

## Unbiased Proteomics Can Open New Doors in Cancer Research

**Proteomics can reveal the molecular mechanisms that directly impact disease**

Mar 16, 2022 MARGARET DONOVAN, PHD

Margaret Donovan, PhD, joined Seer Bio in June 2020 as a bioinformatics scientist. Following her core interest in better understanding the molecular mechanisms driving human disease, she focused on diving into deep plasma proteomics and discovering how protein complexities and proteoforms relate to disease. Donovan received her doctorate in bioinformatics and system biology from UC San Diego in 2019, where she studied how genetic variation impacts the transcriptome and its relationship to disease in fetal-like and adult tissues.



Though genomics paved the way for personalized health care, proteomics has revealed the molecular

Margaret Donovan, PhD, bioinformatics scientist at Seer Bio

mechanisms that directly impact disease. Profiling the proteome (the complete set of proteins in a cell at a given time) has the potential to significantly expand our understanding of human health and show us new channels to finding more precise therapeutics that may ultimately treat diseases like cancer.

With genomics, we have been able to sequence hundreds of thousands of whole genomes and several million exomes, revealing over 1 billion genomic variants. But less than 1 percent of those variants have been functionally characterized or understood; for example, determining changes variants could cause in a cell or in a biological pathway.

Conversely, understanding how these genomic variants drive biological change can be achieved through proteomics studies, including measuring proteins, peptides with and without modifications, proteoforms, and protein regulation and degradation. Due to the complexity of the proteome, this has largely been technologically insurmountable until recently.

## A new gateway to the proteome

Unbiased, deep, scalable proteomics technologies are now removing the barriers to open up a new gateway to the proteome. With these new tools, researchers can interrogate the proteome in the same way they did the genome and transcriptome, making new discoveries to further understand how the cellular machinery of our cells works.

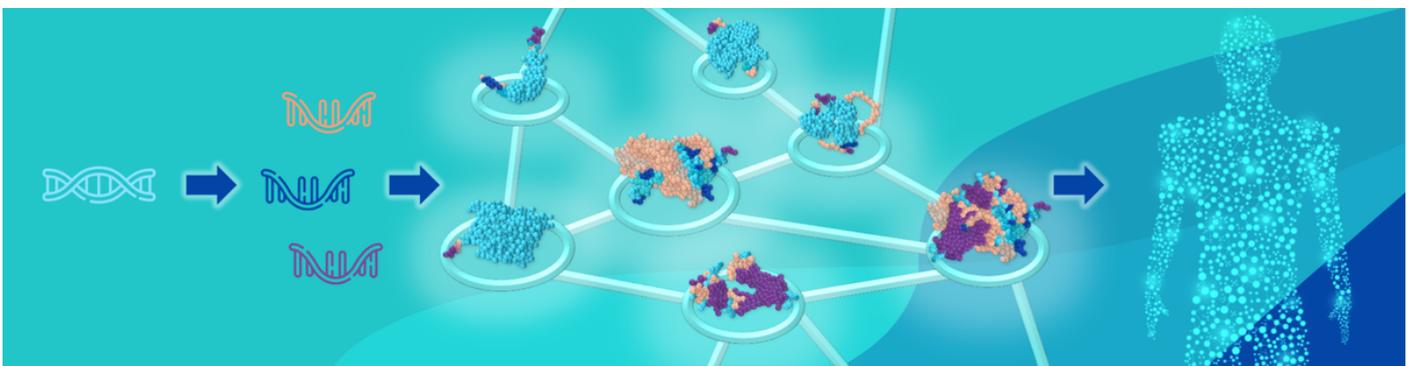
Taking an unbiased approach to studying the proteome means we are not limited to targeting specific proteins already linked to disease types or phenotypes to validate known biology. The plasma proteome has a large dynamic range with more than 10 orders of magnitude of difference between the most abundant proteins, like albumin, and those that are scarce but may be of interest in disease, like cytokines. Most proteomics technologies don't do a good job of accessing that complete range in a scalable way.

Some solutions are scalable, but they only target predetermined proteins to confirm their presence and, as such, are biased to what you already know about a disease. Other solutions can look more deeply at the proteome but are fundamentally not scalable, as they are time-consuming and cumbersome (e.g., mass spectrometry coupled with protein chromatography, depletion, and fractionation methods). Although they are unbiased, these methods do not scale to larger sample sizes and the largest studies to date using these methods have included fewer than 40 samples.

## Developing plasma-based biomarkers

When it comes to diseases like cancer, early detection is key for better patient outcomes. Plasma and other blood-based samples are promising for cancer diagnostics, as they can be regularly obtained using noninvasive methods, compared to the alternative approach of invasive biopsies—but analyzing the plasma proteome has historically been difficult.

We recently developed a plasma-based biomarker discovery platform for non-small cell lung cancer, or NSCLC, using an unbiased proteogenomic approach and analyzed early-stage cancer samples and healthy controls to dissect the differences between protein variants arising from a single gene. Performing our untargeted proteomics technology at scale, we were able to identify lung cancer-associated protein variants.



COURTESY OF SEER BIO

## Exploring the dynamic range of proteins

Multiple forms of a given protein (proteoforms) arise in different ways, including genetic variation, alternative splicing, and post-translational modifications, any of which can change a protein's function and its underlying role in human health and disease. It's estimated that there are more than a million potential proteoforms, therefore, it is impractical and likely infeasible to target all proteoforms and cover the diversity needed to further deconvolute biology.

Rapid, scalable, deep, and unbiased plasma proteomics lets us explore the dynamic range of proteins to generate peptide-level identification and quantification to infer proteoform presence and abundance. In [our recent preprint](#) in bioRxiv, we show that peptide-centric analyses identify disease-linked proteoforms that would not have been discovered using protein-level information. Though it has yet to be peer reviewed, this study further shows that peptide-level resolution enables us to identify hundreds of proteoforms, several of which are significantly associated with NSCLC, notably including BMP-1.

## The potential of unbiased proteomics

Through this work and other studies, we've found that unbiased proteomics methods can generate deep plasma proteome profiles that have the potential to reveal novel biological and mechanistic insights. Advancements in proteomics workflows are changing the way scientists are thinking about larger and more comprehensive studies to obtain a more complete picture of what's happening in the human body.

CLINICAL **LM**  
Lab Manager

<https://www.clinicallabmanager.com>

© 1986 – 2022 CLINICAL LAB MANAGER. ALL RIGHTS RESERVED.