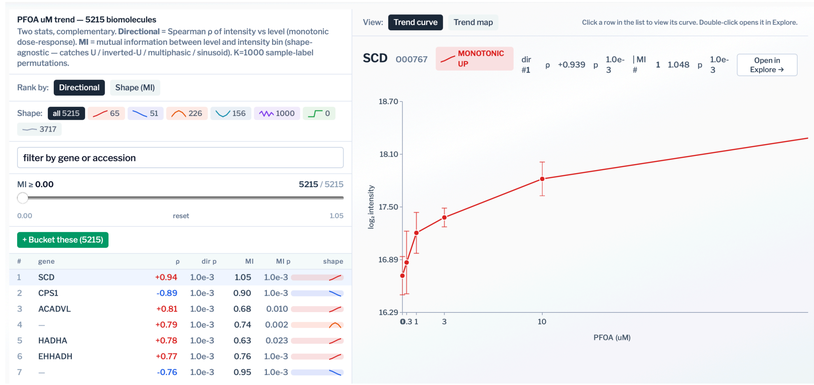


Fig. 2. The SimpliFi trend explorer. Features are ranked by directional correlation or by shape (MI), filtered by trend archetype, and any feature opens its trend curve with confidence intervals.

MI IN PRACTICE

A PFAS concentration series in HuH-7 cells provides a working demonstration. Known PFAS-affected proteins and planted arbitrary-shape controls were both recovered by MI, showing how MI behaves rather than making a biological claim.

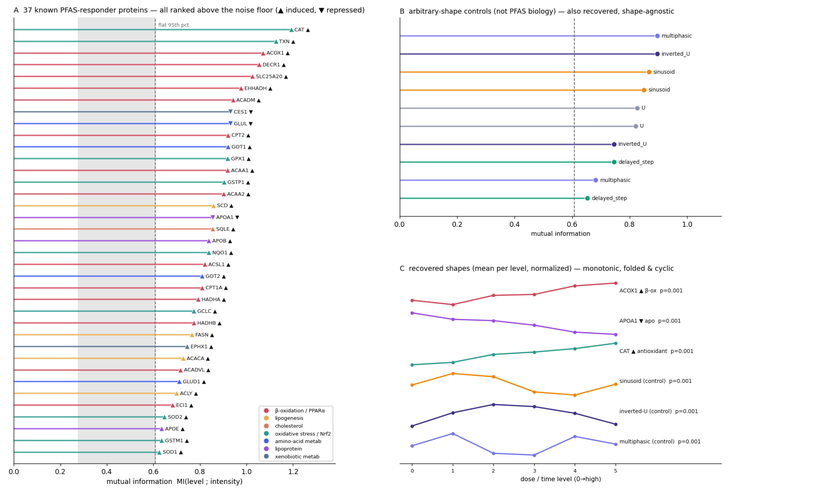


Fig. 3. MI ranks informative features and recovers their trend shapes, including monotonic, folded and cyclic responses, without being told what shape to expect.

DATA TO MEANING

Newly integrated external databases annotate each biomolecule and place it in biological context. Features flagged by MI carry into pathway views such as Reactome, so coordinated changes across proteins, metabolites and lipids read at the pathway level.

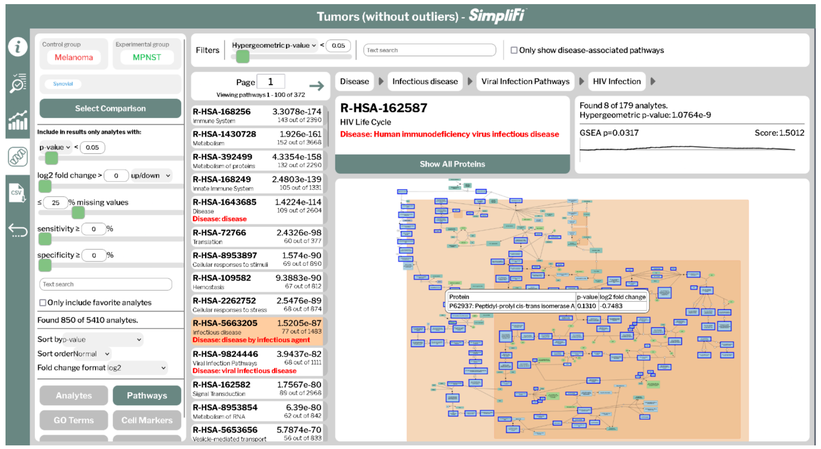


Fig. 4. MI-selected features shown in a Reactome pathway view inside SimpliFi, linking detected trends to mechanism without leaving the interface.

CONCLUSIONS

SimpliFi brings assumption-free, non-parametric analysis and mutual-information trend discovery to multi-omics time-course and concentration-course studies. Mutual information detects informative trends of any shape, a directional correlation supplies the sign of monotonic responses, and resampling places a confidence interval on every estimate. Users move from quality control to trend discovery to biological mechanism, making rigorous multi-omics series analysis accessible to non-omics users.

ProtiFi LLC, Fairport, NY