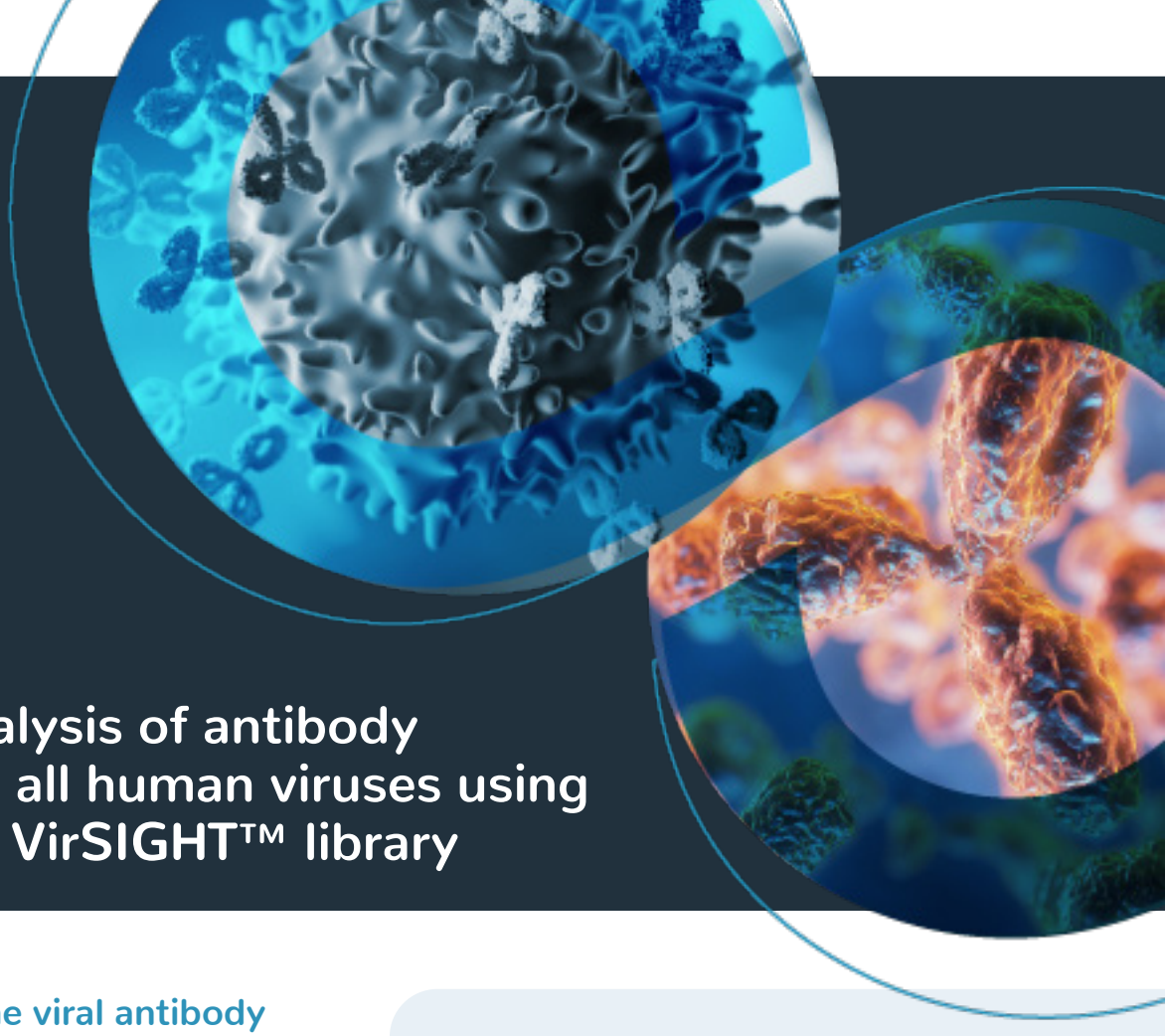


WHITE PAPER

# Unbiased analysis of antibody responses to all human viruses using Infinity Bio's VirSIGHT™ library

## Why analysis of the viral antibody reactome is important

The viral antibody reactome encompasses the set of all possible binding interactions between antibodies and viral antigens. Each individual's viral antibody reactome is unique and carries a wealth of information about prior and ongoing immune responses to viruses. By cataloging the different antibodies that an individual has produced in response to viral infections and vaccinations, scientists can learn more about how the immune system recognizes and eliminates viruses, and how viruses or the immune responses might be involved in complex human diseases.

> antibody reactome: **the set of all possible binding interactions between antibodies and viral antigens**

> MIPSA: **Molecular Indexing of Proteins by Self-Assembly is Infinity Bio's novel technology for antibody reactome profiling**


## Summary

**Sensitive and specific:**
in our study, VirSIGHT detects all positive infections with little to no signal from non-infecting viruses.

**Comprehensive and efficient:**
over 2 million viral protein sequences were clustered to optimally capture strain-level diversity with minimal redundancy.

**Completely unbiased profiling**
of an individual's antibody reactome can elucidate how the adaptive immune system recognizes viral pathogens, while uncovering novel mechanisms driving complex human diseases.

Detecting antibodies targeting unknown, transient, chronic or latent viral infections can lead to the discovery of novel associations with disease phenotypes in longitudinal (Fig 1A), and cross-sectional (Fig 1B) studies. For example, a longitudinal study showed that Epstein-Barr Virus (EBV) infection always precedes the onset of multiple sclerosis (MS) (Bjornevik et al., 2022), a chronic demyelinating disease of the central nervous system. Another cross-sectional study of children with a mysterious severe acute hepatitis revealed a pathologic infection by adeno-associated dependoparvovirus A (AAV-A) (Mitchell et al., 2023). VirSIGHT, which covers over 500 viruses known to infect humans, can also enhance both surveillance and pandemic preparedness efforts.

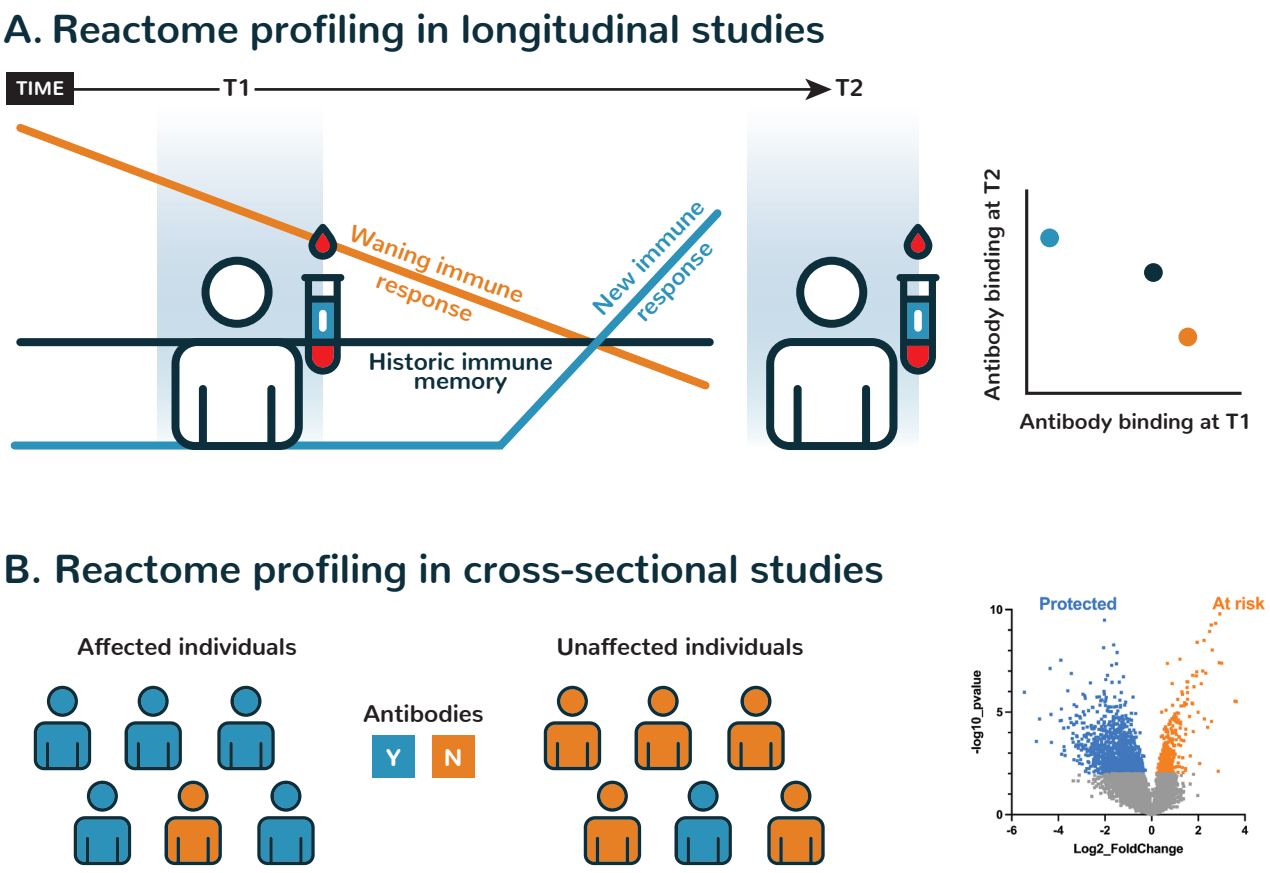



**Fig 1.** Reactome profiling in **(A)** longitudinal and **(B)** cross-sectional studies enables detailed characterization of reactome profiles across a population and the discovery of novel disease associations with viral infections.

**Whether looking at acute, previous, or latent viral infections, VirSIGHT can identify novel disease associations, immune phenotypes (such as dysregulation or amnesia), aid in serosurveys/surveillance, and pandemic preparedness.**

## VirSIGHT is the most comprehensive human virome library ever designed

The VirSIGHT library design incorporates over 2 million viral protein sequences deposited in the UniProt database that cover all known human viruses (over 94,000 unique taxonomic IDs corresponding to over 500 viral species).

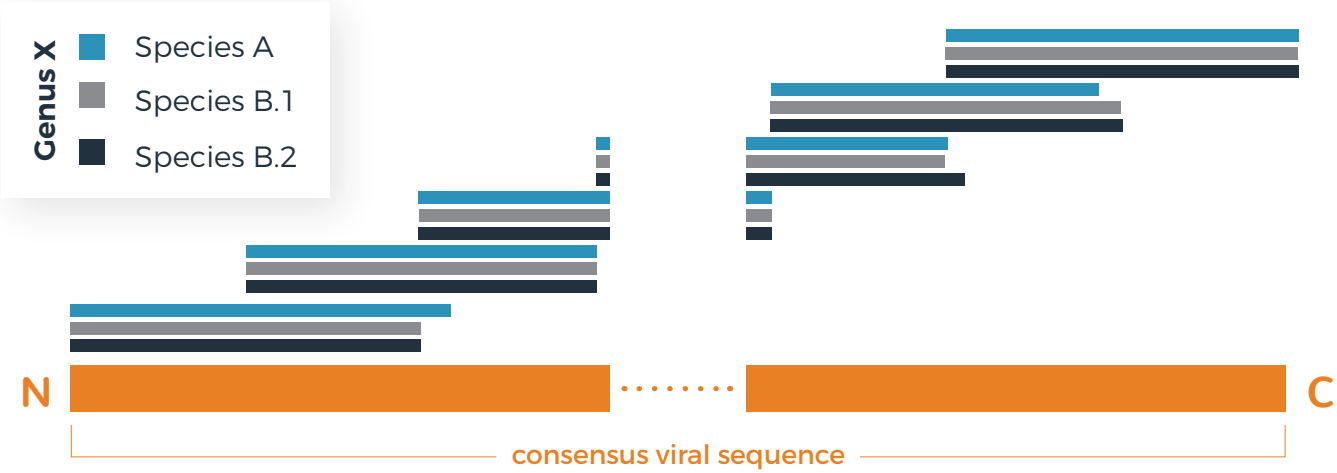
Our proprietary multiple sequence alignment (MSA)-based tiling approach organizes viral proteins into a standardized genus-level coordinate system for coverage with overlapping ~ 60 amino acid long peptide tiles, using an adjacent tile overlap length of 20 amino acids, representing a minimally inclusive epitope size (Liebhoff, et al. 2024; Jayaraman, et al. 2024) (Fig 2).

After clustering at 95% sequence identity (90% for human immunodeficiency virus 1), 285,669 unique peptides were designed to cover the entire human virome, thus providing the most comprehensive pan-viral antibody reactome assay available.

**2M+**
viral protein sequences

**94K+**
unique taxonomic IDs

**500+**
viral species




Genus-level MSA-based coordinate system

**Fig 2.** VirSIGHT library design via genus-level multiple sequence alignment (MSA). The orange bar at the bottom represents the consensus sequence of a viral protein from N- to C-terminus. Overlapping peptide tiles from three strains (A, B.1, B.2) and two species (A and B) are represented as thin lines along the MSA.

## Antibody reactome profiling with VirSIGHT is sensitive and specific

We assessed the performance of IgG-specific antibody reactome profiling with the VirSIGHT library using a set of samples which had previously been characterized using standard viral antibody testing (Table 1).

Using our antibody reactome profiling assay with the VirSIGHT library, each sample was individually compared against a complete set of mock IPs (assays with no antibody input). This comparison uses the edgeR software package for statistical testing and fold-over-background estimation.

For each sample, "hits" are defined as peptides with FDR-adjusted p-values < 0.05. Fig 3 provides total viral peptide reactivities detected in samples from individuals positive or negative for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2, Fig 3A), human immunodeficiency virus 1 (HIV-1, Fig 3B) and herpes simplex virus 1 (HSV1, Fig 3C).

Samples were run in 6x replicate as part of a robust performance qualification of the assay. The average coefficient of variation (CV) of total hits across both positive and negative sample replicates, which were processed in different plates and by different operators (Fig 3), was 7.08%.

> **MIPSA has high reproducibility and low inter- and intra-operator variability.**

| Sample Serostatus | Number | Sample Type | Validation Test |
|---|---|---|---|
| Human Immunodeficiency Virus 1 (HIV-1) Positive | 11 | Serum (11) | LFIA - Allere Determine |
| Human Immunodeficiency Virus 1 (HIV-1) Negative | 14 | Serum (14) | NAT - Roche Cobas MPX |
| Herpes Simplex Virus (HSV1) Positive | 11 | Serum (11) | CLIA - DiaSorin Liason |
| Herpes Simplex Virus (HSV1) Negative | 12 | Serum (12) | CLIA - DiaSorin Liason |
| Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) Positive | 12 | Serum (10), Plasma (2) | CLIA - DiaSorin Liason, CLIE - Beckman Access, CLIA - Diazyme |
| Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) Negative | 8 | Serum (8) | CLIA - DiaSorin Liason, CLIE - Beckman Access, CLIA - Diazyme |

**Table 1.** Seropositive and seronegative samples purchased from Precision for Medicine (Bethesda, MD). LFIA, Rapid lateral flow immunoassay. NAT, Nucleic acid test. CLIA, Chemiluminescent Immunoassay. ELFA, Enzyme-linked fluorescent assay.

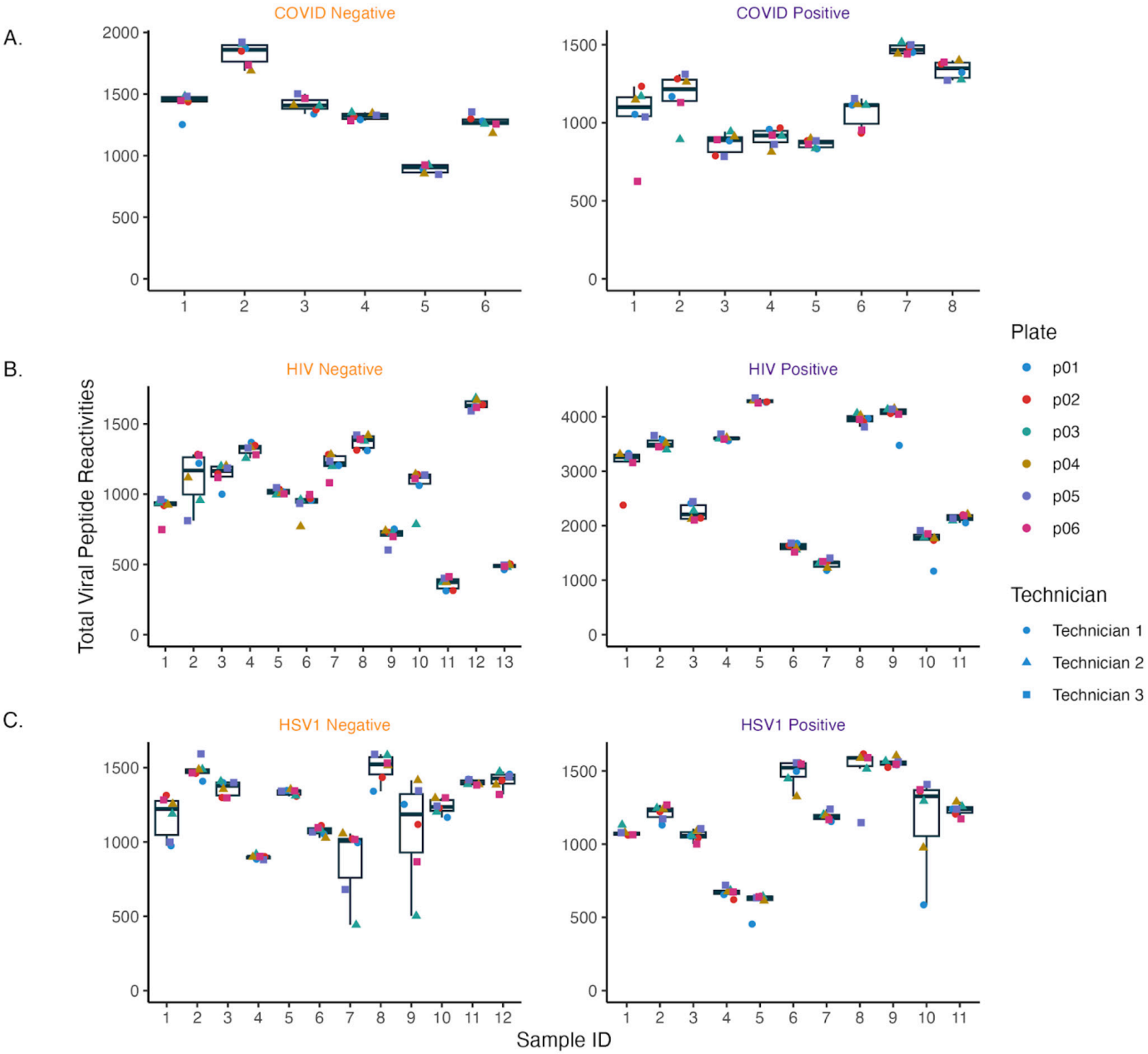
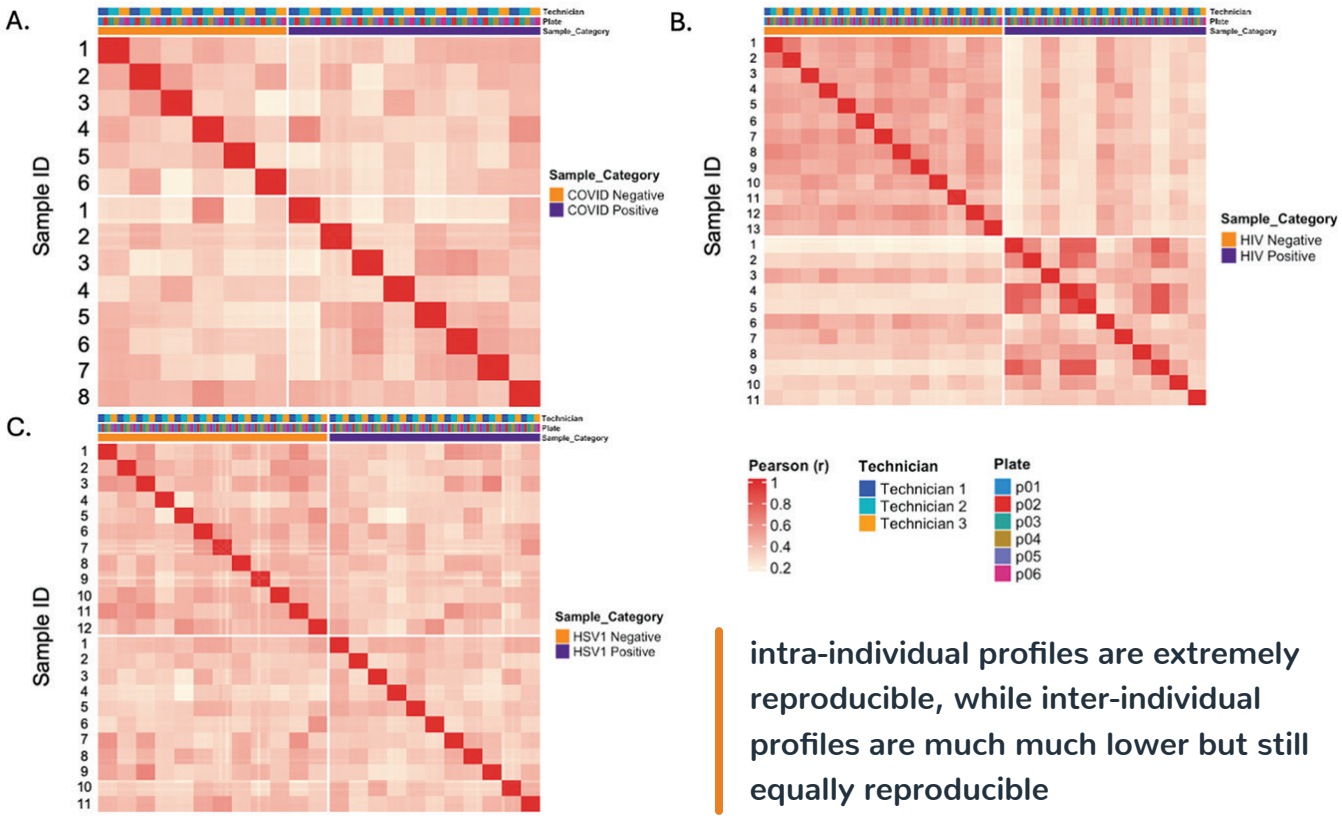



**Fig 3.** Total VirSIGHT reactivities by samples run in 6x replicate. Boxplots highlight the number of peptide reactivities for each sample across plates and operators. **(A)** SARS-CoV-2 seronegative and SARS-CoV-2 seropositive samples. **(B)** HIV-1 seronegative and HIV-1 seropositive samples. **(C)** HSV1 seronegative and HSV1 seropositive samples.

**robust reproducibly for total antibody reactivity across plates and operators**

We further assessed the Pearson correlation of the magnitude of the specific reactivities within and across all samples. Whether run on different plates or by different technicians, these $R^2$ values were close to 1 (average of 0.989). While each individual's VirSIGHT antibody reactome is highly reproducible across replicas, it is very unique compared to different individuals, observable in the checkerboard pattern seen in the correlation matrix of Fig 4 for SARS-CoV-2 (Fig 4A), HIV-1 (Fig 4B) and HSV1 (Fig 4C) samples.




**intra-individual profiles are extremely reproducible, while inter-individual profiles are much much lower but still equally reproducible**

**Fig 4.** Pearson correlation coefficient values of all -log10($p^{adj}$) peptide reactivity values across all samples. **(A)** Correlation matrix depicting pan-viral reactivity correlations across SARS-CoV-2 seronegative and SARS-CoV-2 seropositive samples, **(B)** HIV-1 seronegative and HIV-1 seropositive samples, and **(C)** HSV1 seronegative and HSV1 seropositive samples. Metadata labels provide plate number, technician number and virus status.

## VirSIGHT has a broad dynamic range and is quantitative

The samples' EdgeR-normalized fold-over-background (FOB) value of each peptide's barcode count versus that from the mock IP controls can be considered as the strength of each detected reactivity. A higher FOB corresponds to a higher antibody titer, clonality and/or binding strength. The broad range of FOBs we report (typically 2- to >100-fold) reflects the broad dynamic range and the quantitative nature of the VirSIGHT assay.

Top peptide reactivities corresponding to the known infections are shown in the heatmaps of Fig 5. The SARS-CoV-2 spike glycoprotein and nucleocapsid proteins are clearly reactive in the SARS-CoV-2 seropositive samples, but not in the SARS-CoV-2 seronegative samples (Fig 5A). Similarly, antibodies targeting HIV-1 envelope glycoproteins and polyproteins were identified in HIV-1 seropositive samples but not in HIV-1 seronegative samples (Fig 5B).

Finally, HSV1 seropositive samples were characterized by antibodies targeting structural proteins that were not identified in the HSV1 seronegative samples (Fig 5C). While replicate data was highly reproducible, notably individual donor distinct patterns of reactivity can be observed for each virus.

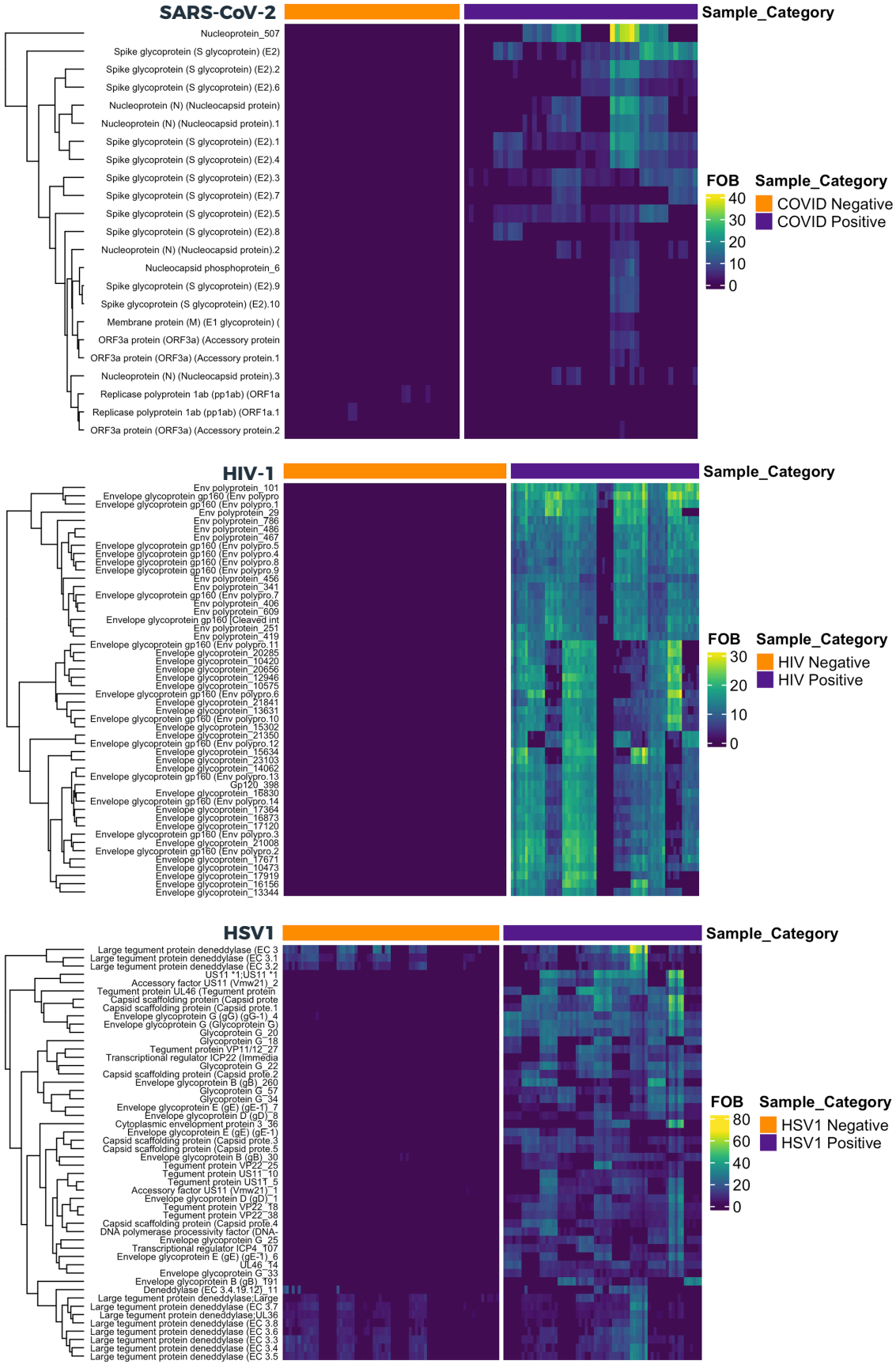



**Fig 5.** Fold Over Background (FOB) data for the top peptide reactivities corresponding to the known infections. **(A)** FOB data for SARS-CoV-2 seronegative and SARS-CoV-2 seropositive samples. **(B)** FOB data for HIV-1 seronegative and HIV-1 seropositive samples. **(C)** FOB data for HSV1 negative and HSV1 positive samples.

As a standard output of VirSIGHT analyses, we adapted the Viral Aggregate Reactivity score (VARscore) algorithm (Morgenlander et. al., 2024). Each VARscore serves as an aggregate measure of the reactivity values across all peptides associated with an individual virus at the species or strain level. Standard VARscores do not give preference or weight to any of the peptides, meaning the scores are completely unbiased.

However, weighted VARscores may be developed to preferentially weight specific peptide reactivities that are found to be more indicative of infection, while other, more cross-reactive peptides may be conversely down weighted or excluded from the score. Preferential peptide weighting depends on training data, however, meaning that weighted VARscores are inherently biased by the genetics of the population and/or the circulating viral serotypes.

It is important to keep in mind that an optimal cutoff to define VARscore seropositivity will be population and virus specific. Furthermore, VARscores can be falsely elevated due to cross-reactivity from conserved epitopes of related viruses. In the absence of a priori training data, we recommend using a default VARscore cutoff of 1 to determine if a sample is seropositive for a given virus. Using our ground truth sample labels, we used receiver operator characteristics (ROC) to determine the optimal cutoffs for each of the three viruses (Fig 6). In each case, an optimal VARscore cutoff could distinguish between all positives and negatives. These results illustrate how VirSIGHT's completely unbiased assay can identify individuals' immune responses to viral infections with high sensitivity and specificity.

**VARscore: aggregate measure of reactivity values across all peptides associated with an individual virus that is completely unbiased. This is a novel statistical measure of normalized species- or strain-level aggregate reactivity.**

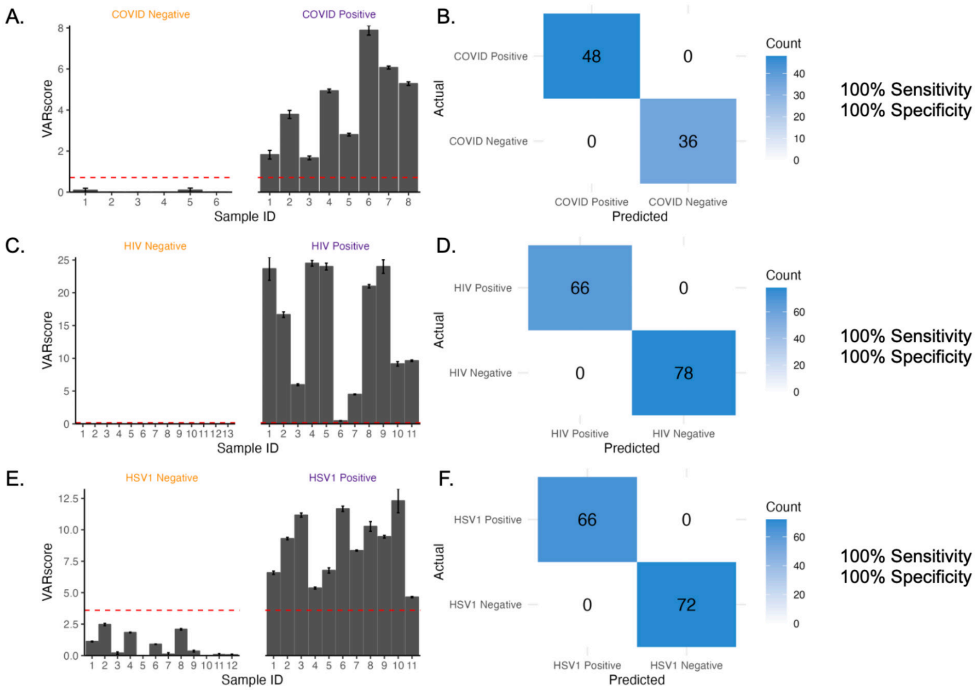



**Fig 6.** VARscore bar plots and Receiver Operating Characteristic (ROC) analysis confusion matrices. **(A)** Bar plots of VARscores for SARS-CoV-2 seronegative and SARS-CoV-2 seropositive samples. **(B)** Optimal VARscore cutoff (0.71) for SARS-CoV-2 determined using ROC analysis. **(C)** Bar plots of VARscores for HIV-1 seronegative and HIV-1 seropositive samples. **(D)** Optimal VARscore cutoff (0.17) for HIV-1 determined using ROC analysis. **(E)** Bar plots of VARscores for HSV1 seronegative and HSV1 seropositive samples. **(F)** Optimal VARscore cutoff (3.6) for HSV1 determined using ROC analysis.

The ability to map antibody reactivities onto each virus's proteome is a powerful feature of VirSIGHT data. Fig 7 illustrates the prevalence of peptide reactivities in seropositive and seronegative individuals across the viral proteomes of SARS-CoV-2, HIV-1 and HSV1.

**Powerful Discovery Capabilities: the ability to map proteome wide reactivity, which allows for mapping differences in antibody responses following vaccination versus infection; conserved reactivities across viruses for pandemic and epidemic preparedness; identification of the molecular targets of one's antibodies which helps define their functionality**

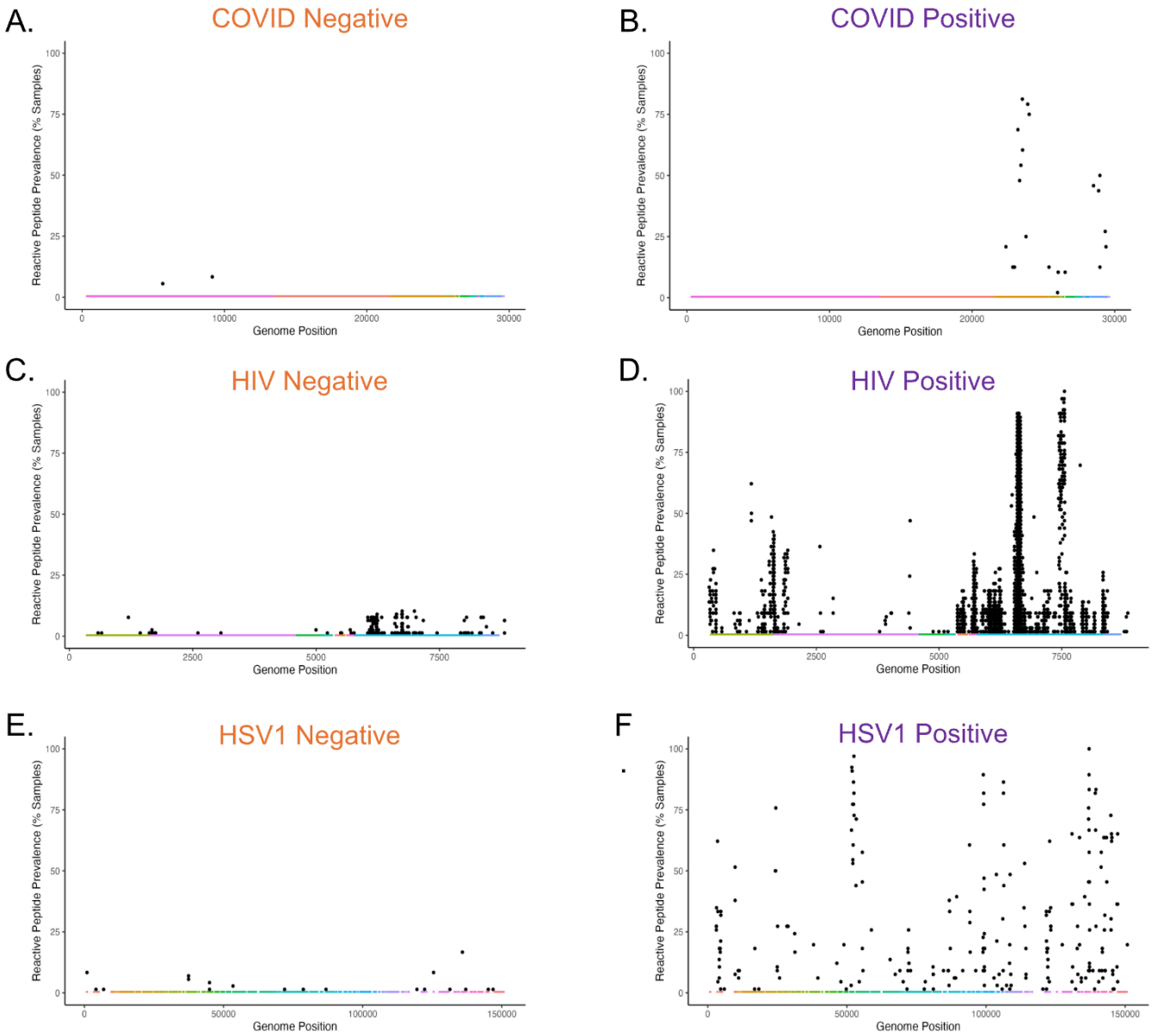



**Fig 7.** Genome-mapped peptide reactivity prevalence plots for seronegative and seropositive samples. Colored rectangles indicate discrete protein-coding sequences along the target genome. Plots are shown for **(A)** SARS-CoV-2 seronegative samples, **(B)** SARS-CoV-2 seropositive samples, **(C)** HIV-1 seronegative samples, **(D)** HIV-1 seropositive samples, **(E)** HSV1 seronegative samples and **(F)** HSV1 seropositive samples.

## MIPSA-based assays are an advance over prior generation technologies

Prior generation technologies have been used for viral antibody reactome profiling, but have suffered from significant limitations. One such approach involves immobilizing peptides or proteins on a surface to form a microarray. Various types of technical artifacts lead to false positive and false negative results (Sutandy et al., 2013), while high per-sample costs tend to limit the scope and power of the studies. Alternatively, molecular display techniques like Phage ImmunoPrecipitation Sequencing (PhIP-Seq/ VirScan) involve oligonucleotide-encoded peptide libraries and high-throughput DNA sequencing (Henson et al., 2023; Shrock et al., 2022; Xu et al., 2015). PhIP-Seq provides cost and throughput advantages over microarrays, but are limited by the size of the protein fragments that can be displayed and tends to generate false positives due to the multivalent display and proximity of the phage capsid protein.

Molecular Indexing of Proteins by Self-Assembly (MIPSA) was developed to overcome the limitations of the prior generation technologies. MIPSA libraries are in-solution DNA-barcoded peptides and proteins used to map the repertoire of antibody reactivities by antigen immunocapture and sequencing of the DNA barcodes (Credle et al., 2022). The MIPSA technology was compared to the PhIP-Seq technology in terms of library quality and assay performance. For direct comparison, the originally designed ~100,000 member 56-amino acid VirScan peptide library was constructed in both PhIP-Seq and MIPSA formats. Aliquots of each input library were amplified and sequenced to the same depth for comparison (Fig 8).

**MIPSA libraries feature uniform distributions of library members, which alleviates the false negative detection associated with other techniques with highly-skewed distributions of library members.**

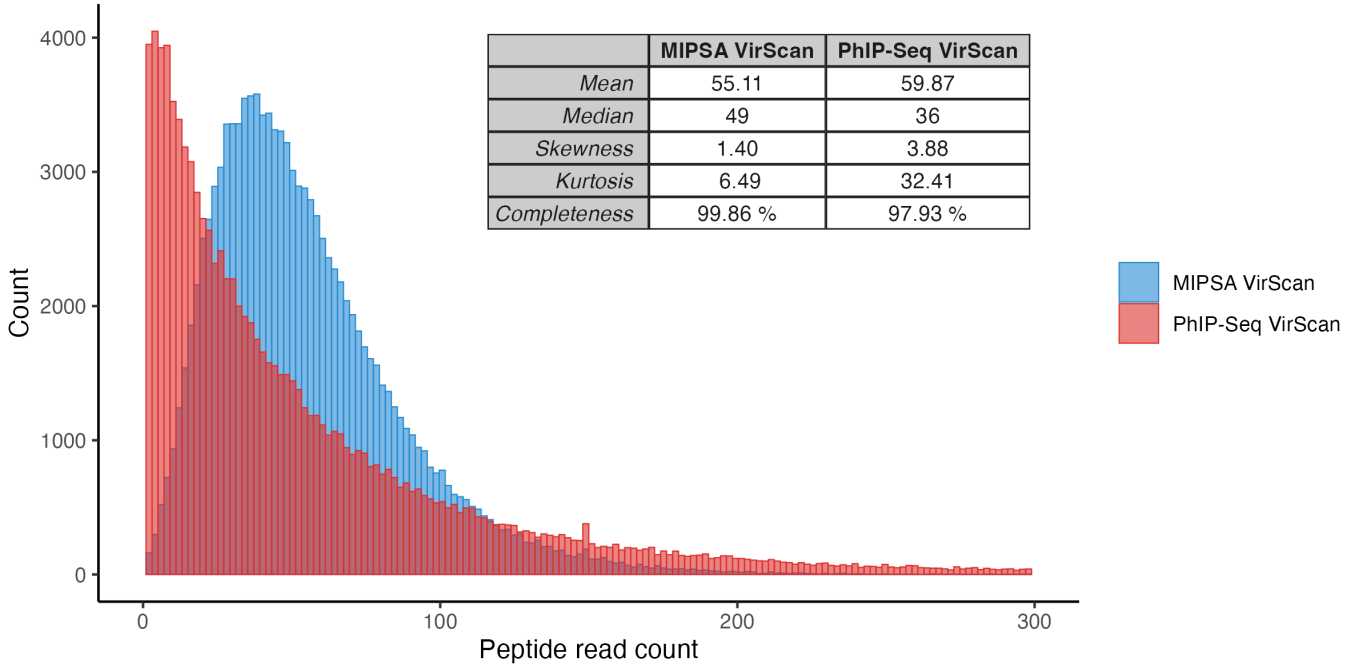



**Fig 8.** MIPSA libraries are more complete and less skewed versus PhIP-Seq libraries. Properties of the distributions, including library completeness, are provided in the table inset. Libraries were sequenced to the same depth.

In a head-to-head comparison of assay performance, we examined the number of HIV-1 peptide reactivities in samples from HIV-1 seronegative versus HIV-1 seropositive individuals and the number of HSV1 peptide reactivities in samples from HSV1 seronegative versus HSV1 seropositive individuals. This analysis revealed a preponderance of false positive reactivities detected using the PhIP-Seq assay (Fig 9A, 9B). We further examined reactivities associated with Ebola and Rabies virus peptides, which should be absent from this US-based cohort. While PhIP-Seq typically generated between 10 and 100 false positives, MIPSA detected few or no reactivities to these peptides (Fig 9C). For these comparisons, identical peptide sequences were used to create the MIPSA VirScan and PhIP-Seq VirScan libraries, samples were processed using the same workflow, and the same informatics pipeline.

**In a head-to-head comparison of MIPSA and PhIP-Seq, far fewer false positives were detected with MIPSA.**

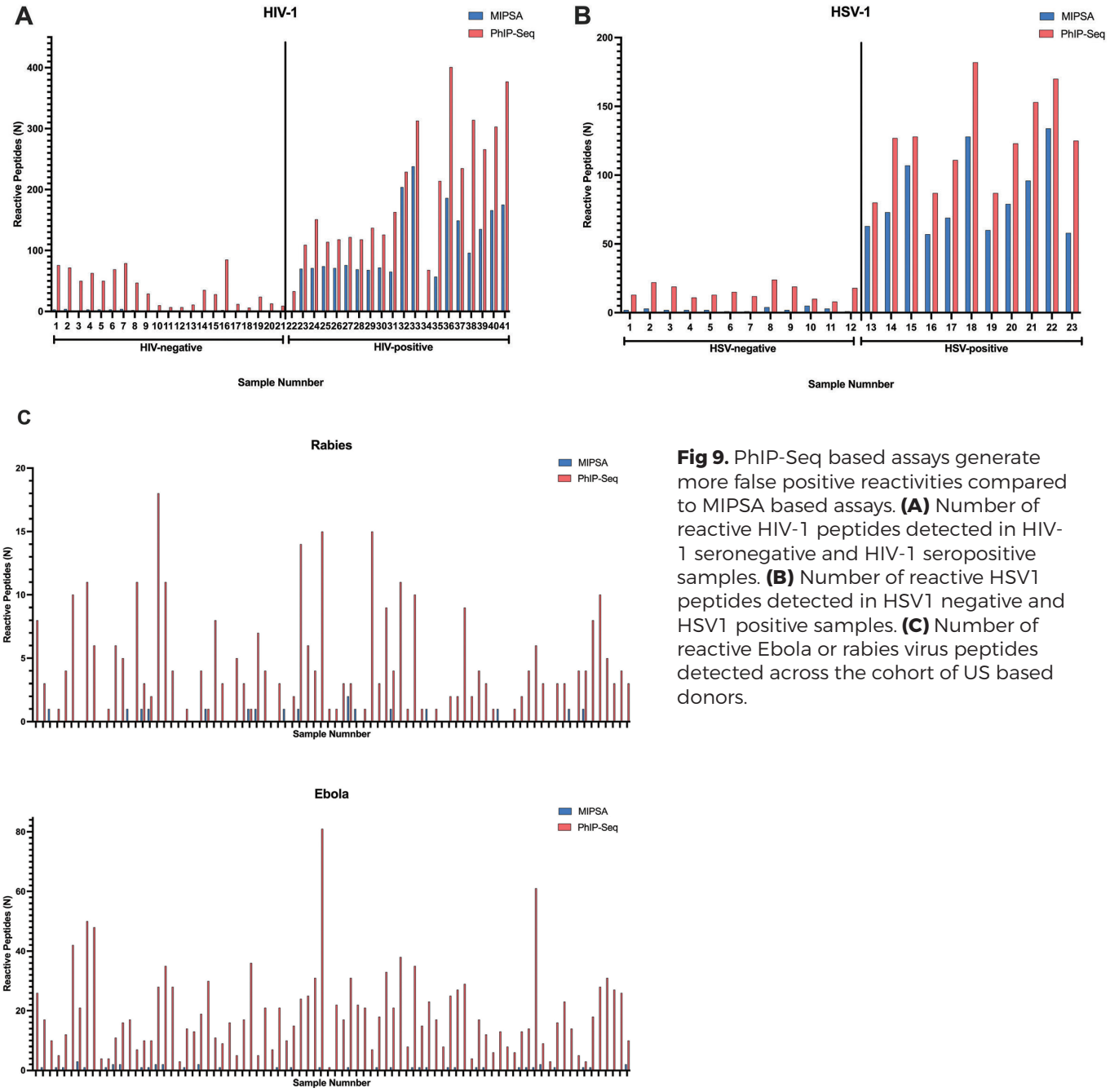
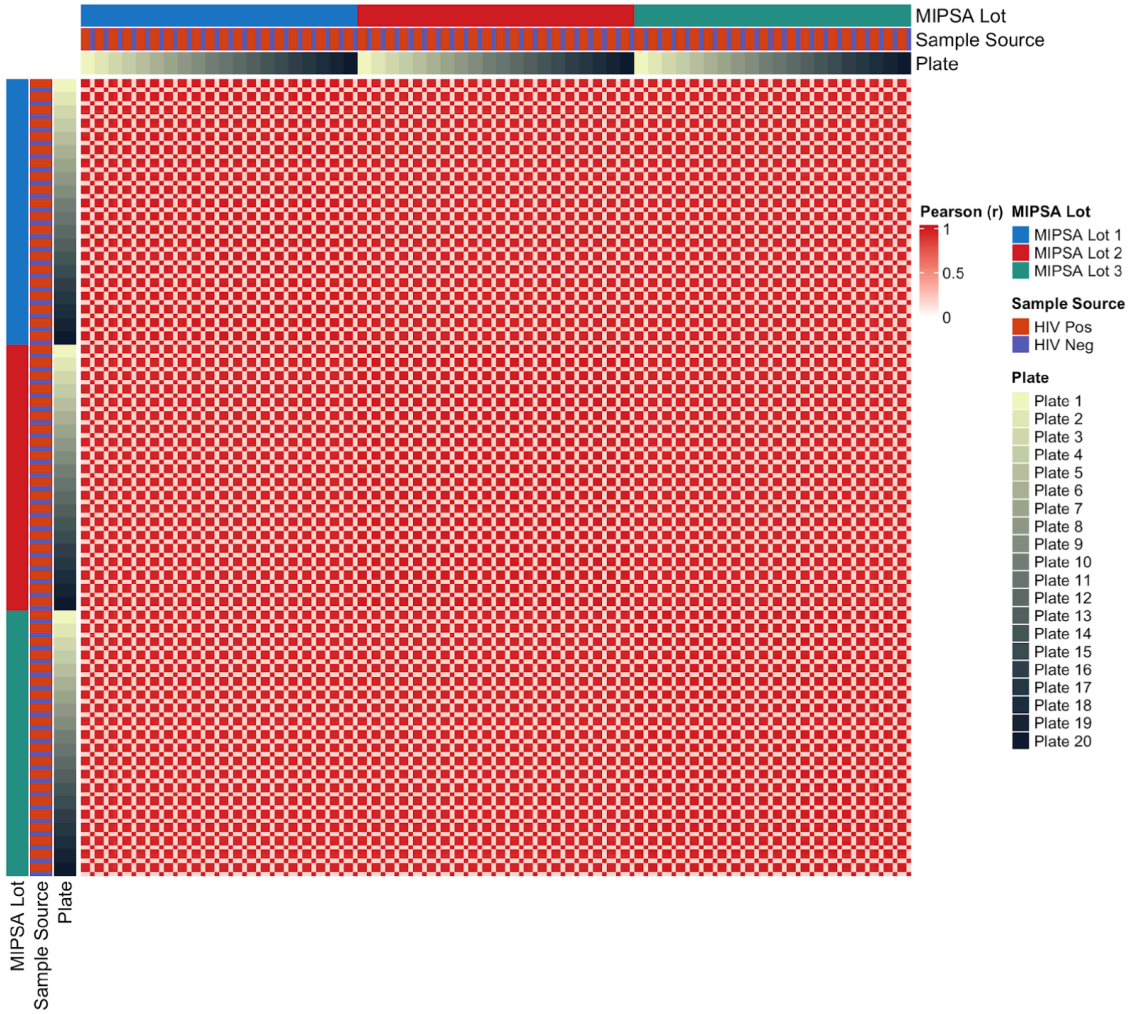



**Fig 9.** PhIP-Seq based assays generate more false positive reactivities compared to MIPSA based assays. **(A)** Number of reactive HIV-1 peptides detected in HIV-1 seronegative and HIV-1 seropositive samples. **(B)** Number of reactive HSV1 peptides detected in HSV1 negative and HSV1 positive samples. **(C)** Number of reactive Ebola or rabies virus peptides detected across the cohort of US based donors.

A key strength of the fully in vitro MIPSA technology is the extremely low lot-to-lot variability and thus the reliability of the assay performance. To assess assay reproducibility across three different MIPSA library lots, we tested the same two serum control samples using the VirSIGHT assay run across 20 different plates per library lot, for a total of 60 plates. The HIV-1 seropositive sample was run in duplicate on each plate (n=120), while the HIV-1 seronegative sample was run once per plate (n=60). The Pearson correlation coefficient for each pairwise comparison of all 180 complete VirSIGHT reactomes (comparing the peptide-level -log10($p^{adj}$) values) were calculated and visualized in the correlation matrix of Fig 10A. When comparing the HIV-1 seropositive samples to themselves, whether across plates or MIPSA library lots, the Pearson correlation coefficient was always very close to 1. Conversely, little to no correlation was observed between the profiles of the HIV-1 seropositive and seronegative samples. Finally, the average and standard deviations of the FOB values of the top 50 most significant peptide reactivities are shown for the HIV-1 seropositive individual. The average coefficient of variation (CV) of these FOB scores was 20.4% (Fig 10B).

A.



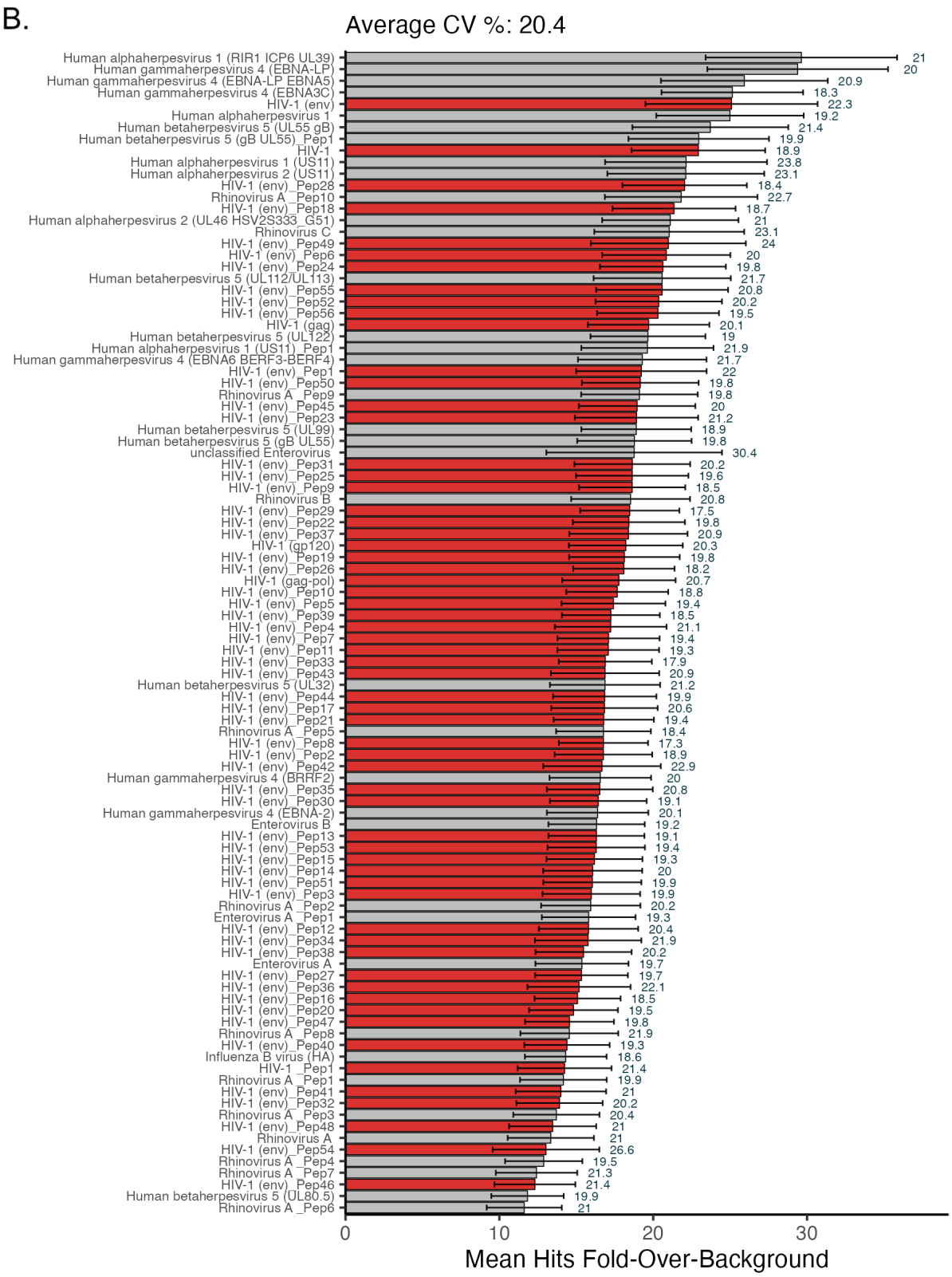
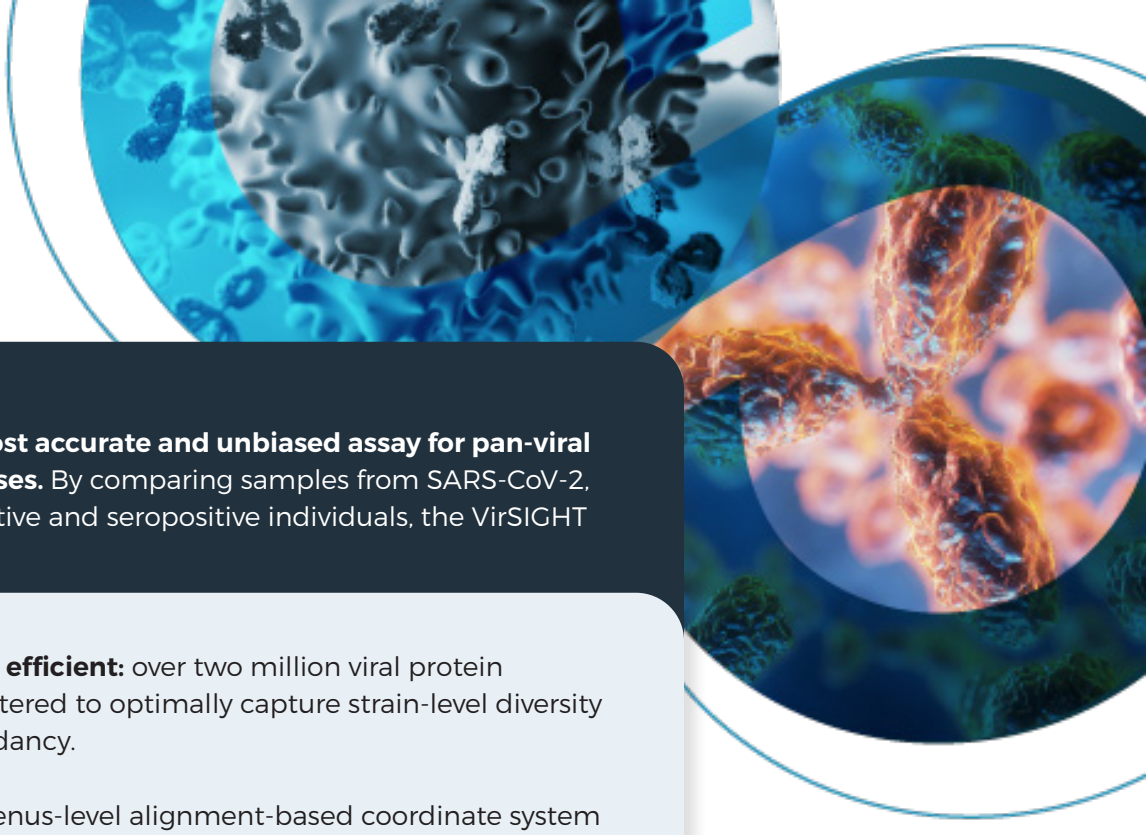



**Fig 10.** VirSIGHT assay reproducibility across 60 plates and three lots of MIPSA library using two donor serum samples. **(A)** Pearson correlation of the peptide -log10($p^{adj}$) reactivities presented as correlation matrix. 120 HIV-1 seropositive replicate profiles and 60 HIV-1 seronegative replicate profiles are included in an all-vs-all correlation comparison (n=180x180 comparisons). **(B)** The top 50 peptide reactivities by average -log10($p^{adj}$) values were selected to display their FOB mean and standard deviations across the 120 HIV-1 seropositive replicas. The average coefficient of variation (CV) for these FOBs was 20.4%.



## Conclusions

**VirSIGHT provides the most accurate and unbiased assay for pan-viral antibody reactome analyses.** By comparing samples from SARS-CoV-2, HIV-1 and HSV1 seronegative and seropositive individuals, the VirSIGHT assay was found to be:

- **Comprehensive and efficient:** over two million viral protein sequences were clustered to optimally capture strain-level diversity with minimal redundancy.
- **Interpretable:** the genus-level alignment-based coordinate system enables apples-to-apples epitope comparisons across strains and species of the same genus. Visualizing reactivities in the context of the viral genome can provide additional insight.
- **Sensitive:** completely unbiased detection of known infections at 100% sensitivity.
- **Specific:** no positive VARscores among the known seronegative samples. Minimal to no reactivities associated with known negative infections.
- **Reproducible:** highly reproducible reactivity levels (at the peptide and virus levels), with CVs of total reactivities averaging ~7% and of peptide FOBs averaging ~20%. Replicate profiles, across plates, technicians, and MIPSA library lots, are tightly correlated with Pearson $R^2$ values >0.95.
- **Superiority over PhIP-Seq:** higher quality libraries and improved assay performance, particularly with greatly reduced false positivity.

INFINITY BIO
THE ANTIBODY REACTOME COMPANY

## References


Bjornevik, K., Cortese, M., Healy, B. C., Kuhle, J., Mina, M. J., Leng, Y., Elledge, S. J., Niebuhr, D. W., Scher, A. I., Munger, K. L., & Ascherio, A. (2022). Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis. Science, 375(6578). https://doi.org/10.1126/science.abj8222

Credle, J. J., Gunn, J., Sangkhapreecha, P., Monaco, D. R., Zheng, X. A., Tsai, H. J., Wilbon, A., Morgenlander, W. R., Rastegar, A., Dong, Y., Jayaraman, S., Tosi, L., Parekkadan, B., Baer, A. N., Roederer, M., Bloch, E. M., Tobian, A. A. R., Zyskind, I., Silverberg, J. I., … Benjamin Larman, H. (2022). Unbiased discovery of autoantibodies associated with severe COVID-19 via genome-scale self-assembled DNA-barcoded protein libraries. Nature Biomedical Engineering, 6(8). https://doi.org/10.1038/s41551-022-00925-y

Henson, S. N., Elko, E. A., Swiderski, P. M., Liang, Y., Engelbrektson, A. L., Piña, A., Boyle, A. S., Fink, Z., Facista, S. J., Martinez, V., Rahee, F., Brown, A., Kelley, E. J., Nelson, G. A., Raspet, I., Mead, H. L., Altin, J. A., & Ladner, J. T. (2023). PepSeq: a fully in vitro platform for highly multiplexed serology using customizable DNA-barcoded peptide libraries. Nature Protocols, 18(2). https://doi.org/10.1038/s41596-022-00766-8

Jaroszewicz, W., Morcinek-Orłowska, J., Pierzynowska, K., Gaffke, L., & Wegrzyn, G. (2022). Phage display and other peptide display technologies. In FEMS Microbiology Reviews (Vol. 46, Issue 2). https://doi.org/10.1093/femsre/fuab052

Jayaraman S, Tiniakou E, Morgenlander WR, Na M, Christopher-Stine L, Larman HB. Comprehensive Enteroviral Serology Links Infection and Anti-Melanoma Differentiation-Associated Protein 5 Dermatomyositis. ACR Open Rheumatol. 2025 Jan;7(1):e11752. doi: 10.1002/acr2.11752. Epub 2024 Nov 7. https://doi.org/10.1002/acr2.11752

Liebhoff, A. M., Venkataraman, T., Morgenlander, W. R., Na, M., Kula, T., Waugh, K., Morrison, C., Rewers, M., Longman, R., Round, J., Elledge, S., Ruczinski, I., Langmead, B., & Larman, H. B. (2024). Efficient encoding of large antigenic spaces by epitope prioritization with Dolphyn. Nature Communications, 15(1). https://doi.org/10.1038/s41467-024-45601-8

Mitchell, M. M., Leng, Y., Boppana, S., Britt, W. J., Sanchez, L. H. G., & Elledge, S. J. (2023). Signatures of AAV-2 immunity are enriched in children with severe acute hepatitis of unknown etiology. Science Translational Medicine, 15(706). https://doi.org/10.1126/scitranslmed.adh9917

Morgenlander WR, Chia WN, Parra B, Monaco DR, Ragan I, Pardo CA, Bowen R, Zhong D, Norris DE, Ruczinski I, Durbin A, Wang LF, Larman HB, Robinson ML. Precision arbovirus serology with a pan-arbovirus peptidome. Nat Commun. 2024 Jul 11;15(1):5833. https://10.1038/s41467-024-49461-0.

Shrock, E. L., Shrock, C. L., & Elledge, S. J. (2022). VirScan: High-throughput Profiling of Antiviral Antibody Epitopes. Bio-Protocol, 12(13). https://doi.org/10.21769/BioProtoc.4464

Sutandy, F. X. R., Qian, J., Chen, C. S., & Zhu, H. (2013). Overview of protein microarrays. Current Protocols in Protein Science, SUPPL.72. https://doi.org/10.1002/0471140864.ps2701s72

Xu, G. J., Kula, T., Xu, Q., Li, M. Z., Vernon, S. D., Ndung'u, T., Ruxrungtham, K., Sanchez, J., Brander, C., Chung, R. T., O'Connor, K. C., Walker, B., Larman, H. B., & Elledge, S. J. (2015). Comprehensive serological profiling of human populations using a synthetic human virome. Science, 348(6239). https://doi.org/10.1126/science.aaa0698