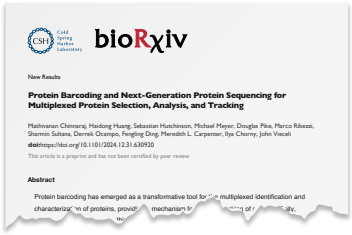


Protein Barcoding and Next-Generation Protein Sequencing for Multiplexed Protein Selection, Analysis, and Tracking

Mathivanan Chinnaraj, Haidong Huang, Sebastian Hutchinson, Michael Meyer, Douglas Pike, Marco Ribezzi, Sharmin Sultana, Derrek Ocampo, Fengling Ding, Meredith L. Carpenter, Ilya Chorny, John Vieceli

This study establishes a scalable, sensitive, and accessible protein barcoding framework that pairs seamlessly with Next-Gen Protein Sequencing™ (NGPS™) using the Platinum® benchtop instrument, offering a powerful alternative to mass spectrometry for quantitative, multiplexed protein analysis in both research and pre-clinical drug development.

Key Innovations

NGPS Platform: Sequencing offering single-molecule resolution, avoiding mass spectrometry limitations

Direct protein detection: Quantify expression and differentiate functional performance of proteins

Barcoding System: Barcodes link to proteins using common affinity tags

Workflow Optimization: Total workflow time <6 hours with 1 hour hands-on time

Applications

Drug Screening: Screen nucleic acid therapy targets by relative expression

Drug Delivery: Track therapeutic delivery by direct protein expression in model organisms

Protein Engineering: Track and identify variants in pooled functional screens

Proteomics: Map protein-protein interactions and expression

Performance Highlights

Metric	Value
Limit of detection	50 fmol per single barcode
Dynamic range	10-fold
Reproducibility	Across kits, instruments, users

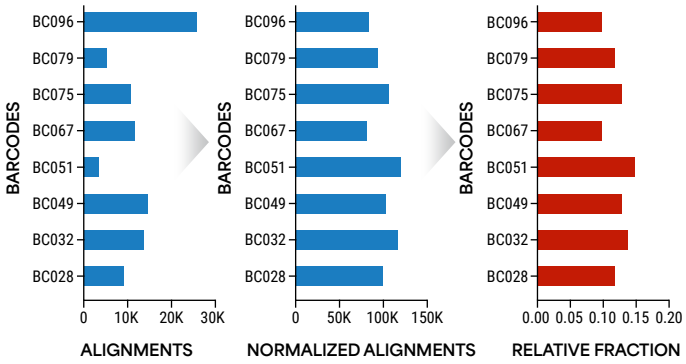
Key Results

Sensitive and quantitative read-out: Demonstrated the ability to accurately normalize barcode abundance across a **10-fold** dynamic range and detect barcodes at concentrations as low as **50** fmol for individual barcodes and ~**400** fmol within an 8-plex mixture.

Broad dynamic range: The barcoding method achieves a **10-fold** dynamic range, emphasizing its sensitivity in detecting low-abundance variants.

Reproducibility and robustness: All proteins in a mixture of five proteins expressed in *E. coli* across eight runs with multiple lots of reagents/kits and two operators were successfully identified with a false discovery rate (FDR) of **less than 10%**.

Normalization of Barcodes



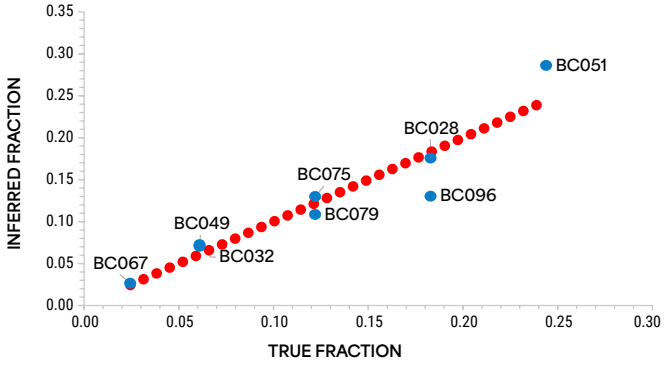
Normalization factors were derived from over 25 sequencing runs. When normalization is applied, mean absolute percent error (MAPE) is <20% and FDR is <0.1, indicating high accuracy and reproducibility for differentiating high-performance candidates.

Limit of Detection

Barcode	LOD (pmol)
BC028	0.41
BC032	0.41
BC049	0.35
BC051	0.41
BC067	0.35
BC075	0.35
BC079	0.41
BC096	0.35

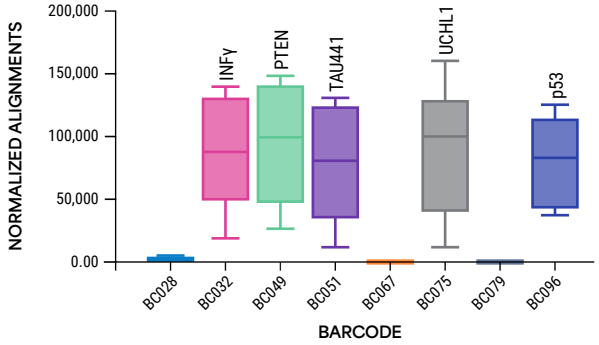
Barcodes were identified at LOD as low as ~400 fmol within an 8-plex mixture and at 50 fmol for individual barcodes.

Ten-fold Dynamic Range



In an 8-plex mixture, titration from 1x down to 0.1x produces a linear dynamic range (R2 of 0.9) with low MAPE and FDR below 0.1.

Recovery of a Mixture of Proteins



Eight runs with multiple lots of five proteins expressed in *E. coli* demonstrate that barcodes can accurately recover relative abundances in a mixture of full-length proteins; barcode presence did not impede expression.

Conclusion

The protein barcoding workflow – now supported by the Barcoding Kit – enables multiplexed protein selection, analysis, and tracking, with applications ranging from protein engineering to nucleic acid therapy development.

Future Directions

Barcode design is optimized for scale and expansion and future ambitions include expanding barcode libraries and applying barcoding in live systems.

Explore the manuscript detailing the development and validation of a robust protein barcoding system with NGPS on Platinum for high-sensitivity, multiplexed analysis of proteins →