EARLY ACCESS

# Illumina Protein Prep

## NGS-based solution for deeper insights into the proteome

- Measure > 6K unique human proteins in a single plasma or serum sample using the SomaScan® proteomics assay with NGS readout
- Go from sample to results in < 2.5 days with just 4 hours hands-on time with a streamlined, automated workflow
- Analyze proteomic data with integrated secondary analysis via DRAGEN™ Protein Quantification and Illumina Connected Analytics

illumina®

## Introduction

Proteins play a key functional role in human biology, reflecting a real-time snapshot into health and disease states. Insights from proteomics serve as a critical link between genotype and phenotype, enable a deeper understanding of disease mechanisms, and ultimately help with the prediction, monitoring, and prevention of disease escalation. High-throughput proteomics assays with next-generation sequencing (NGS)–based readouts enable large-scale proteomic studies and have the ability to connect genomic and proteomic data sets, accelerating multiomics research.

Illumina Protein Prep is a comprehensive high-throughput proteomics solution that integrates trusted Illumina sequencing by synthesis (SBS) chemistry with the high-sensitivity of the Standard BioTools SomaScan proteomics assay. This innovative proteomics assay uses SOMAmer® (slow off-rate modified aptamer) Reagents for protein capture to achieve high specificity for target proteins compared to antibody-based approaches.[1] Combining this advanced proteomics assay with NGS-based readout and the bioinformatics power of Illumina data analysis software gives researchers a streamlined sample-to-results solution for evaluating more than 6K* unique human proteins in a single plasma or serum sample (Figure 1).

* > 9K unique human protein targets, corresponding to 11K SOMAmers, will be available in early 2025.

## Comprehensive content

The Illumina Protein Prep assay enables discovery and quantification of > 6K unique proteins in plasma or serum samples. The assay includes 7002 SOMAmer Reagents, of which 6831 target a total of 6056 unique human proteins and 171 target nonhuman proteins and controls. This comprehensive content† targets human proteins across a diverse set of pathophysiological processes, including cancer, inflammation, immunology, and cardio-metabolic function. The content covers major molecular targets, including receptors, kinases, growth factors, and hormones, spanning secreted, intracellular, and extracellular proteins across more than 200 biological pathways.

All SOMAmer Reagents used in Illumina Protein Prep have undergone rigorous characterization and the specificity of the SOMAmer Reagent for its cognate protein has been confirmed by orthogonal methods, such as mass spectroscopy and enzyme-linked immunosorbent assay (ELISA).[1] With the content included in the Illumina Protein Prep solution, researchers can access deep insights into the proteomics landscape, enabling the discovery of novel biomarkers, drug targets, and insights into disease mechanisms.

† Complete menu of proteins available upon request.

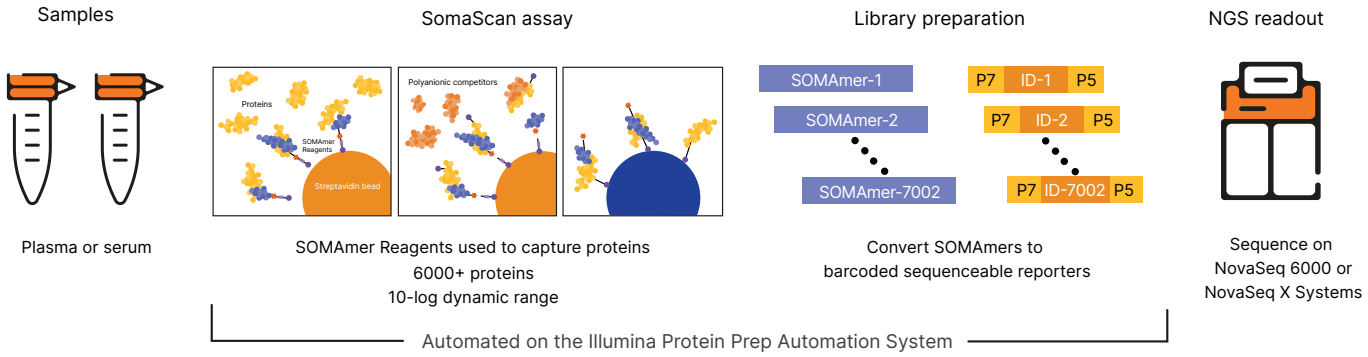


Figure 1: Overview of the Illumina Protein Prep solution—The high-throughput proteomics solution uses the SomaScan proteomics assay for high-plex protein detection from serum or plasma samples, followed by Illumina library preparation and sequencing on the NovaSeq™ 6000 or NovaSeq X Systems. Data analysis is performed with integrated secondary analysis via DRAGEN Protein Quantification and Illumina Connected Analytics. The entire workflow can be automated on the Illumina Protein Prep Automation System, a custom Tecan Fluent 780, for consistent and reproducible results.

# Sensitive and precise protein detection

Illumina Protein Prep uses innovative slow off-rate modified aptamers, or SOMAmer Reagents, as protein affinity molecules to achieve exceptional sensitivity and specificity for protein detection across a broad dynamic range.[2] These aptamers, which are short pieces of single-stranded DNA with hydrophobic modifications, provide a high degree of shape matching to the protein target, allowing discernment between nearly identical proteins. Data obtained using NGS readout demonstrate a high degree of concordance with microarray data (Figure 2). Additionally, unlike polyclonal antibodies, which can be variable in structure and performance, SOMAmer Reagents are based on binding kinetics, providing femtomolar sensitivity[2] and excellent reproducibility, with a low median coefficient of variance (CV) under 10% (Table 1, Figure 3).

Table 1: Summary of precision metrics

| Sample | Intra-run CV | Inter-run CV | Total-run CV | Single-site CV |
|---|---|---|---|---|
| Plasma | 7.63% | 4.92% | 9.17% | 8.43% |
| Serum | 7.86% | 4.42% | 9.21% | 8.45% |

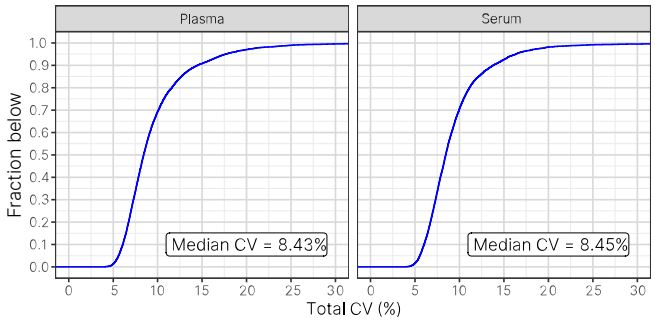


Figure 3: Cumulative distribution plots for total CV for plasma and serum—A pooled sample for precision measurements was created by pooling 10 plasma and 10 serum samples from healthy donors. Twelve replicates of each of the plasma or serum pooled sample were included on each automated run of the Illumina Protein Prep assay. Eight individual automation runs were performed for both serum and plasma across three different laboratory locations. Libraries were sequenced and data were normalized using the proteomic secondary data analysis pipeline, DRAGEN Protein Quantification. The intra-run, inter-run, and total CV values were calculated for each of the 6831 SOMAmers targeting human proteins across all eight runs and at all of the locations. The total median CV is shown for a single site reflecting performance from a single laboratory setup of the automated Illumina Protein Prep assay.

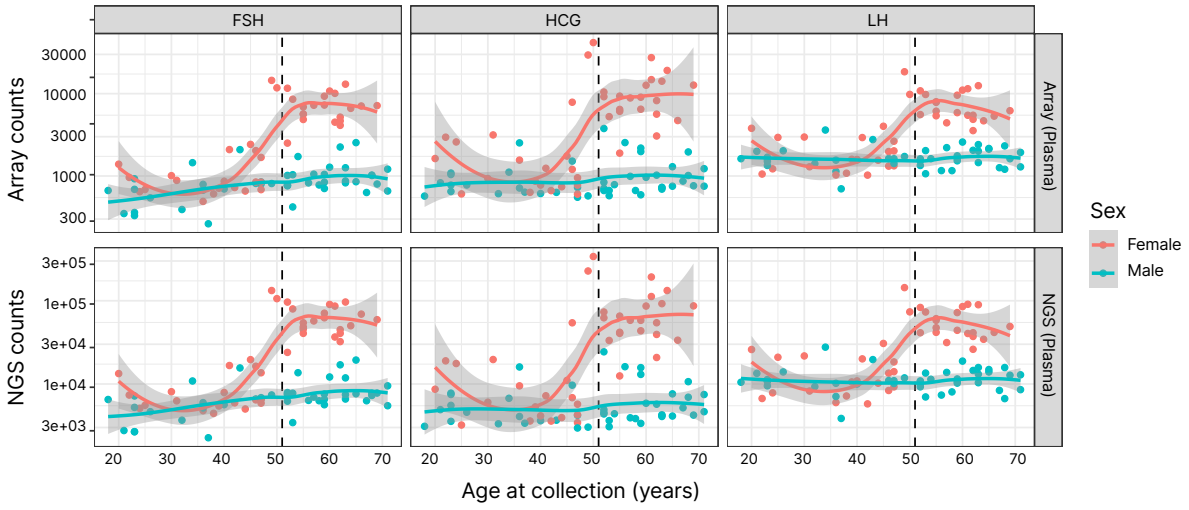


Figure 2: Detecting protein abundance changes associated with menopause in plasma samples using microarrays and NGS readouts—Differential expression of three known regulatory genes involved in menopause were detected using microarray and NGS readouts. Plasma samples from 85 healthy males and females were tested across a range of ages. For all three genes, protein abundances were higher in older females compared to males and data were comparable between microarray and NGS readouts. FSH, follicle stimulating hormone; HCG, human chorionic gonadotropin; LH, leutenizing hormone.

# Hybridization-based NGS assay

The first step in the Illumina Protein Prep workflow is the SomaScan assay, which quantitatively transforms the protein epitope availability in a biological sample into a specific SOMAmer Reagent-based DNA signal.[3] The initial SOMAmer Reagent–protein binding step is followed by a series of bead capture and wash steps to convert relative protein concentrations into SOMAmer Reagent abundancies (Figure 4). Next, SOMAmer Reagents are converted into barcoded sequencing libraries using a hybridization-based approach. Pairs of probes are hybridized to SOMAmer Reagents in an overnight incubation step before being captured on magnetic beads. Each SOMAmer Reagent has a unique pair of probes, one of which carries a barcode corresponding to a particular SOMAmer Reagent. Unbound probes are washed away to ensure the relative abundance of the SOMAmer Reagents is converted to the abundance of the barcoded probe. Index PCR primers are added to append sample indexes for sequencing and amplified to create individually indexed-barcoded libraries. The entire Illumina Protein Prep assay is automated on a single platform, the Illumina Protein Prep Automation System. Samples are pooled and sequenced on the NovaSeq 6000 or NovaSeq X System (Figure 5).

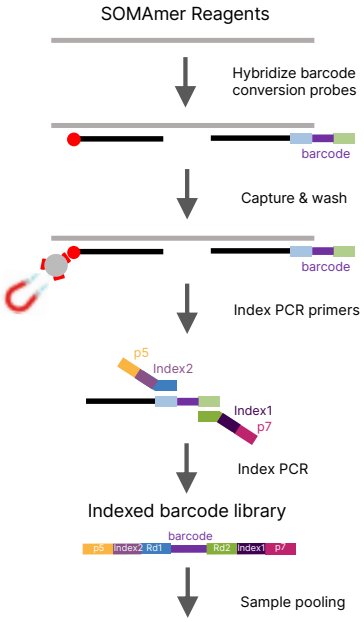


Figure 5: Illumina Protein Prep NGS conversion chemistry—Following the SomaScan assay, SOMAmer Reagent abundancies are converted to barcoded sequencing-ready libraries using a hybridization-based approach.

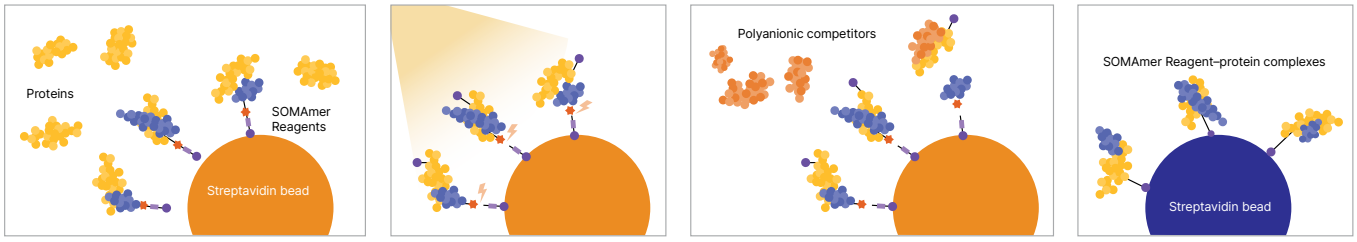


Figure 4: Protein-capture using the SOMAScan assay—SOMAmer Reagents (blue) contain a photocleavable linker and biotin. SOMAmer Reagents bound to streptavidin beads are used to capture specific proteins (yellow) from a complex mixture of proteins in serum or plasma. Unbound proteins are washed away and bound proteins are tagged with biotin. Next, UV light is used to break the photocleavable linker and release SOMAmer–protein complexes back into solution. During the incubation step, nonspecific complexes dissociate while specific complexes remain bound. The inclusion of polyanionic competitors during the incubation step prevents nonspecific rebinding of dissociated proteins. Specific protein-SOMAmer Reagent complexes are captured on new streptavidin beads and are eluted for relative quantification using NGS.

## Scalable, streamlined workflow

The Illumina Protein Prep solution uses a streamlined end-to-end workflow (Figure 6) that starts with plasma or serum samples, followed by highly sensitive protein capture using innovative SOMAmer Reagents and Illumina library preparation. Libraries are sequenced on the NovaSeq 6000 System (S4 flow cell) or the NovaSeq X System (10B flow cell), with 170 samples and 22 controls per run. The entire workflow is automated on the Illumina Protein Prep Automation System, enabling labs to go from samples to results in less than 2.5 days with ~4 hours of hands-on time (Table 2). Integrated proteomic analysis powered by DRAGEN Protein Quantification and Illumina Connected Analytics delivers highly accurate results for deeper biological insights.

Table 2: Illumina Protein Prep specifications

| Parameter | Specification |
|---|---|
| Sample type | Plasma or serum |
| Input volume | 55 µl |
| Total no. of SOMAmer Reagents | 7002 |
| No. of human proteins targeted | 6056 unique human proteins |
| Dynamic range | 10-log (fM to mM) |
| Throughput | 170 samples + 22 controls per run on the NovaSeq 6000 S4 flow cell or NovaSeq X 10B flow cell |
| Total workflow time | 2.5 days |
| Hands-on time | ~4 hours |

## Integrated data analysis

Labs can analyze sequencing data easily using a fully integrated proteomic secondary data analysis pipeline that includes NGS and proteomic assay–specific normalization methods. This analysis pipeline can be accessed on DRAGEN Protein Quantification and Illumina Connected Analytics. Integration with Illumina Connected Analytics offers a secure, streamlined, cloud-based platform to scale up secondary analysis and reduce manual touchpoints, with data streaming and autolaunch capabilities.

## Summary

The Illumina Protein Prep solution is a comprehensive, end-to-end proteomics workflow with NGS-readout for large-scale protein studies. This high-performance assay leverages innovative SOMAmer Reagents to detect > 6K unique human proteins from a single plasma or serum sample with femtomolar sensitivity and excellent reproducibility. The comprehensive content of Illumina Protein Prep contains rigorously validated protein-affinity reagents, covering key biological processes, including cancer, inflammation, immunity, cardiometabolic function, and more. By combining this high-plex proteomics assay with an NGS readout, Illumina Protein Prep solution enables the integration of proteomic data with genomic and transcriptomic data, paving the way for high-impact multiomics research.

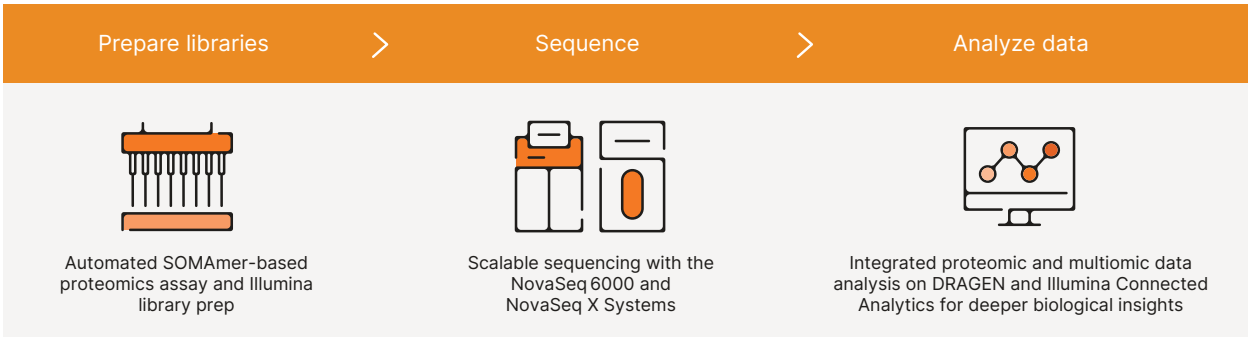


Figure 6: Overview of the streamlined, sample-to-results Illumina Protein Prep workflow—This NGS-based proteomics solution combines SOMAmer technology with trusted Illumina sequencing and DRAGEN data analysis.

## Learn more

Illumina Protein Prep

NGS-based proteomics

NovaSeq 6000 System

NovaSeq X System

DRAGEN secondary analysis

Illumina Connected Analytics

## Ordering information

| Product | Catalog no. |
|---|---|
| Illumina Protein Prep 6K Plasma (96 samples) | 20104377 |
| Illumina Protein Prep 6K Serum (96 samples) | 20104378 |
| Illumina Protein Prep Automation System | 20116818 |

## References


1. SomaLogic. SomaScan platform. somalogic.com/somascan-platform. Updated December 2023. Accessed January 2, 2024.
2. SomaLogic. SomaScan 7K Assay v4.1 technical note. somalogic.com/wp-content/uploads/2023/01/Rev4_2022-01_SomaScan_Assay_v4.11.pdf. Updated January 2022. Accessed January 2, 2024.
3. Gold L, Ayers D, Bertino J, et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One*. 2010;5(12):e15004. doi:10.1371/journal.pone.0015004.



illumina®

1.800.809.4566 toll-free (US) | +1.858.202.4566 tel
techsupport@illumina.com | www.illumina.com



M-GL-02533 v2.0