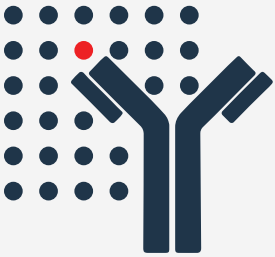




STUDY REPORT



HuProt™ Report

Human Proteome Microarray

Human Proteome Microarray Dual Channel Seromics

HuProt Study Report

Prepared For:

GC_IgG_IgA

September 29, 2025

Contents

1	Overview of HuProt	3
2	Library Creation	4
3	Experimental Overview	5
4	Informatics Methods	6
5	Core Pipeline Output Files	7
6	Processed Output Records Files	8
7	Pipeline Output Structure	9
8	Control Arrays	10
9	Array Pseudocolor Images	12
10	Array QC Summary	13
11	Control Spots	16
12	Individual Sample Z-Score Barplots	23
13	Summary of Z-score Hits	24
14	Fisher Exact Tests	26
15	Summary Heatmaps: Top Public	39
16	Summary Heatmaps: Fisher Exact Tests	40
17	UniProt BLAST Annotations	41
18	References	43

1 Overview of HuProt

Protein microarrays are a versatile and sensitive platform for unbiased immunoprofiling of the autoantibody serome. They are additionally useful for various interaction assays such as monoclonal antibody specificity evaluation, protein–protein, protein–small molecule, protein–DNA, protein–RNA, and protein–lipid interactions. The first large-scale protein microarray, the ‘ProtoArray’ was developed by Heng Zhu and Michael Snyder at Yale in 2001.¹ After becoming faculty at Johns Hopkins University, Prof. Zhu co-founded CDI Labs along with Jef Boeke, Ignacio Pino, Joe Bonaventura, Dan Eichinger, and Seth Blackshaw to create the first proteome-scale microarray: HuProt.

HuProt, The Human Proteome Microarray, contains >21,000 human proteins and isoforms - the largest collection of unique, full-length, individually-purified human proteins in a single assay platform. Compared to the discontinued ProtoArray, HuProt has 2.9x more unique amino acid content and is the first true ‘proteome-scale’ human protein microarray. HuProt arrays allow thousands of interactions to be profiled in a high-throughput manner across ~80% of the known human proteome. Full-length recombinant proteins are expressed in the yeast *S. cerevisiae*, purified via N-terminal GST-tags, and printed on glass slides in duplicate along with a set of control proteins (GST, BSA, histones, IgG, etc.). The HuProt microarray is not restricted to a particular type of surface coating, although the default surface is glass coated with ultra-thin nitrocellulose film for non-covalent, yet irreversible, capture of active proteins onto the surface.

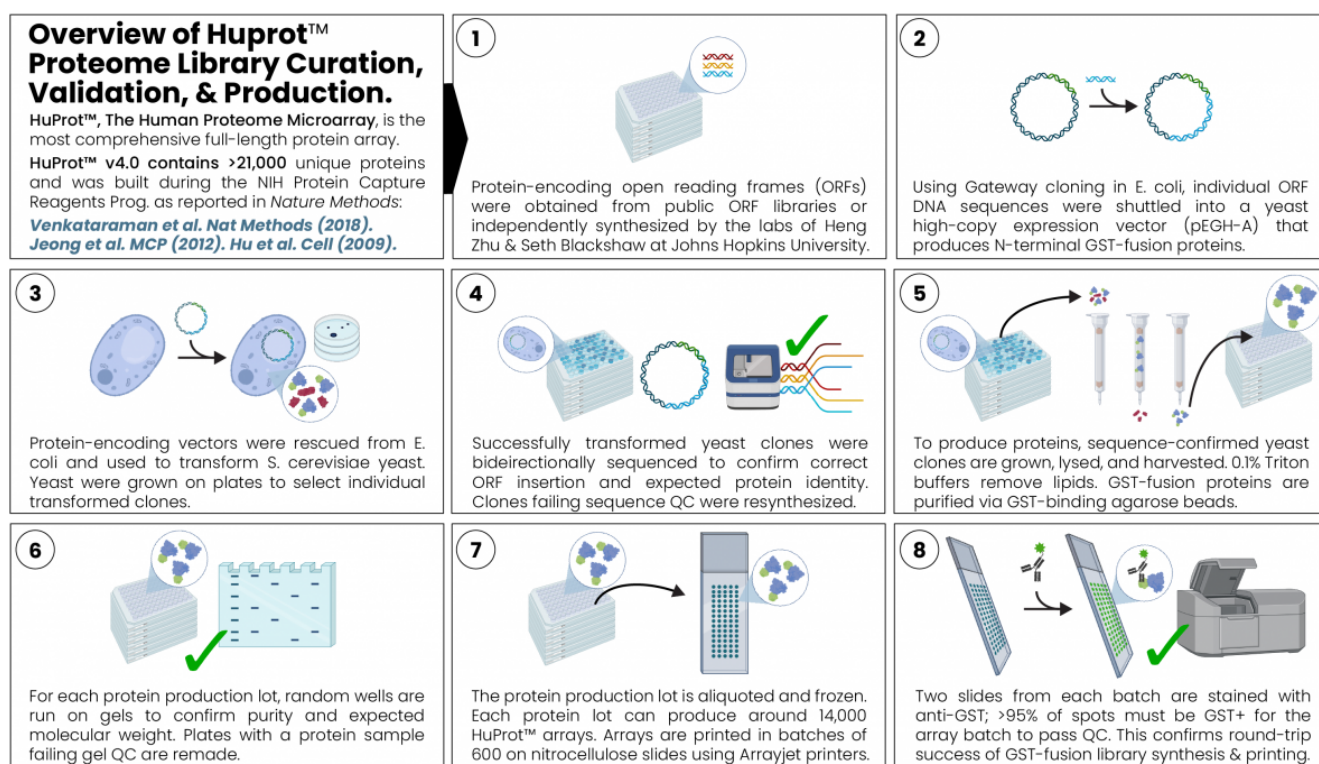


Figure by Tyler Hulett, PhD for CDI Antygen; created with BioRender.

Overview of HuProt library construction. Human proteome library creation was created as-reported: A. Venkataraman, et al. *A toolbox of immunoprecipitation-grade monoclonal antibodies to human transcription factors*, *Nat Methods*. 15 (2018) 330–338. <https://doi.org/10.1038/nmeth.4632>.

2 Library Creation

Creation of HuProt Library. HuProt library clones were derived from established open reading frame (ORF) libraries or independently synthesized; entry clones are from the laboratories of Heng Zhu and Seth Blackshaw (The Johns Hopkins University). Using the Gateway recombinant cloning system (Invitrogen, USA), human ORFs were shuttled from the entry clones to a yeast high-copy expression vector (pEGH-A) that produces GST-His6 fusion proteins under the control of the galactose-inducible GAL1 promoter. Plasmids were rescued into *E. coli* and verified by restriction endonuclease digestion. Plasmids with inserts of correct size were transformed into yeast for protein purification.^{1,2}

Validating and Curating Clones in HuProt. To check and confirm the identity of each human ORF in the HuProt library, bidirectional Sanger sequencing was conducted on both the entry clones and the yeast expression vectors that were derived from them.³ Failing clones are labeled 'ND' for 'No Data.' All QC-passing sequences have a 'JHU ID' and a gene name provided via BLAST versus UniProt. Complete BLAST results for each CDI-made clone are included in your **annotations** folder. Methods are further described in the **UniProt BLAST Annotations** section.

Protein Purification from the HuProt Library. Proteins are purified from yeast transformed with expression vectors encoding the human ORFs. Human proteins are purified as GST-His6 fusion proteins from yeast using a previously described high-throughput purification protocol (Hu S et al, 2009; Zhu et al., 2001). Using a 96-well format, the samples are purified from yeast extracts using glutathione-agarose beads. 0.1% Triton is included in the lysis buffer and washers to ensure that the purified proteins are free of lipids.

Protein Microarray Production & Testing. The purified human proteins are arrayed in a 384-well format and printed on nitrocellulose PATH slides (GraceBio, USA), using an Arrayjet UltraMarathon printer (Arrayjet, UK). Arrays that show >95% of the spots with a foreground/background signal (F/B) ratio of at least 1.5 in an anti-GST assay pass printing QA/QC. Positive and negative control spots for various staining and secondary detection reagents are additionally included. Controls include titrated GST protein, anti-IgG, histones, mouse and rabbit anti-biotin, mouse IgM, and biotin-tagged control for streptavidin detection. Each of the x20 printed blocks contains a row of control spots as well as Alexa Fluor 555/647 dye landmarks.

3 Experimental Overview

Both HuProt arrays and samples are blocked for 1 hour prior to the assay. For blocking, serum or plasma samples were diluted 1:1000 into CDISampleBuffer while arrays were incubated with CDIArrayBlock at room temperature with gentle shaking. Each blocked and diluted sample was then probed onto a HuProt microarray at room temperature for 1 hour with gentle shaking. After probing, the arrays were washed three times with TBST (1xTBS / 0.1% Tween 20) for 10 min, and then probed with secondaries at room temperature for 1 hour in a light-proof box with gentle shaking, followed by three washes with TBST for 10 min each and three rinses with ddH2O. The arrays are then dried with compressed CO2 and scanned using a GenePix 4000B scanner for data collection. Scanned .tif images are aligned to the array layout using a .gal (GenePix layout) file. Raw fluorescence intensity values are extracted using GenePix software to create .gpr (GenePix results) files. Raw .gpr files are the start of all downstream data processing.

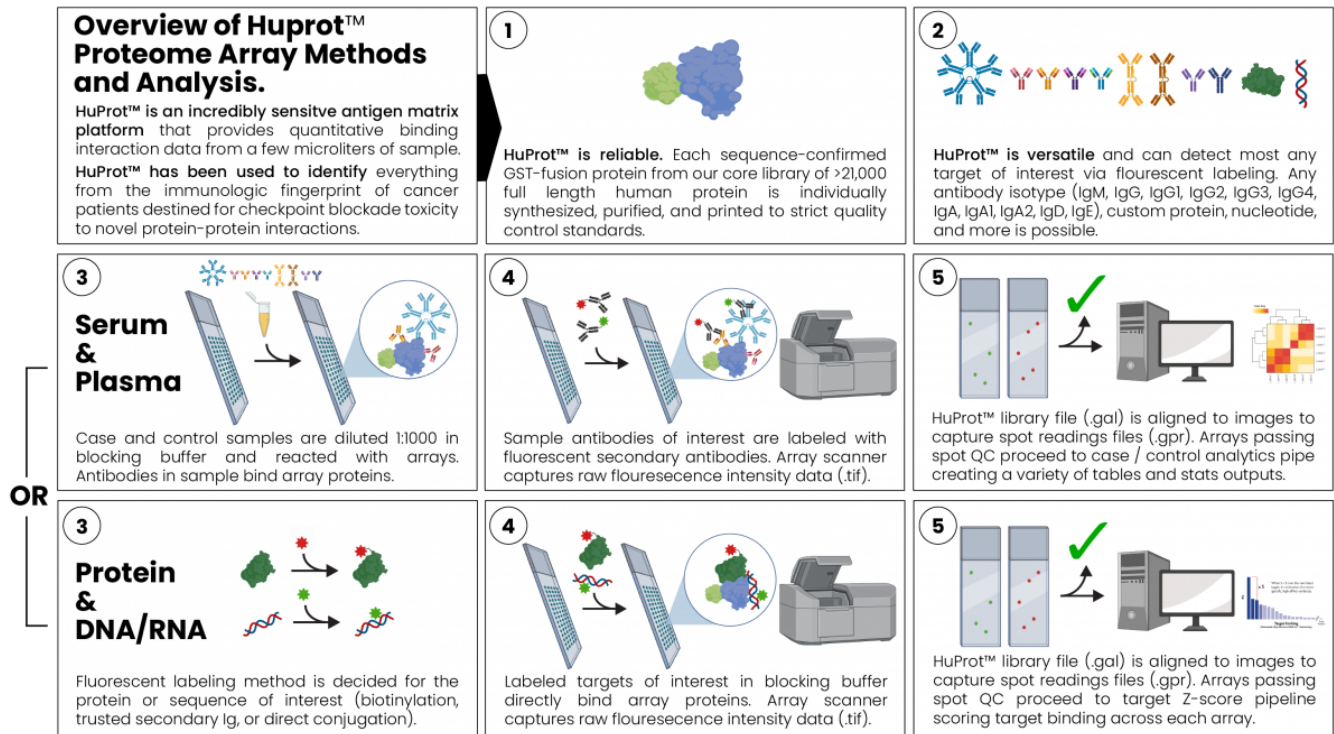


Figure by Tyler Hulett, PhD for CDI Antygen; created with BioRender.

Overview of typical HuProt assay types. Specific secondary stains and other modifications are performed according to client request.

4 Informatics Methods

Raw .gpr files from study arrays are pre-processed alongside secondary-only (SecondaryOnly) and anti-GST (GST) control arrays.

The median foreground fluorescence intensity (MFI) of each replicate spot pair in the .gpr file is averaged to produce a matrix of raw MFI readings for each staining channel, one column per array. This creates stitched **RawMFI** files which are used to generate this report. Median values which diverge by more than two-fold between replicate spots are recorded as 'outs' and used to catch misaligned array data extraction; all CDI-processed arrays have <1% of spot pairs called as 'outs' and this is our primary staining and alignment QC metric for individual arrays.

Median values are used for analysis rather than average values. Unlike older pin microarray printers, CDI's use of piezoelectric protein printing ensures each printed spot is relatively homogeneous. Median values therefore can extract accurate readings even in instances of imperfect array alignment or microscopic dye artifacts, situations which affect average readings. We do not subtract or recommend subtracting the 'background' signals recorded in the .gpr files because occasionally the amount of printed protein will exceed the binding capacity of the nitrocellulose array surface; this occasionally creates real-signals in the 'background' region. Additionally, the background of a well-stained HuProt array is small enough as to be irrelevant to most analyses, typically 40-100 MFI. Real signals are much larger and range up to 65,536 MFI, the maximum-recordable signal in a 16-bit microarray scanner grayscale .tif image.

Downstream processing and reporting is primarily created from log2 transformed raw data. While patient samples vary in signal intensity array to array, background signals are typically so low - and real signals so much higher - that a simple log2 transformation of the raw data is often adequate to stabilize variance across cohorts. In our experience, other methods such as quantile normalization are more likely to create false-positive signals than a direct analysis of log2 transformed raw data because such normalization methods artificially enforce similarity across samples which do not in fact share similar distributions. A log2 transform is performed on the raw data for each output channel which transforms the range of raw MFI readings to a log2 range of 0 to 16, each increase of 1 log2 MFI representing a doubling of the overall signal. This creates **RawLog2MFI** files.

Before Z-Scores are calculated, secondary-only (SecondaryOnly) array readings are used to call and remove problematic 'caveat' spots with false-positive readings >4-fold ($\log_2 > 2$) above the total secondary-only array median. Spots labeled ND which failed sequencing QC and control spots are additionally removed. Filtered raw log2 MFI data are recorded in the **processed_outputs** folder as **RawLog2MFI_Filtered** files. These filtered files are then used to compute Z-scores within each sample array for all non-caveat spots and are recorded as **RawLog2MFI_ZScore** files.

The Z-score output as well as the raw log2 MFI readings are then used to call staining-channel-specific 'hits.' Pipeline-called hits are >4-fold ($\log_2 > 2$) above the secondary-only median reading for that specific spot and receive a Z-score >3 within that individual sample array. The secondary-only filtering prevents staining background variance from creating false hits and the Z-score threshold ensures the spot is a real outlier within that sample and not just from an array with higher staining background. Downstream figures and analyses are then either created with this binary hit status (1/0), semi-continuous Z-score hits data representing the Z-scores of all passing hits, or directly from the raw log2 MFI data. The Z-score of all passing-hits are recorded as **RawLog2MFI_ZScore_Hits** files.

5 Core Pipeline Output Files

These are the most important files produced from HuProt microarray data according to the pipeline described previously. These are in the root directory.

gpr (folder)

This folder contains raw data from the microarray scanner. One .gpr (GenePix results) file is generated for each stained array. Only the median-foreground values are extracted and used for downstream analysis. Contains data for both 635nm and 532nm channels even if one channel is unused. More about .gpr files can be found on the Molecular Devices website:

<https://support.moleculardevices.com/s/article/GenePix-File-Formats>

RawMFI_635nm_red.csv ; RawMFI_532nm_green.csv

The linear spot-pair averaged data in its most raw format. One column per study sample generated from its .gpr file; median foreground fluorescence intensity (MFI) is summarized as an average for each spot pair. This is created from the F635 Median and F532 Median columns of the .gpr file.

RawLog2MFI_635nm_red.csv ; RawLog2MFI_532nm_green.csv

A direct log2 transform of the linear spot-pair averaged data in its most raw format. One column per study sample generated from its .gpr file; median foreground fluorescence intensity (MFI) is summarized as an average for each spot pair. This log2 transformation reduces sample to sample variance and is the recommended starting point for analysis in place linear raw data or other normalization methods.

Rawlog2MFI_ZScore_Hits_635nm_red.csv ; Rawlog2MFI_Zscore_Hits_532nm_green.csv

Z-score readings for all pipeline-called hits; non-passing values set to 0. This semi-continuous hits file is created from both the filtered raw log2 MFI data and Z-scores. All sample hits are >4x above raw secondary-only readings for that specific spot; hits also require a Z-score >3 across all non-caveat log2 transformed data from that individual array.

6 Processed Output Records Files

These are important processing records and intermediate files but are generally less-useful for starting from for custom analysis. These are in the **processed_outputs** subfolder.

Top_Public_635nm_red.csv ; Top_Public_532nm_green.csv

The most frequent publicly reactive proteins across the sample cohort; this file is output even for blinded studies. For each row in the *RawLog2MFI_ZScore_Hits* file, the total hit frequency is calculated and sorted in descending order. Average log2 MFI of hits is used to break ties. These are used to create unbiased figures in the **summary_heatmaps** folder labeled as *Top_Public*.

RemovedCaveats_635nm_red.csv ; RemovedCaveats_532nm_green.csv

A record of the spots removed to create the filtered raw data. These spots are removed prior to Z-score analysis. They are sequencing failure ND (no data) spots, controls, or assay-specific secondary-only false-positives.

Rawlog2MFI_Filtered_635nm_red.csv ; Rawlog2MFI_Filtered_532nm_green.csv

The log2 transformed spot-pair averaged data with filtered caveats removed. This is the parent file used for hit-calling and Z-score calculations.

Rawlog2MFI_ZScore_635nm_red.csv ; Rawlog2MFI_Zscore_532nm_green.csv

Sample-specific Z-scores computed down each sample column from the filtered raw log2 MFI data. This is computed for each sample column on the caveat-filtered raw log2 MFI data. It records all Z-scores including for spots that are not 'hits.'

Group_Hit_Counts_635nm_red.csv ; Group_Hit_Counts_532nm_green.csv

Tally of pipeline-called hits for each provided group. This is a true/false tally of the number of *RawLog2MFI_ZScore_Hits_635nm_red.csv* and *RawLog2MFI_ZScore_Hits_532nm_green.csv* samples from each group which react to each protein on the array. One *Hit* and *No_Hit* column pair per provided group described in the *studyManifest*; only one pair will be present if the study groups are blinded. If more than two groups are provided, these are used for Fisher exact tests.

7 Pipeline Output Structure

All data files have a similar metadata layout.

Block: Describes which of the x20 array 'blocks' this spot-pair is included in. Corresponds to physical spot-pair location in the microarray layout.

Column: Describes which of the columns this spot-pair is included in. Corresponds to physical spot-pair location in the microarray layout.

Rows: Describes the pair of rows used to create this spot-pair-averaged value. Corresponds to physical spot-pair location in the microarray layout.

Name: For the core JHU collection, this represents our best current Gene Name for this protein sequence followed by the best UniProt ID (or NA if unavailable). Description used in the case of controls and third party antigens.

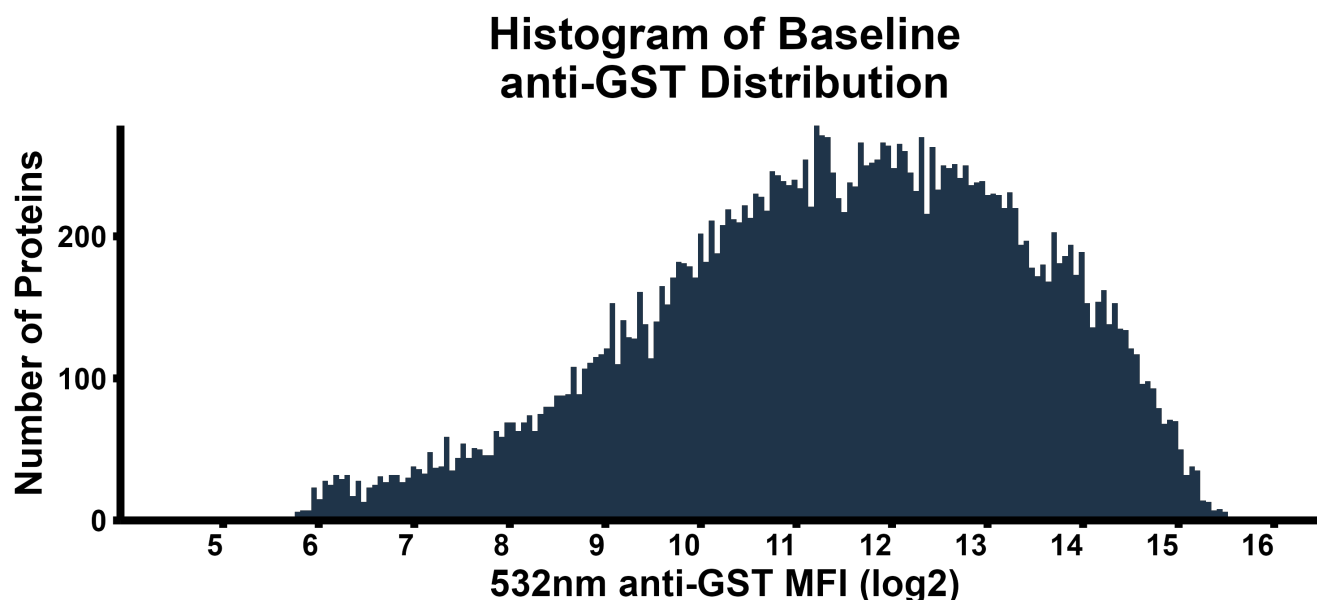
ID: The unique JHU clone identifier for this protein, status as control, third party antigen, or sequencing-fail no-data 'ND' protein. All unique JHU values correspond to a unique sequence variant or isoform - even if they correspond to the same gene name. Some literal JHU clone ID duplicates exist; while the parent ORF sequence is identical for all such entries - each are independently synthesized and purified from independent yeast clones and represent a technical replicate of our entire protein production process. JHU numbered clones are made by CDI Labs and contain a GST-tag.

Sample Columns: All subsequent columns are samples in this assay. Any column defined as 'SecondaryOnly' are staining controls used for hit calling; any column strings containing 'GST' are anti-GST QC staining results for successful protein production (most proteins in the HuProt collection contain a GST-tag). One column should exist for each sample provided to us at assay onboarding and correspond to each individual raw .gpr array file delivered. Later stages of processing (ie hits tables and Z-scores) have the control secondary-only and GST arrays removed.

8 Control Arrays

All proteins manufactured in-house by CDI Labs have a glutathione S-transferase (GST) purification tag; anti-GST stained positive control data from the array lot are provided with the experimental results. Proteins are printed as-purified and can range in concentration from 0.001 to 2.5 mg/mL. After printing, each array lot is stained for anti-GST in order to confirm successful protein synthesis and printing of each protein in the collection. Our productions COAs guarantee >95% of all spots are positive for GST above array background and typical print lots are >99% GST-positive.

Note: Due to the wide experimental variance in sample concentration and affinity affecting signal strength versus individual HuProt spot pairs, we do not typically find attempts to normalize HuProt data via anti-GST signals useful. We find anti-GST data useful for interrogating false-negatives in the assay output and as a quality control for array production.

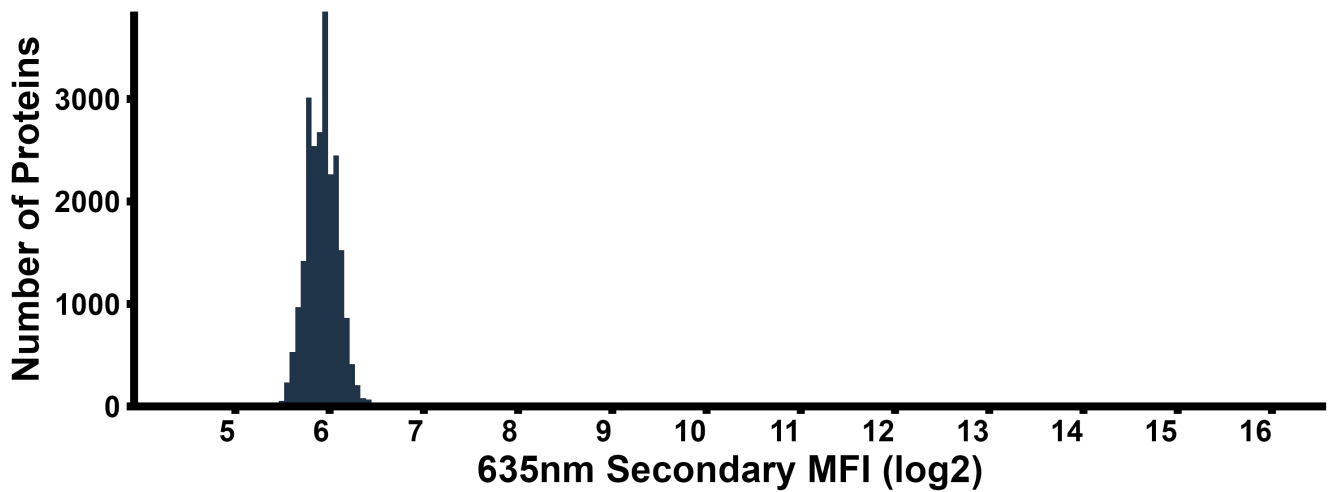


Frequency histogram of an anti-GST protein production and print QC array. This histogram plots the number of proteins present at different binned anti-GST median fluorescence intensity (MFI) readings. Created from the filtered raw log2 MFI data. Demonstrates that most proteins in the HuProt collection were successfully synthesized and printed. If multiple anti-GST arrays were included in the project this plot represents an average of all 'GST' labeled arrays labeled in the *studyManifest*.

plots/GST_histogram_plot.png

Secondary-only staining controls are important for subtracting secondary-derived false positives. These secondary-only stains are used to construct this output report.

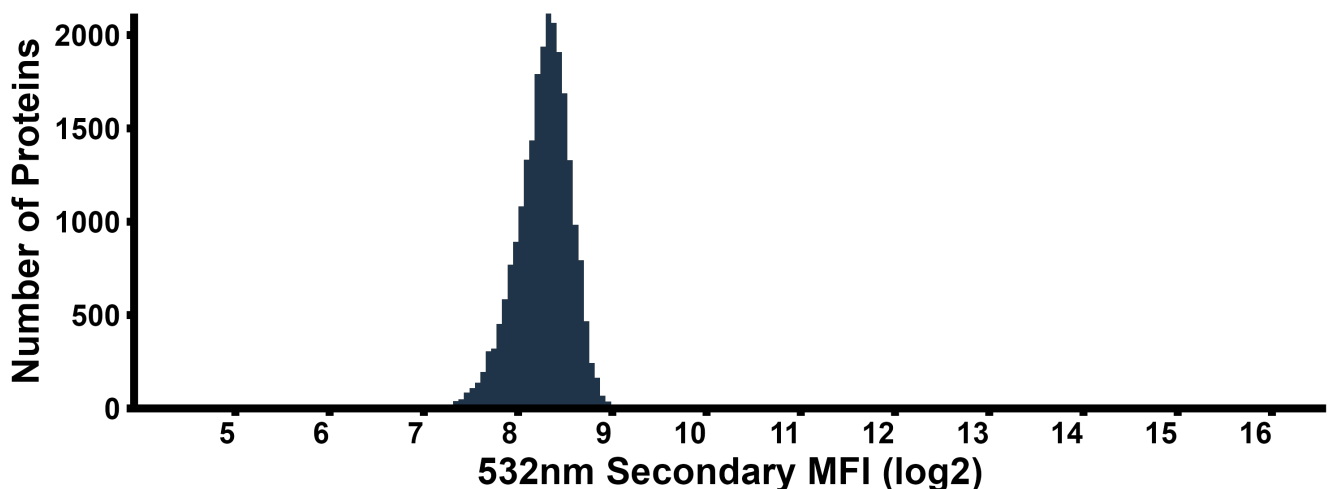
Histogram of Baseline 635nm Secondary Distribution



Frequency histogram of a 635nm secondary-only staining QC array. This histogram plots the number of proteins present at different binned secondary-only median fluorescence intensity (MFI) readings. Demonstrates a typical 'negative' background signal distribution. If multiple anti-635nm secondary-only arrays were included in the project this plot represents an average of all 'SecondaryOnly' labeled arrays labeled in the *studyManifest*.

plots/SecondaryOnly_635_histogram_plot.png

Histogram of Baseline 532nm Secondary Distribution



Frequency histogram of a 532nm secondary-only staining QC array. This histogram plots the number of proteins present at different binned secondary-only median fluorescence intensity (MFI) readings. Demonstrates a typical 'negative' background signal distribution. If multiple anti-532nm secondary-only arrays were included in the project this plot represents an average of all 'SecondaryOnly' labeled arrays labeled in the *studyManifest*.

plots/SecondaryOnly_532_histogram_plot.png

9 Array Pseudocolor Images

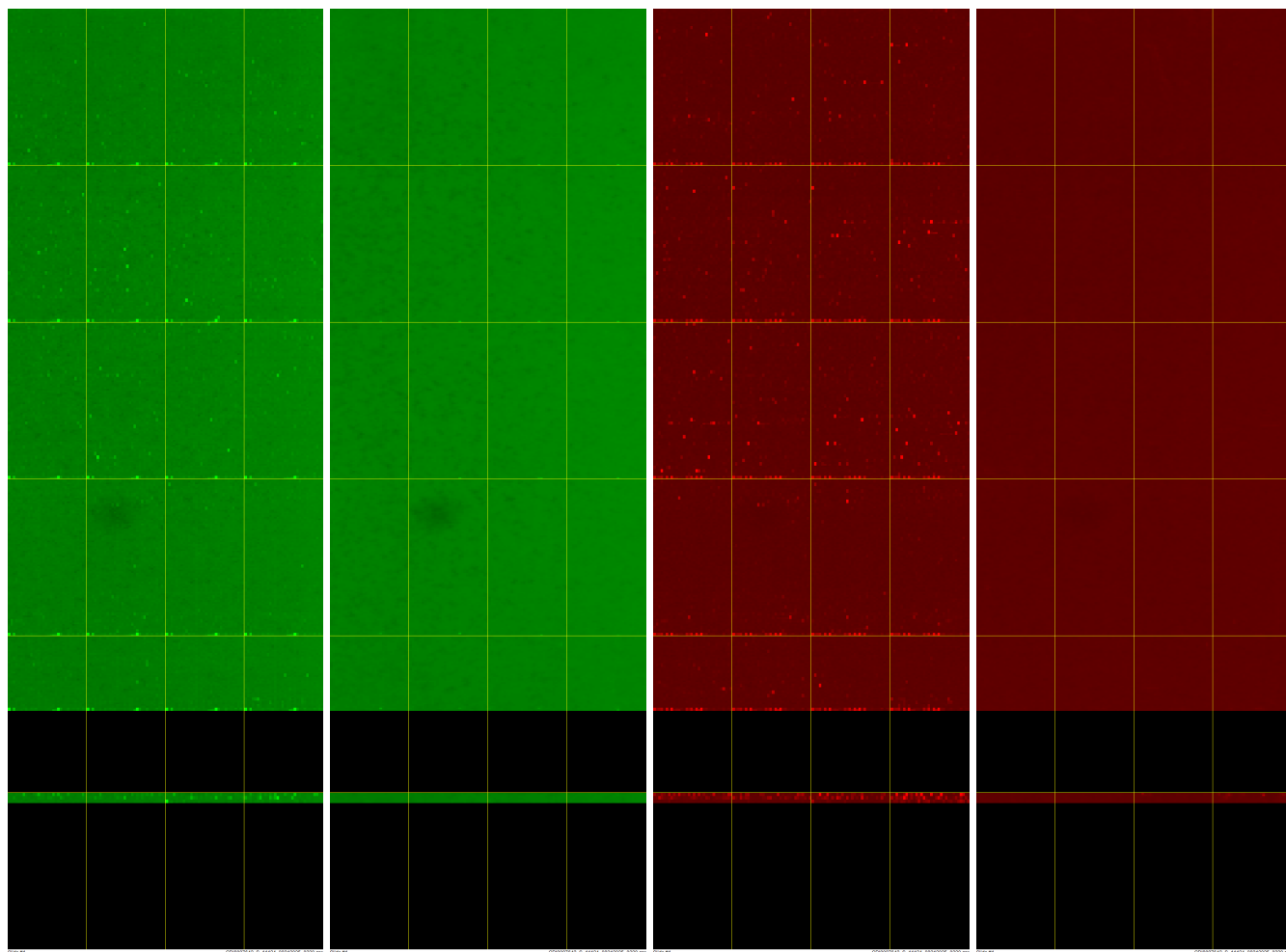
Extracted foreground and background data from individual .gpr data files are used to reconstruct pseudocolor HuProt array images according to their original layout. These images demonstrate that no large visual artifacts affect the overall study data; small visual artifacts may remain on some arrays. A set of four images are provided per study array in the **array_images** folder.

An example set of images from a single study array is shown below. **Note:** If the particular study only included one staining channel, one of these channels will be blank.

These images represent .gpr file data reconstructed into images of the:

532nm green channel (foreground F532 Median .G.png and background B532 Median .Gb.png).

635nm red channel (foreground F635 Median .R.png and background B635 Median .Rb.png).



CDI9097642_S_11124_09212025_0220.G.png
CDI9097642_S_11124_09212025_0220.Gb.png
CDI9097642_S_11124_09212025_0220.R.png
CDI9097642_S_11124_09212025_0220.Rb.png

10 Array QC Summary

To pass staining and alignment quality control, individual HuProt arrays must pass three main parameters:

- 1) **No large visual artifacts** in each staining channel by manual review.
- 2) **A correlation coefficient >0.95 across all replicate spot pairs** in the 635nm staining channel.
- 3) **Fewer than 1% of spot pairs flagged as 'outs'** having >2-fold signal divergence between replicate spot pairs.

Correlation coefficients and 'outs' are calculated using the raw foreground median fluorescence intensity of each individual raw .gpr file. The data in this report are also provided as an **Alignment_Report.csv** file in your parent directory. The following HuProt arrays were used for the study and have passed all QC parameters.

#	Array	Correlation 635nm	Outs 635nm	Correlation 532nm	Outs 532nm
1	CDI9097642_S_11124_09212025_0220.gpr	0.960	130	0.916	60
2	CDI9097695_S_11107_09212025_0198.gpr	0.952	163	0.879	79
3	CDI9097697_S_11122_09212025_0199.gpr	0.971	108	0.880	75
4	CDI9097699_S_11110_09212025_0200.gpr	0.972	121	0.853	72
5	CDI9097701_S_11095_09212025_0201.gpr	0.969	133	0.869	73
6	CDI9097703_S_11068_09212025_0202.gpr	0.967	153	0.860	90
7	CDI9097705_S_11089_09212025_0203.gpr	0.961	189	0.844	57
8	CDI9097707_S_11023_09212025_0204.gpr	0.973	94	0.918	94
9	CDI9097709_S_11103_09212025_0205.gpr	0.973	118	0.826	87
10	CDI9097711_S_11094_09212025_0206.gpr	0.972	145	0.890	73
11	CDI9097712_S_11098_09212025_0207.gpr	0.967	139	0.959	67
12	CDI9097714_S_11051_09212025_0208.gpr	0.973	113	0.853	66
13	CDI9097716_S_11061_09212025_0209.gpr	0.970	101	0.886	61
14	CDI9097718_S_11083_09212025_0210.gpr	0.973	104	0.864	54
15	CDI9097720_S_11100_09212025_0211.gpr	0.973	106	0.915	63
16	CDI9097722_S_11108_09212025_0212.gpr	0.972	140	0.827	75
17	CDI9097724_S_11118_09212025_0213.gpr	0.973	99	0.825	71
18	CDI9097726_S_11054_09212025_0214.gpr	0.971	111	0.865	65
19	CDI9097728_S_11065_09212025_0215.gpr	0.976	98	0.876	72
20	CDI9097730_S_11093_09212025_0216.gpr	0.974	108	0.858	69
21	CDI9097732_S_11074_09212025_0217.gpr	0.973	103	0.841	72
22	CDI9097734_S_11117_09212025_0218.gpr	0.973	101	0.875	63
23	CDI9097736_S_11126_09212025_0219.gpr	0.969	140	0.874	68
24	CDI9097739_S_11044_09212025_0221.gpr	0.970	95	0.927	39
25	CDI9097741_S_11046_09212025_0222.gpr	0.976	82	0.963	33
26	CDI9097743_S_11048_09212025_0223.gpr	0.974	84	0.947	31
27	CDI9097745_S_11010_09212025_0224.gpr	0.975	81	0.945	32
28	CDI9097747_S_11075_09212025_0225.gpr	0.975	81	0.947	31
29	CDI9097749_S_11085_09212025_0226.gpr	0.965	118	0.929	38
30	CDI9097751_S_11086_09212025_0227.gpr	0.970	112	0.947	40
31	CDI9097753_S_11014_09212025_0228.gpr	0.974	80	0.947	36
32	CDI9097755_S_11039_09212025_0229.gpr	0.973	91	0.936	32
33	CDI9097757_S_11088_09212025_0230.gpr	0.975	90	0.947	32
34	CDI9097759_S_11011_09212025_0231.gpr	0.975	83	0.935	35
35	CDI9097761_S_11067_09212025_0232.gpr	0.975	107	0.932	30
36	CDI9097762_S_11069_09212025_0233.gpr	0.977	80	0.939	38
37	CDI9097764_S_11113_09212025_0234.gpr	0.977	75	0.931	31
38	CDI9097766_S_11058_09212025_0235.gpr	0.974	56	0.952	30
39	CDI9097768_S_11008_09212025_0236.gpr	0.977	70	0.941	26
40	CDI9097772_S_11090_09212025_0238.gpr	0.978	62	0.951	30
41	CDI9097774_S_11125_09212025_0239.gpr	0.978	46	0.941	28
42	CDI9097776_S_11097_09212025_0240.gpr	0.981	42	0.955	25
43	CDI9097787_S_13046_09252025_0009.gpr	0.968	190	0.918	65
44	CDI9097789_S_13168_09252025_0015.gpr	0.971	146	0.839	83
45	CDI9097791_S_13023_09252025_0102.gpr	0.975	93	0.815	72
46	CDI9097793_S_13114_09252025_0114.gpr	0.977	68	0.924	23
47	CDI9097795_S_12013_09252025_0123.gpr	0.976	73	0.954	28
48	CDI9097797_S_12044_09252025_0157.gpr	0.975	87	0.884	52
49	CDI9097799_S_11119_09252025_0237.gpr	0.972	68	0.880	46
50	CDI9097891_S_13131_09182025_0001.gpr	0.976	50	0.912	14
51	CDI9097893_S_13136_09182025_0002.gpr	0.982	25	0.958	15
52	CDI9097895_S_13169_09182025_0003.gpr	0.985	18	0.936	8
53	CDI9097897_S_13001_09182025_0004.gpr	0.989	15	0.882	7
54	CDI9097899_S_13070_09182025_0005.gpr	0.981	22	0.922	8
55	CDI9097901_S_13075_09182025_0006.gpr	0.979	30	0.879	14
56	CDI9097902_S_12025_09212025_0180.gpr	0.980	38	0.893	14
57	CDI9097903_S_13146_09182025_0007.gpr	0.976	26	0.897	6
58	CDI9097904_S_11015_09212025_0181.gpr	0.980	40	0.895	9
59	CDI9097905_S_13162_09182025_0008.gpr	0.977	43	0.917	15
60	CDI9097906_S_11040_09212025_0182.gpr	0.982	26	0.887	21
61	CDI9097909_S_13110_09182025_0010.gpr	0.981	27	0.880	3
62	CDI9097911_S_13152_09182025_0011.gpr	0.981	26	0.888	7
63	CDI9097913_S_13005_09182025_0012.gpr	0.979	26	0.897	9
64	CDI9097915_S_13072_09182025_0013.gpr	0.980	23	0.877	7
65	CDI9097916_S_13145_09182025_0014.gpr	0.967	127	0.850	60

(continued)

#	Array	Correlation 635nm	Outs 635nm	Correlation 532nm	Outs 532nm
66	CDI9097920_S_13049_09182025_0016.gpr	0.965	129	0.877	21
67	CDI9097922_S_13111_09182025_0017.gpr	0.965	117	0.929	36
68	CDI9097924_S_13171_09182025_0018.gpr	0.970	88	0.896	4
69	CDI9097926_S_13076_09182025_0019.gpr	0.968	93	0.921	13
70	CDI9097928_S_13095_09182025_0020.gpr	0.972	114	0.851	25
71	CDI9097930_S_13115_09182025_0021.gpr	0.967	94	0.942	0
72	CDI9097932_S_13129_09182025_0022.gpr	0.970	129	0.926	3
73	CDI9097934_S_13143_09182025_0023.gpr	0.966	108	0.926	11
74	CDI9097936_S_13144_09182025_0024.gpr	0.968	106	0.927	3
75	CDI9097938_S_13153_09182025_0025.gpr	0.973	246	0.953	13
76	CDI9097940_S_13022_09182025_0026.gpr	0.966	108	0.863	11
77	CDI9097941_S_13026_09182025_0027.gpr	0.959	187	0.915	6
78	CDI9097943_S_13050_09182025_0028.gpr	0.966	134	0.931	13
79	CDI9097944_S_11055_09212025_0183.gpr	0.965	166	0.920	67
80	CDI9097945_S_13105_09182025_0029.gpr	0.966	100	0.951	10
81	CDI9097946_S_11006_09212025_0184.gpr	0.966	138	0.924	53
82	CDI9097947_S_13135_09182025_0030.gpr	0.966	213	0.923	42
83	CDI9097948_S_11102_09212025_0185.gpr	0.966	105	0.937	53
84	CDI9097949_S_13016_09182025_0031.gpr	0.966	117	0.906	7
85	CDI9097950_S_11019_09212025_0186.gpr	0.963	141	0.939	57
86	CDI9097951_S_13083_09182025_0032.gpr	0.967	137	0.921	14
87	CDI9097952_S_11071_09212025_0187.gpr	0.964	126	0.921	57
88	CDI9097953_S_13106_09182025_0033.gpr	0.963	104	0.938	4
89	CDI9097954_S_11047_09212025_0188.gpr	0.963	110	0.893	52
90	CDI9097955_S_13156_09182025_0034.gpr	0.965	164	0.920	31
91	CDI9097956_S_11080_09212025_0189.gpr	0.966	152	0.822	109
92	CDI9097957_S_13002_09182025_0035.gpr	0.967	160	0.940	8
93	CDI9097958_S_11020_09212025_0190.gpr	0.965	176	0.912	62
94	CDI9097959_S_13038_09182025_0036.gpr	0.966	148	0.938	11
95	CDI9097960_S_11104_09212025_0191.gpr	0.965	148	0.912	48
96	CDI9097961_S_13058_09182025_0037.gpr	0.964	120	0.945	7
97	CDI9097962_S_11025_09212025_0192.gpr	0.972	135	0.909	56
98	CDI9097963_S_13113_09182025_0038.gpr	0.959	104	0.882	21
99	CDI9097964_S_11028_09212025_0193.gpr	0.969	92	0.910	48
100	CDI9097965_S_13007_09182025_0039.gpr	0.963	107	0.910	21
101	CDI9097966_S_13041_09182025_0040.gpr	0.963	114	0.914	15
102	CDI9097968_S_13054_09182025_0041.gpr	0.968	161	0.935	17
103	CDI9097970_S_13056_09182025_0042.gpr	0.970	80	0.944	3
104	CDI9097972_S_13065_09182025_0043.gpr	0.962	72	0.902	20
105	CDI9097974_S_13085_09182025_0044.gpr	0.965	107	0.918	21
106	CDI9097976_S_13100_09182025_0045.gpr	0.964	80	0.929	23
107	CDI9097978_S_13163_09182025_0046.gpr	0.961	108	0.915	23
108	CDI9097980_S_13025_09182025_0047.gpr	0.966	95	0.925	32
109	CDI9097982_S_13052_09182025_0048.gpr	0.966	136	0.885	40
110	CDI9097984_S_13061_09182025_0049.gpr	0.966	151	0.880	47
111	CDI9097986_S_13063_09182025_0050.gpr	0.966	125	0.894	58
112	CDI9097988_S_13125_09182025_0051.gpr	0.967	99	0.877	42
113	CDI9097990_S_13133_09182025_0052.gpr	0.966	135	0.869	41
114	CDI9097991_S_13158_09182025_0053.gpr	0.963	100	0.867	86
115	CDI9097992_S_11035_09212025_0194.gpr	0.968	121	0.943	164
116	CDI9097993_S_13164_09182025_0054.gpr	0.962	140	0.843	63
117	CDI9097994_S_11057_09212025_0195.gpr	0.967	127	0.873	79
118	CDI9097995_S_13010_09182025_0055.gpr	0.964	179	0.863	73
119	CDI9097996_S_11034_09212025_0196.gpr	0.965	213	0.891	85
120	CDI9097997_S_13015_09182025_0056.gpr	0.964	166	0.861	57
121	CDI9097998_S_11018_09212025_0197.gpr	0.969	131	0.843	73
122	CDI9097999_S_13019_09182025_0057.gpr	0.960	174	0.848	68
123	CDI9098001_S_13034_09182025_0058.gpr	0.958	101	0.794	154
124	CDI9098003_S_13090_09182025_0059.gpr	0.966	116	0.808	174
125	CDI9098005_S_13006_09182025_0060.gpr	0.964	110	0.781	159
126	CDI9098011_S_13017_09192025_0061.gpr	0.970	130	0.833	90
127	CDI9098013_S_13031_09192025_0062.gpr	0.968	134	0.795	78
128	CDI9098015_S_13081_09192025_0063.gpr	0.968	102	0.811	91
129	CDI9098016_S_13120_09192025_0064.gpr	0.968	70	0.816	81
130	CDI9098018_S_13148_09192025_0065.gpr	0.968	84	0.819	107
131	CDI9098020_S_13149_09192025_0066.gpr	0.970	96	0.850	31
132	CDI9098022_S_13157_09192025_0067.gpr	0.974	106	0.861	43
133	CDI9098024_S_13116_09192025_0068.gpr	0.971	76	0.904	48
134	CDI9098026_S_13134_09192025_0069.gpr	0.972	70	0.870	32
135	CDI9098028_S_13167_09192025_0070.gpr	0.974	89	0.871	33
136	CDI9098030_S_13011_09192025_0071.gpr	0.972	89	0.887	30
137	CDI9098032_S_13013_09192025_0072.gpr	0.973	66	0.905	46
138	CDI9098034_S_13024_09192025_0073.gpr	0.974	66	0.876	31
139	CDI9098036_S_13140_09192025_0074.gpr	0.971	91	0.873	37
140	CDI9098038_S_13088_09192025_0075.gpr	0.972	66	0.912	31
141	CDI9098040_S_13097_09192025_0076.gpr	0.970	67	0.922	37
142	CDI9098041_S_13123_09192025_0077.gpr	0.965	122	0.851	54
143	CDI9098043_S_13128_09192025_0078.gpr	0.967	99	0.883	42
144	CDI9098045_S_13161_09192025_0079.gpr	0.974	74	0.886	45
145	CDI9098047_S_13018_09192025_0080.gpr	0.972	75	0.880	33
146	CDI9098049_S_13027_09192025_0081.gpr	0.970	52	0.868	31
147	CDI9098051_S_13053_09192025_0082.gpr	0.973	56	0.874	26
148	CDI9098053_S_13089_09192025_0083.gpr	0.970	84	0.841	33
149	CDI9098055_S_13029_09192025_0084.gpr	0.971	55	0.888	39
150	CDI9098057_S_13109_09192025_0085.gpr	0.972	76	0.880	37
151	CDI9098059_S_13141_09192025_0086.gpr	0.971	63	0.877	29
152	CDI9098061_S_13060_09192025_0087.gpr	0.971	76	0.841	41
153	CDI9098063_S_13069_09192025_0088.gpr	0.974	58	0.891	41
154	CDI9098065_S_13037_09192025_0089.gpr	0.976	77	0.884	34
155	CDI9098066_S_13121_09192025_0090.gpr	0.975	41	0.860	31
156	CDI9098068_S_13160_09192025_0091.gpr	0.972	80	0.855	35
157	CDI9098070_S_13033_09192025_0092.gpr	0.972	80	0.860	38
158	CDI9098072_S_13071_09192025_0093.gpr	0.972	97	0.891	42
159	CDI9098074_S_13079_09192025_0094.gpr	0.973	54	0.850	37
160	CDI9098076_S_13159_09192025_0095.gpr	0.975	54	0.849	41
161	CDI9098078_S_13064_09192025_0096.gpr	0.971	80	0.864	38
162	CDI9098080_S_13093_09192025_0097.gpr	0.972	62	0.822	43
163	CDI9098082_S_13150_09192025_0098.gpr	0.972	80	0.872	38
164	CDI9098084_S_13151_09192025_0099.gpr	0.973	91	0.865	33
165	CDI9098086_S_13165_09192025_0100.gpr	0.971	75	0.878	22
166	CDI9098088_S_13008_09192025_0101.gpr	0.977	54	0.864	37
167	CDI9098091_S_13047_09192025_0103.gpr	0.964	78	0.872	33
168	CDI9098093_S_13057_09192025_0104.gpr	0.974	37	0.819	46
169	CDI9098095_S_13094_09192025_0105.gpr	0.975	55	0.847	23
170	CDI9098097_S_13104_09192025_0106.gpr	0.976	54	0.868	24
171	CDI9098099_S_13127_09192025_0107.gpr	0.976	46	0.860	20
172	CDI9098101_S_13032_09192025_0108.gpr	0.975	48	0.868	22
173	CDI9098103_S_13092_09192025_0109.gpr	0.977	59	0.892	15
174	CDI9098105_S_13108_09192025_0110.gpr	0.977	31	0.895	22
175	CDI9098107_S_13003_09192025_0111.gpr	0.979	50	0.871	25

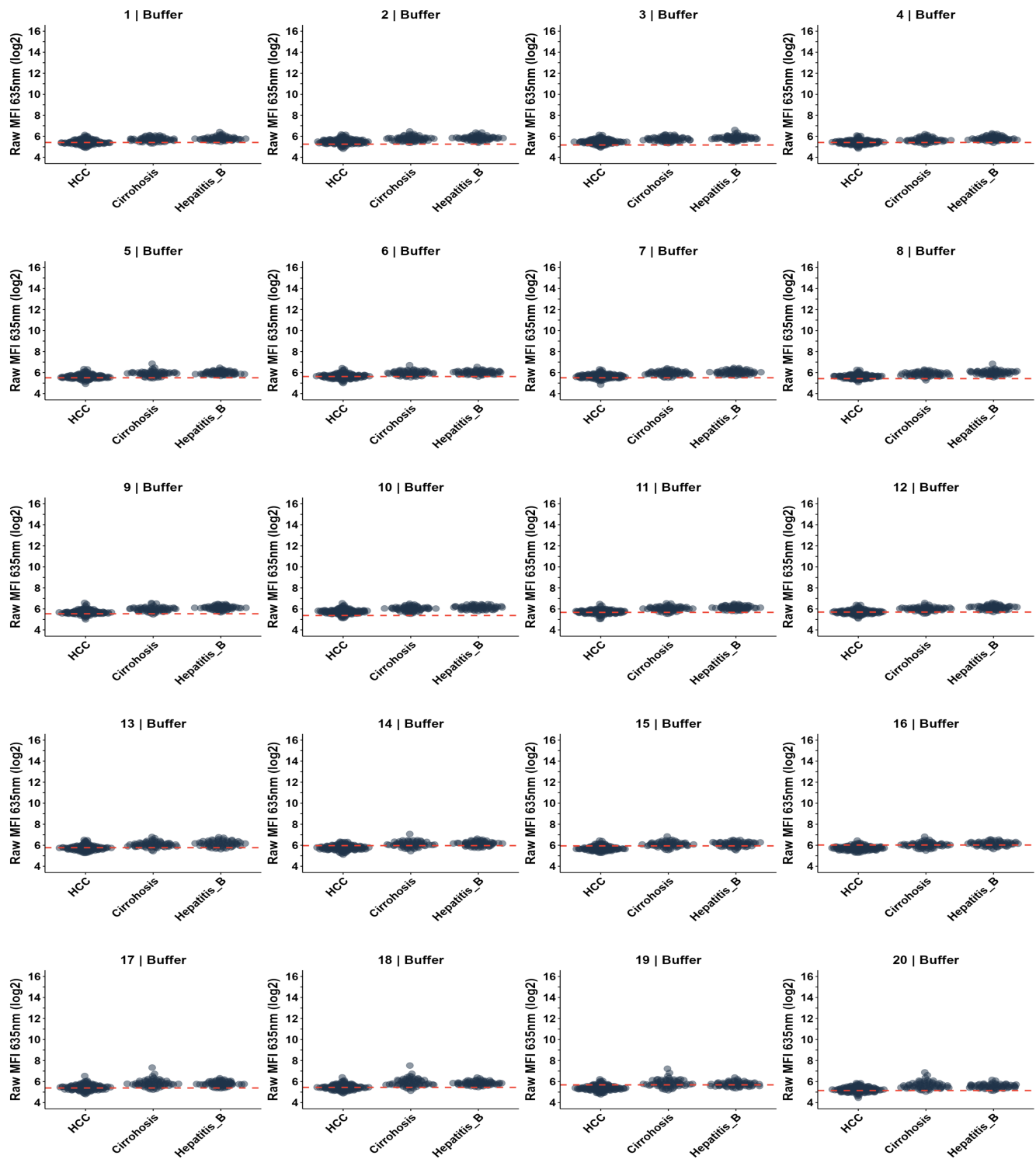
(continued)

#	Array	Correlation 635nm	Outs 635nm	Correlation 532nm	Outs 532nm
176	CDI9098109_S_13039_09192025_0112.gpr	0.975	51	0.873	23
177	CDI9098111_S_13062_09192025_0113.gpr	0.977	41	0.892	22
178	CDI9098115_S_13009_09192025_0115.gpr	0.979	54	0.884	33
179	CDI9098116_S_13130_09192025_0116.gpr	0.976	36	0.876	25
180	CDI9098118_S_13119_09192025_0117.gpr	0.978	41	0.871	21
181	CDI9098120_S_13061_09192025_0118.gpr	0.977	32	0.879	19
182	CDI9098122_S_13004_09192025_0119.gpr	0.979	37	0.906	14
183	CDI9098124_S_13059_09202025_0120.gpr	0.978	32	0.904	15
184	CDI9098126_S_12036_09202025_0121.gpr	0.978	45	0.918	32
185	CDI9098128_S_12002_09202025_0122.gpr	0.978	29	0.936	17
186	CDI9098132_S_12014_09202025_0124.gpr	0.978	20	0.876	12
187	CDI9098136_S_12022_09202025_0125.gpr	0.980	50	0.953	21
188	CDI9098138_S_12009_09202025_0126.gpr	0.978	36	0.945	35
189	CDI9098140_S_12035_09202025_0127.gpr	0.983	27	0.923	22
190	CDI9098141_S_12042_09202025_0128.gpr	0.980	24	0.950	14
191	CDI9098143_S_12067_09202025_0129.gpr	0.982	25	0.955	12
192	CDI9098145_S_12097_09202025_0130.gpr	0.979	22	0.962	21
193	CDI9098147_S_12106_09202025_0131.gpr	0.980	21	0.943	7
194	CDI9098149_S_12108_09202025_0132.gpr	0.981	21	0.972	17
195	CDI9098151_S_12011_09202025_0133.gpr	0.980	21	0.930	17
196	CDI9098153_S_12069_09202025_0134.gpr	0.978	66	0.923	14
197	CDI9098155_S_12027_09202025_0135.gpr	0.981	32	0.920	20
198	CDI9098157_S_12030_09202025_0136.gpr	0.979	48	0.889	14
199	CDI9098159_S_12004_09202025_0137.gpr	0.980	41	0.963	13
200	CDI9098161_S_12012_09202025_0138.gpr	0.978	33	0.944	32
201	CDI9098163_S_12060_09202025_0139.gpr	0.982	23	0.910	8
202	CDI9098165_S_12047_09202025_0140.gpr	0.978	34	0.915	11
203	CDI9098166_S_12100_09202025_0141.gpr	0.978	46	0.907	10
204	CDI9098168_S_12112_09202025_0142.gpr	0.983	17	0.899	9
205	CDI9098170_S_12010_09202025_0143.gpr	0.981	21	0.907	13
206	CDI9098172_S_12034_09202025_0144.gpr	0.982	25	0.860	10
207	CDI9098174_S_12092_09202025_0145.gpr	0.983	16	0.900	5
208	CDI9098176_S_12017_09202025_0146.gpr	0.979	35	0.928	6
209	CDI9098178_S_12041_09202025_0147.gpr	0.982	24	0.866	11
210	CDI9098180_S_12048_09202025_0148.gpr	0.982	40	0.931	26
211	CDI9098182_S_12006_09202025_0149.gpr	0.979	38	0.884	6
212	CDI9098183_S_12018_09202025_0154.gpr	0.982	28	0.874	14
213	CDI9098184_S_12007_09202025_0150.gpr	0.980	33	0.934	18
214	CDI9098185_S_12008_09202025_0178.gpr	0.983	25	0.889	13
215	CDI9098186_S_12028_09202025_0151.gpr	0.982	46	0.938	22
216	CDI9098187_S_12064_09202025_0179.gpr	0.982	29	0.890	23
217	CDI9098188_S_12038_09202025_0152.gpr	0.982	43	0.884	17
218	CDI9098190_S_12005_09202025_0153.gpr	0.964	55	0.912	17
219	CDI9098193_S_12102_09202025_0155.gpr	0.958	197	0.843	87
220	CDI9098195_S_12039_09202025_0156.gpr	0.957	217	0.909	93
221	CDI9098199_S_12056_09202025_0158.gpr	0.961	184	0.894	82
222	CDI9098201_S_12083_09202025_0159.gpr	0.961	214	0.881	81
223	CDI9098203_S_12086_09202025_0160.gpr	0.959	247	0.911	79
224	CDI9098205_S_12088_09202025_0161.gpr	0.961	188	0.923	79
225	CDI9098207_S_12103_09202025_0162.gpr	0.963	145	0.879	64
226	CDI9098209_S_12058_09202025_0163.gpr	0.960	174	0.860	75
227	CDI9098211_S_12061_09202025_0164.gpr	0.965	155	0.894	70
228	CDI9098213_S_12104_09202025_0165.gpr	0.968	160	0.941	101
229	CDI9098215_S_12015_09202025_0166.gpr	0.966	114	0.879	63
230	CDI9098216_S_12040_09202025_0167.gpr	0.979	27	0.930	29
231	CDI9098218_S_12068_09202025_0168.gpr	0.983	20	0.933	18
232	CDI9098220_S_12091_09202025_0169.gpr	0.981	20	0.926	6
233	CDI9098222_S_12026_09202025_0170.gpr	0.983	21	0.873	8
234	CDI9098224_S_12029_09202025_0171.gpr	0.985	11	0.967	4
235	CDI9098226_S_12032_09202025_0172.gpr	0.982	31	0.914	8
236	CDI9098228_S_12081_09202025_0173.gpr	0.983	16	0.900	20
237	CDI9098230_S_12001_09202025_0174.gpr	0.985	10	0.880	12
238	CDI9098232_S_12057_09202025_0175.gpr	0.984	13	0.893	5
239	CDI9098234_S_12090_09202025_0176.gpr	0.985	20	0.923	18
240	CDI9098236_S_12003_09202025_0177.gpr	0.979	41	0.964	5

11 Control Spots

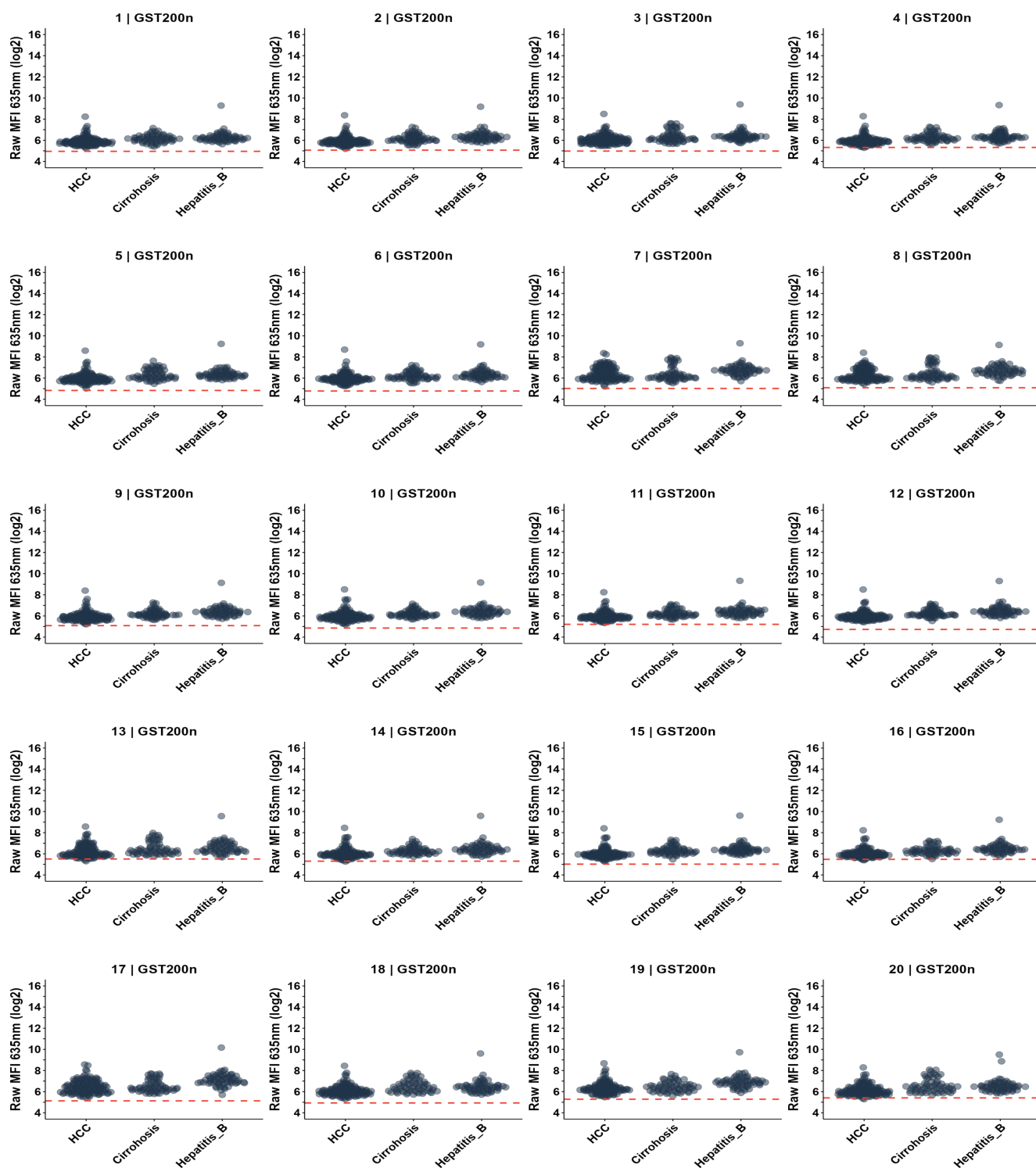
Proteins on HuProt microarrays are arranged into twenty blocks and each block is followed by control spots. Examples of positive and negative control block data across array regions are presented on the following pages. Plots of all control spot types are included in the **control_spots** folder.

Control	Function
H1 - Histone H1 H2 (A+B) - Histone H2A and H2B mixture H3 - Histone H3 H4 - Histone H4	The histones are non-specific binding proteins used as positive controls for certain assays, including DNA, RNA binding assays, etc.
IgG488/594	Alexa Fluor 488/594 labeled IgG, positive control and landmarks for fluorescent detection in 488/594 channels.
Rhodamine + IgG 647	Rhodamine + Alexa Fluor 647 labeled IgG, a positive control and landmarks for fluorescent detection in 532/635 channels.
Anti-human IgA	A positive control for human serum/plasma IgA profiling.
Mouse-anti-biotin	Detects biotin labeled protein probes and serves as a control for anti-mouse antibody detection reagent.
Rabbit-anti-biotin	Detects biotin labeled protein probes and serves as a control for anti-rabbit antibody detection reagent.
BSA - Bovine serum albumin	A negative control for non-specific protein interactions.
Biotin-BSA - biotinylated BSA	A positive control for interaction with streptavidin labeled detection reagent.
Mouse IgM	Positive control for anti-mouse IgM detection.
RanBP2deltaFG	E3 SUMO-Protein Ligase, a positive control for the SUMOylation assay.
hMDM2	E3 Ubiquitin Protein Ligase, a positive control for the ubiquitinylation assay.
Er α	Estrogen receptor alpha. A positive control for ligand binding assays.
Human IgM	A positive control for human serum/plasma IgM profiling.
Human IgG 1.5625 ng/ μ L Human IgG 6.25 ng/ μ L Human IgG 25 ng/ μ L Human IgG 100 ng/ μ L	The human IgG gradient serves as a positive control for human serum/plasma IgG profiling.
Anti-Human IgG 1.5625 ng/ μ L Anti-Human IgG 6.25 ng/ μ L Anti-Human IgG 25 ng/ μ L Anti-Human IgG 100 ng/ μ L	The Anti-human IgG gradient serves as a positive control for human serum/plasma IgG profiling.
GST 10 ng/ μ L GST 50 ng/ μ L GST 100 ng/ μ L GST 200 ng/ μ L	The glutathione S-transferase (GST) protein gradient serves as a negative control for reactivity to the GST-tag.
Buffer	Printing buffer only, negative control.



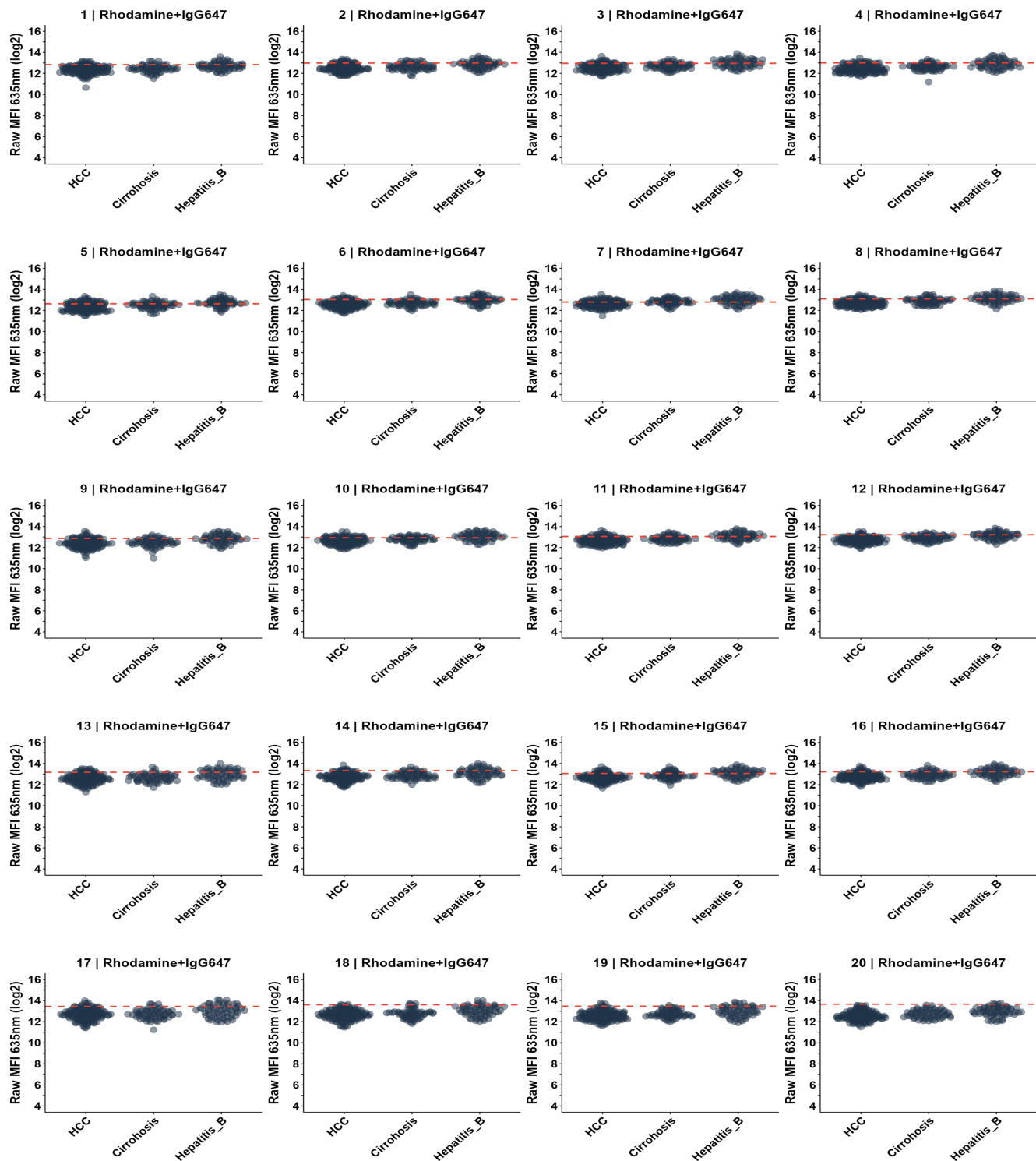
Raw log2 median fluorescence intensity (MFI) of print buffer control spots in 635nm channel broken down by sample group. Print buffer results are a protein-free negative control. Grouped results of individual control-spot results across all twenty control block regions. Dashed orange line represents secondary-only signal.

control_spots/combined_Buffer_635nm_red.png



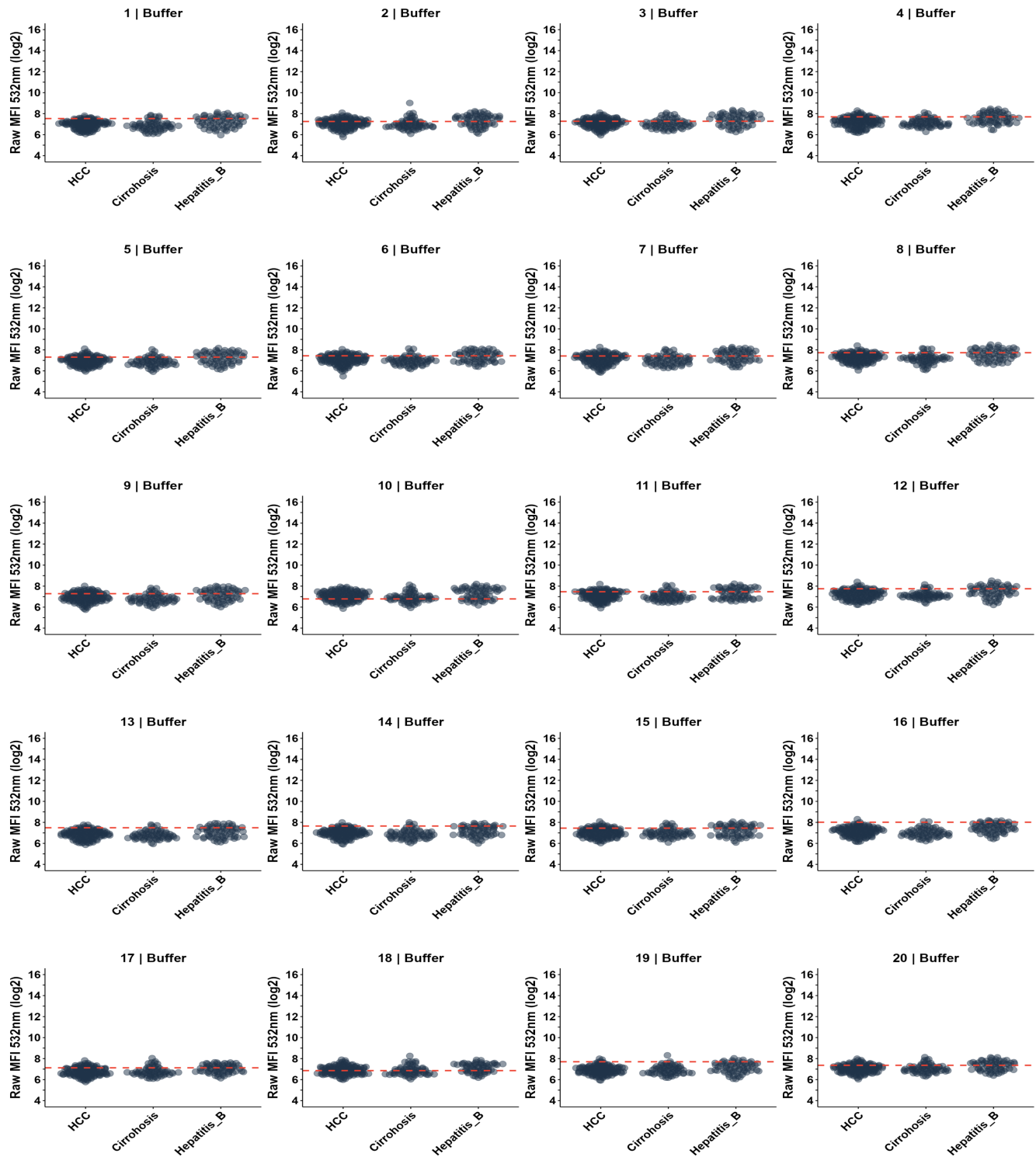
Raw log2 median fluorescence intensity (MFI) of GST (glutathione S-transferase) 200 ng/μL negative control spots in 635nm channel broken down by sample group. Printed GST tags (used to purify all JHU recombinant proteins in the HuProt collection) are a negative control for sample reactivity to GST. Grouped results of individual control-spot results across all twenty control block regions. Dashed orange line represents secondary-only signal.

control_spots/combined_GST200n_635nm_red



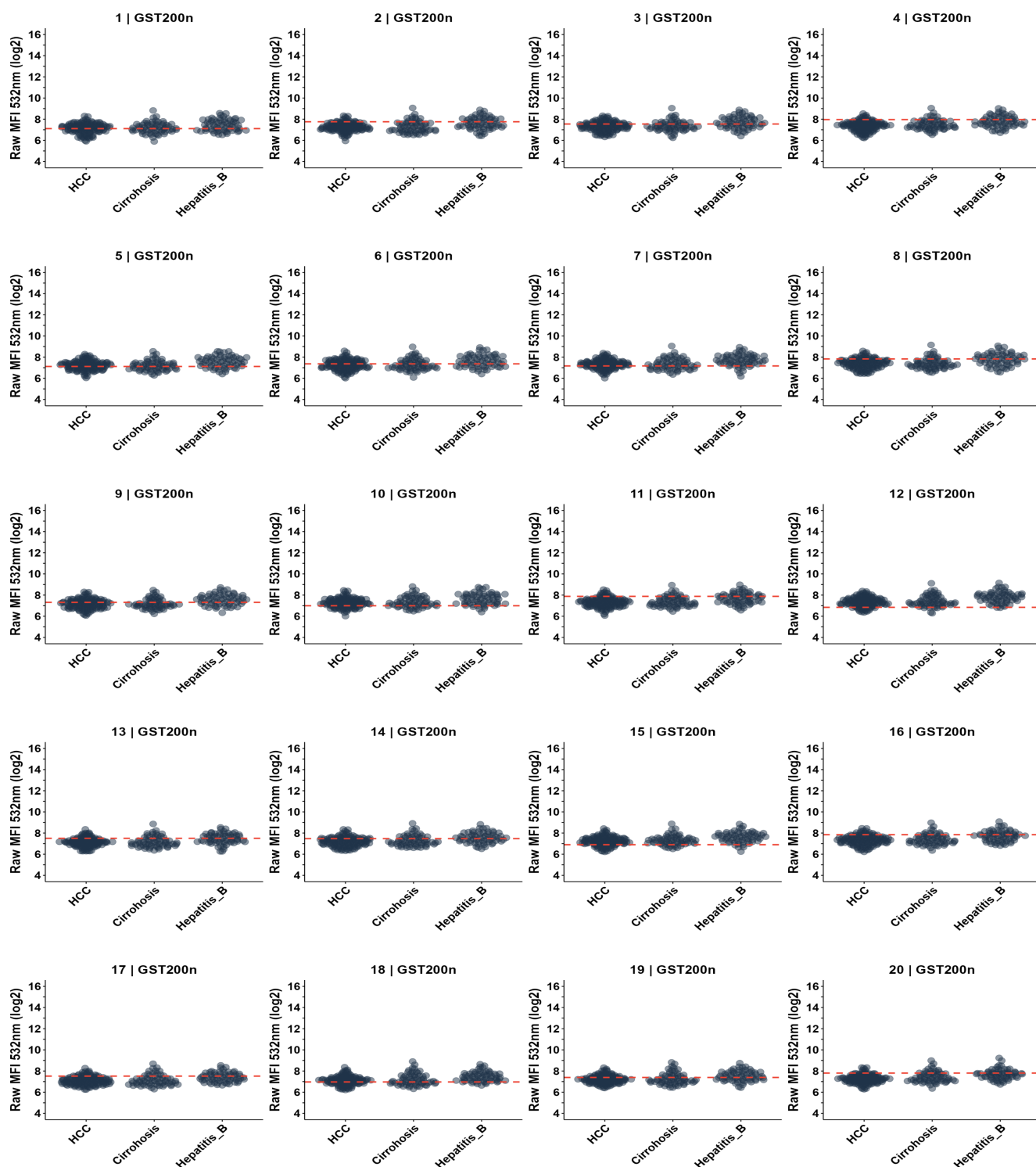
Raw log2 median fluorescence intensity (MFI) of Rhodamine+IgG647 dye control spots in 635nm channel broken down by sample group. Positive control fluorescent dye mimics strongly positive signals. Grouped results of individual control-spot results across all twenty control block regions. Dashed orange line represents secondary-only signal.

control_spots/combined_Rhodamine+IgG647_635nm_red.png



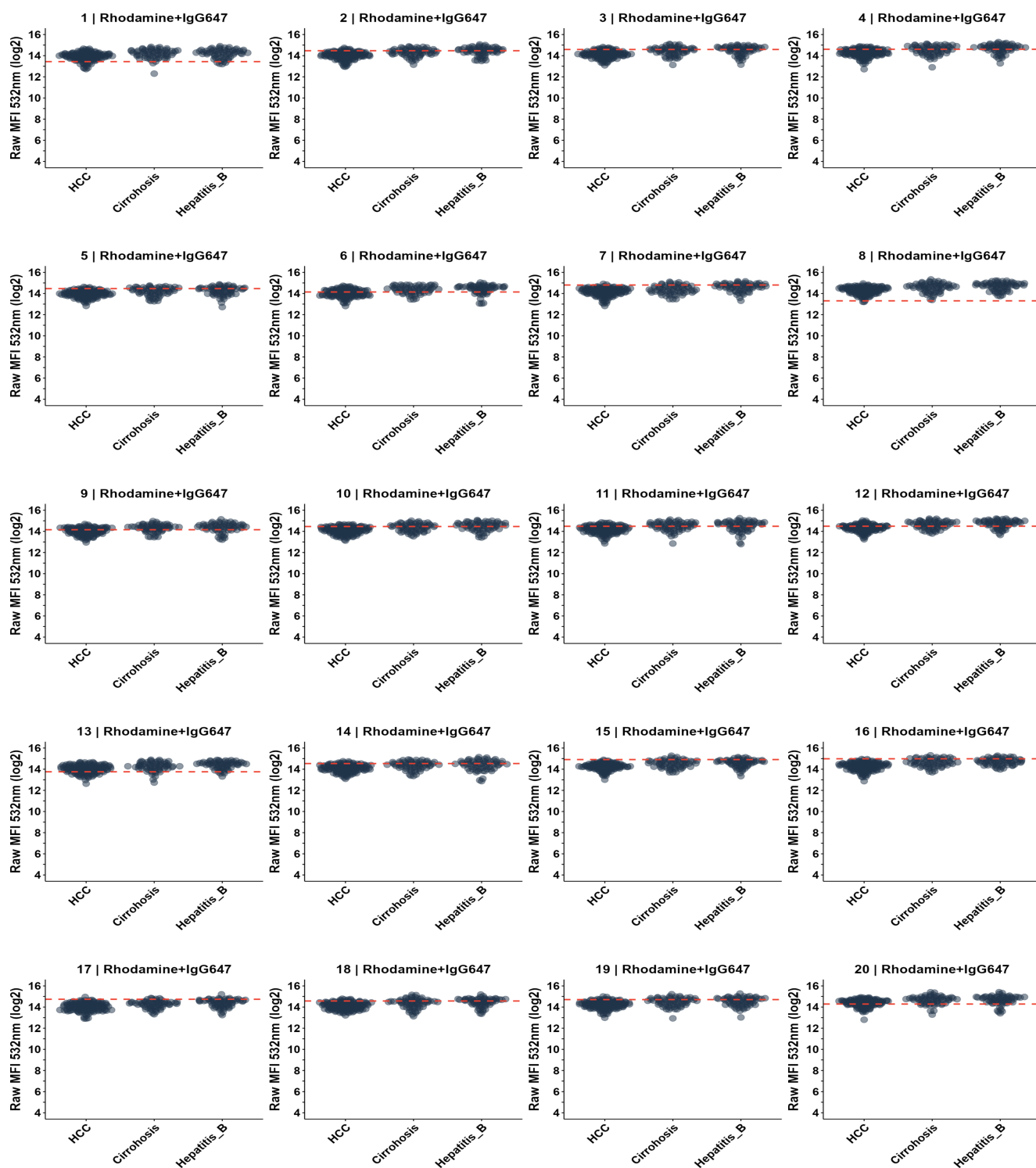
Raw log2 median fluorescence intensity (MFI) of print buffer control spots in 532nm channel broken down by sample group. Print buffer results are a protein-free negative control. Grouped results of individual control-spot results across all twenty control block regions. Dashed orange line represents secondary-only signal.

control_spots/combined_Buffer_532nm_green.png



Raw log2 median fluorescence intensity (MFI) of GST (glutathione S-transferase) 200 ng/ μ L negative control spots in 532nm channel broken down by sample group. Printed GST tags (used to purify all JHU recombinant proteins in the HuProt collection) are a negative control for sample reactivity to GST. Grouped results of individual control-spot results across all twenty control block regions. Dashed orange line represents secondary-only signal.

control_spots/combined_GST200n_532nm_green.png



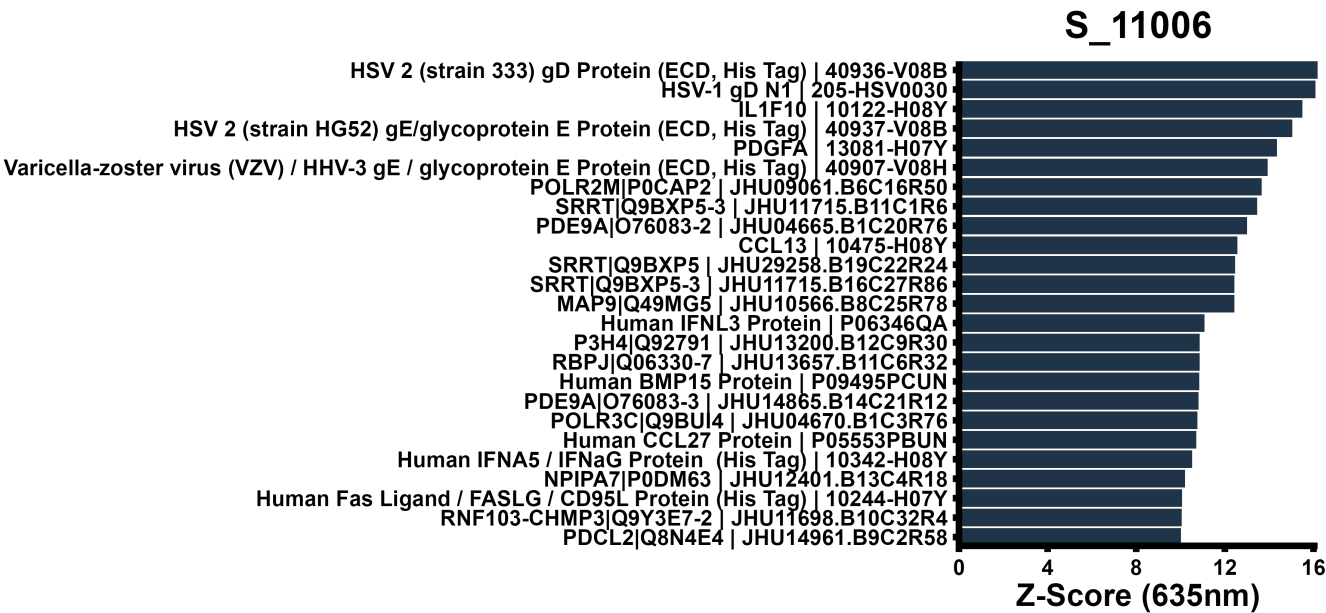
Raw log2 median fluorescence intensity (MFI) of Rhodamine+IgG647 dye control spots in 532nm channel broken down by sample group. Positive control fluorescent dye mimics strongly positive signals. Grouped results of individual control-spot results across all twenty control block regions. Dashed orange line represents secondary-only signal.

control_spots/combined_Rhodamine+IgG647_532nm_green.png

12 Individual Sample Z-Score Barplots

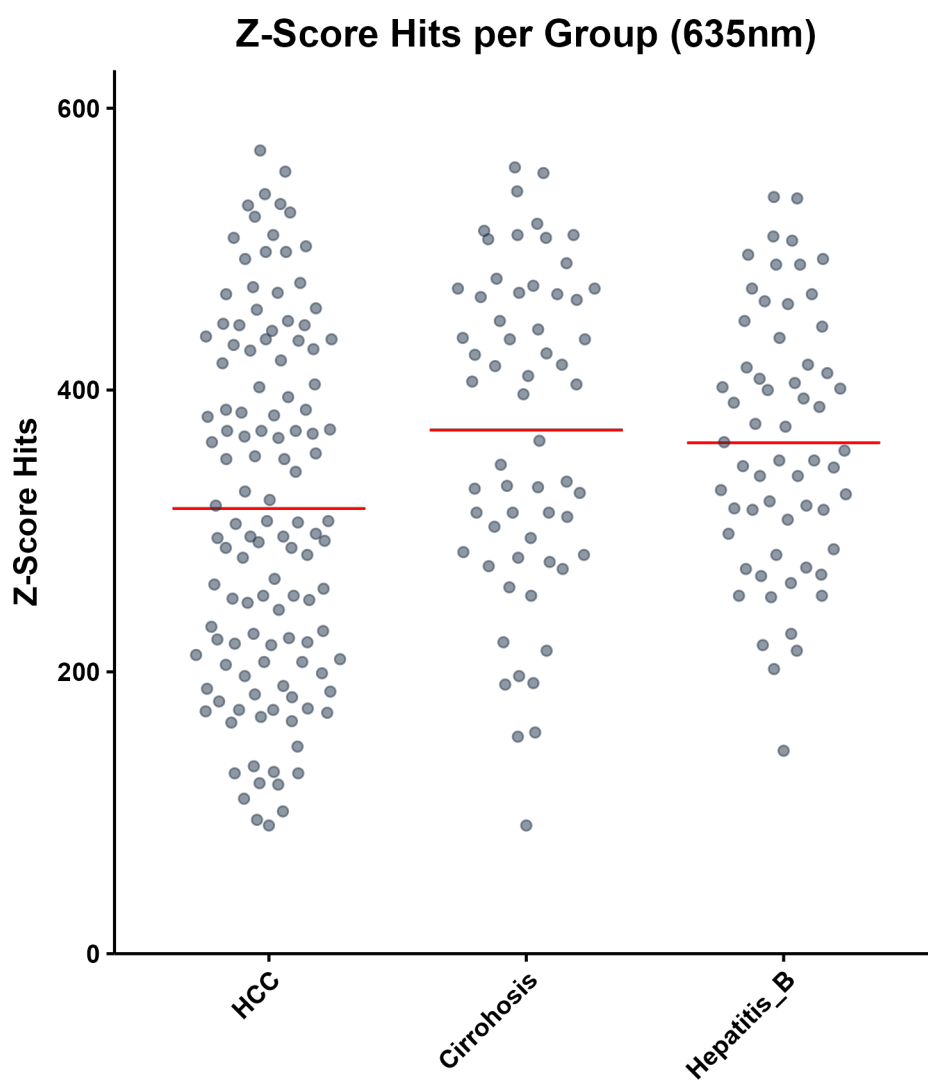
Plots of protein-level top outlier Z-score hits for each sample are provided in the **barplots** folder.

These Z-score barplots are created from the *RawLog2MFI_ZScore_Hits* files and represent the top x25 Z-scores for each individual sample array. If fewer than x25 significant hits are available after caveat and secondary-only filtering, fewer proteins are plotted. An example of one of these images is shown below.



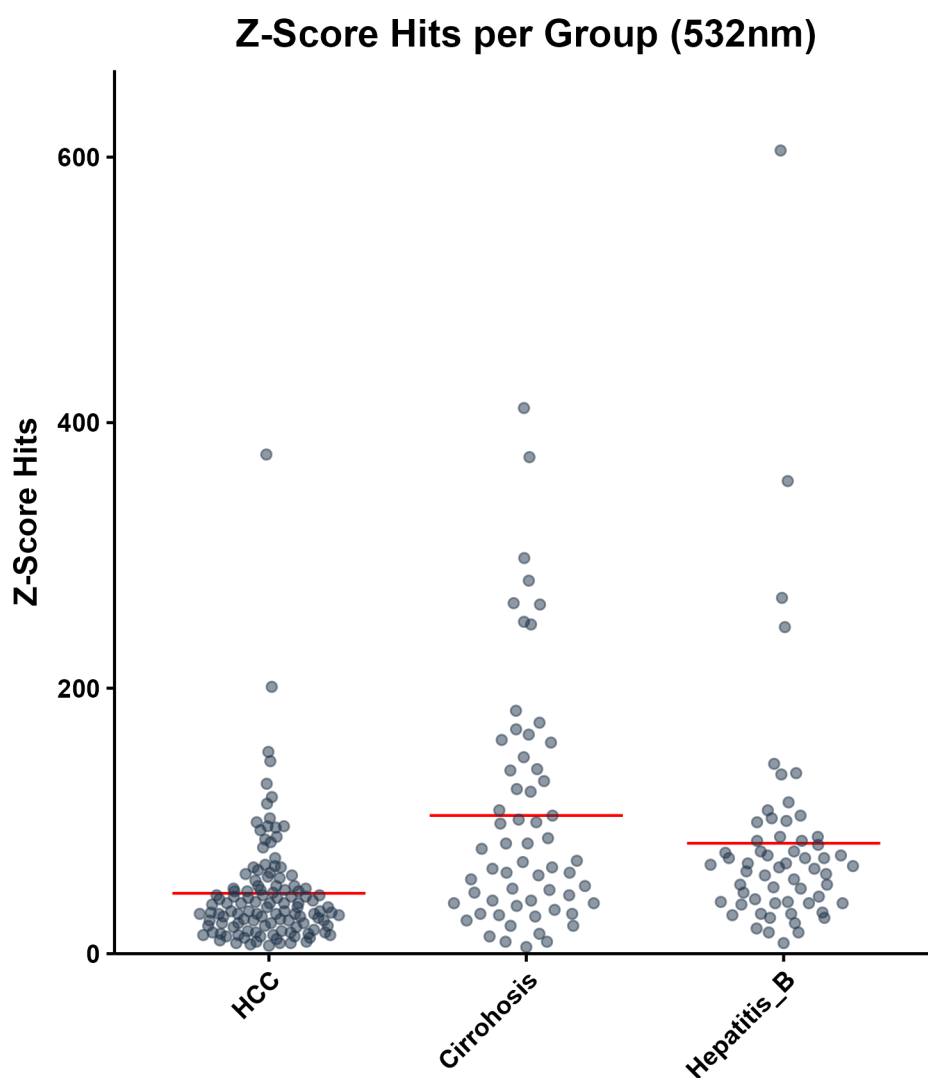
S_11006_635nm_zscore_barplot.png

13 Summary of Z-score Hits



Number of unique protein hits per group in 635nm channel. Grouped results of the total hits called per array. Hits represent raw signals $>4\times$ above spot-specific secondary-only staining and a Z-score >3 within log2 data of each individual array.

plots/hits_per_group_635nm_red.png



Number of unique protein hits per group in 532nm channel. Grouped results of the total hits called per array. Hits represent raw signals >4x above spot-specific secondary-only staining and a Z-score >3 within log2 data of each individual array.

plots/hits_per_group_532nm_green.png

14 Fisher Exact Tests

Binary hit calling followed by case versus control statistics via Fisher exact test have been chosen due to both the wide experimental variance in printed concentration of individual HuProt proteins as well as each reactive sample's variance in concentration and affinity. While HuProt data are continuous from 0 to 65,536 MFI, we typically find that such 'hit' versus 'no hit' comparisons prove more useful for calling candidate outliers.

The following pages contain case versus cohort statistical outputs according to the provided *studyGroup* column entries. If an explicit 'Control' group was provided, all groups are additionally compared to 'Control' only without the rest of the study cohort. For all samples in a group, their status as a *Hit* or *No_Hit* is identified and tallied.

Protein-level Fisher exact tests are performed on a 2x2 matrix of *Hits* and *No_Hits* columns using the *fisher.test* function in R. Resulting p-value calculations are exported and saved in the **fisher_exact_tests** folder to create *fisher_exact_tests* .csv files for each test group and staining channel. To calculate directionality of the p-values, the log odds ratio of the groups is calculated and output in the *log_odds_ratio* column. Positive log odds ratios represent hits over-represented in the *studyGroup* labeled for this set of samples; negative log odds ratios represent hits overrepresented in remainder of the cohort or controls. Note that if hits exist in one group and not the other, an infinite (Inf) or negative infinite (-Inf) ratio results. These values appear appropriately in a text editor or R, but -Inf often displays incorrectly in Excel.

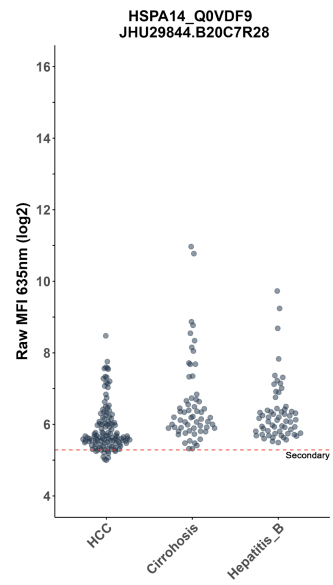
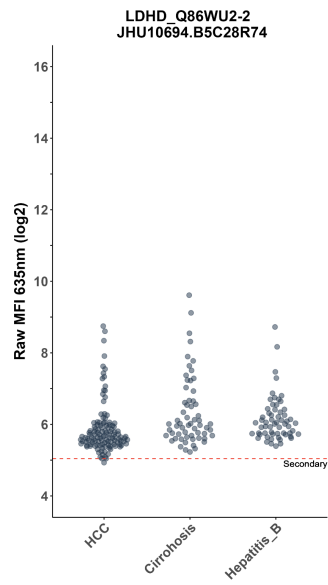
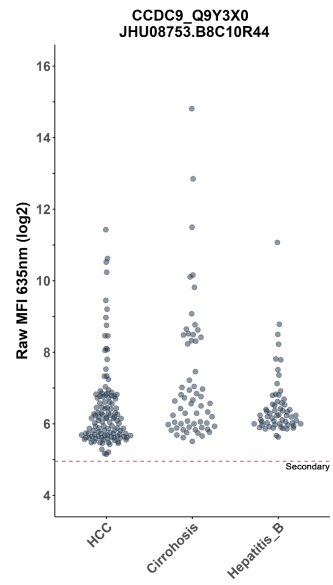
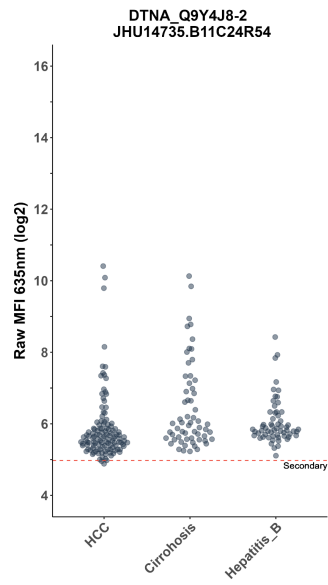
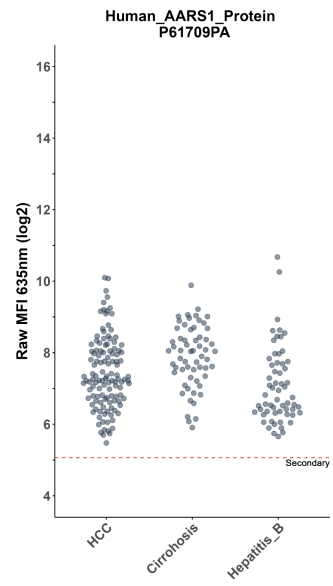
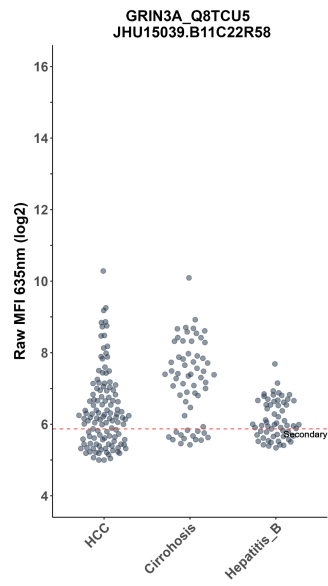
On the following pages you will find tables of the top p-value outlier hits enriched in each provided test group. The top 100 p-values for each test are listed, with full data available in the **fisher_exact_tests** folder.

Protein-level plots of the raw log2 MFI data are displayed for the top 6 p-value outliers for each test; a protein-level plot is additionally created for all the top x100 p-value proteins and provided in the **protein_plots** folder. These plots are useful for quickly visualizing the magnitude of called differences within the underlying raw data.

GC IgG IgA Cirrohosis vs cohort 635nm

Name	ID	Hits Case	Hits Cohort	No Hits Case	No Hits Cohort	p value
GRIN3A Q8TCU5	JHU15039.B11C22R58	16	13	44	167	0.0003210
Human AARS1 Protein	P61709PA	46	93	14	87	0.0008122
DTNA Q9Y4J8-2	JHU14735.B11C24R54	12	9	48	171	0.0009989
CCDC9 Q9Y3X0	JHU08753.B8C10R44	19	21	41	159	0.0010333
LDHD Q86WU2-2	JHU10694.B5C28R74	7	2	53	178	0.0010707
HSPA14 Q0VDF9	JHU29844.B20C7R28	10	6	50	174	0.0011080
EIF1AX P47813	JHU10005.B6C26R70	20	26	40	154	0.0022302
Human RO60 Protein	P00205PB	14	14	46	166	0.0022655
Human GARS1 Protein	P00375PA	41	82	19	98	0.0027494
Human DBT Protein	P65514PA	31	53	29	127	0.0027887
Human RPLP0 Protein	P00135PA	43	89	17	91	0.0027935
PRR22 Q8IZ63	JHU05517.B4C27R86	7	3	53	177	0.0028592
EAF1 Q96JC9	JHU05677.B1C28R88	9	6	51	174	0.0031177
RABEP2 Q9H5N1-2	JHU06787.B8C24R14	9	6	51	174	0.0031177
PPCS Q9HAB8	JHU30058.B18C17R34	17	20	43	160	0.0031583
BCCIP Q9P287	JHU13890.B14C28R10	13	13	47	167	0.0034952
BCCIP Q9P287	JHU04424.B15C32R80	13	13	47	167	0.0034952
RXRBJ P28702	JHU14687.B11C14R54	4	0	56	180	0.0036172
GLMN Q92990	JHU13441.B15C17R86	6	2	54	178	0.0036210
CYP4F2 P78329	JHU30259.B19C30R34	10	8	50	172	0.0038781
PAPSS1 Q43252	JHU29872.B18C21R28	15	17	45	163	0.0038822
WDR62 Q43379	JHU05469.B4C5R86	5	1	55	179	0.0041455
LCN12 Q8IW14	JHU05902.B7C32R4	5	1	55	179	0.0041455
Human CCL17 Protein	P05485QBUN	20	27	40	153	0.0042338
Human FGF8a / FGF-8a Protein	16124-HNAE	21	30	39	150	0.0056433
LDB1 Q86U70-2	JHU02544.B11C26R86	9	7	51	173	0.0057018
ZYX Q15942	JHU08351.B8C25R40	24	38	36	142	0.0060455
Human TRIM21 Protein	P00282PA	52	124	8	56	0.0068445
Human SNRPA Protein	P00187PA	15	18	45	162	0.0080479
EIF1AY Q14602	JHU01949.B4C3R36	11	11	49	169	0.0083165
ZNF232 J3L4J6	JHU09500.B8C4R58	8	6	52	174	0.0083355
TAB2 Q9NYJ8-2	JHU10730.B16C21R10	8	6	52	174	0.0083355
Human IL16 / Interleukin-16 Protein (His Tag)	10944-H07E	8	6	52	174	0.0083355
WASF2 Q9Y6W5	JHU05846.B5C12R4	19	28	41	152	0.0087093
FSIP1 Q8NA03	JHU07906.B5C15R42	6	3	54	177	0.0087215
Asparaginyl-tRNA Synthetase (KS) antigen	A30100	6	3	54	177	0.0087215
Human PPBP Protein	P01722PA	26	45	34	135	0.0090192
PRR27 Q6MZM9	JHU18162.B16C12R44	10	9	50	171	0.0098337
CCDC40 Q4G0X9-5	JHU10191.B5C24R70	18	26	42	154	0.0114290
MGARP Q8TDB4	JHU29573.B18C12R28	15	20	45	160	0.0114880
MNT Q99583	JHU19752.B14C19R2	5	2	55	178	0.0116890
KCNAB1 Q14722-2	JHU05510.B4C29R86	11	12	49	168	0.0117290
NOPCHAP1 Q8N519	JHU08170.B8C8R40	21	33	39	147	0.0117730
RTCA O00442	JHU13663.B9C25R36	7	5	53	175	0.0121180
FBR5 Q9HAH7-1	JHU04346.B2C32R68	13	16	47	164	0.0121540
PDE9A Q76083-10	JHU13549.B11C29R34	13	16	47	164	0.0121540
FAM171A2 Q8N0U1	JHU03899.B10C26R86	8	7	52	173	0.0142990
BIN2 Q9UBW5	JHU29853.B19C10R28	8	7	52	173	0.0142990
SH3BP2 P78314	JHU13094.B11C3R26	12	14	48	166	0.0146280
NADSYN1 Q6IA69	JHU08698.B6C11R46	4	1	56	179	0.0146530
RNPEP Q9H4A4	JHU04008.B16C14R16	4	1	56	179	0.0146530
NDUFA5 Q16718	JHU04570.B2C30R70	3	0	57	180	0.0150400
CELDF2 Q95319-3	JHU05587.B4C6R86	3	0	57	180	0.0150400
LSM3 P62310	JHU07819.B6C27R36	3	0	57	180	0.0150400
BCL7B Q9BQE9-3	JHU11139.B6C5R90	3	0	57	180	0.0150400
C5orf34 Q96MH7	JHU13027.B9C26R28	3	0	57	180	0.0150400
MTURN Q8N3F0	JHU24542.B19C6R6	3	0	57	180	0.0150400
Human IFN-omega/IFNW1	NBP2-35893	3	0	57	180	0.0150400
Human XCL1 Protein	P01494PB	49	116	11	64	0.0154020
Human AMPH Protein	P69786PC	27	50	33	130	0.0166440
MAGEE1 Q9HC15	JHU06955.B5C2R20	14	19	46	161	0.0173860
TBC1D25 B9A6M7	JHU30333.B20C5R38	17	25	43	155	0.0174290
NUDT15 Q9NV35	JHU29945.B20C5R32	12	15	48	165	0.0184250
Human ICA1 Protein	P00477PA	20	33	40	147	0.0195620
Human SNRPD3 Protein	P00442PCS	30	58	30	122	0.0198090
FSIP1 Q8NA03	JHU07906.B6C19R32	9	9	51	171	0.0200080
PDCD6IP Q8WUM4-2	JHU29441.B20C31R20	9	9	51	171	0.0200080
Human SNRPC Protein	P69505PA	24	43	36	137	0.0200610
HMBG2 P26583	JHU04447.B2C24R70	15	21	45	159	0.0201700
Human CCL26 Protein	P05552QBUN	25	45	35	135	0.0210070
CTBP2 P56545-3	JHU05493.B4C32R90	7	6	53	174	0.0210490
FSIP1 Q8NA03	JHU07906.B8C1R42	7	6	53	174	0.0210490
Human EG-VEGF / prokineticin-1 Protein (His Tag)	10183-H08B	7	6	53	174	0.0210490
Human PSIP1 Protein	P00737PC	32	64	28	116	0.0219710
RSRP1 A0A024RAD9	JHU02215.B19C26R16	15	22	45	158	0.0231910
PDE4DIP Q8TBR0	JHU06975.B8C3R24	18	29	42	151	0.0242390
ACOT12 Q8WYK0	JHU30271.B17C28R32	18	29	42	151	0.0242390
Human CCL1 Protein	P05549PCUN	56	146	4	34	0.0248120
Human NGF / NGFB Protein	11050-HNAH	39	86	21	94	0.0251270
LINC01126 NA	JHU08492.B5C15R46	5	3	55	177	0.0251360
SNRPC P09234	JHU10916.B6C27R80	5	3	55	177	0.0251360
LARP6 Q9BRS8	JHU09452.B8C11R56	5	3	55	177	0.0251360
PSMB6 P28072	JHU14684.B11C27R54	5	3	55	177	0.0251360
GPN1 Q9HCN4	JHU14168.B12C30R38	5	3	55	177	0.0251360
IAH1 Q2TAA2	JHU30012.B19C23R34	13	18	47	162	0.0261720
TEPSIN Q96N21	JHU09416.B8C11R60	9	10	51	170	0.0268550
RSPH1 Q8WYR4	JHU16908.B15C4R74	9	10	51	170	0.0268550
VAT1 Q99536	JHU16919.B16C19R78	9	10	51	170	0.0268550
H1-1 Q02539	JHU10506.B8C29R78	19	32	41	148	0.0288020
LMOD3 Q0VAK6	JHU12210.B9C7R12	19	32	41	148	0.0288020
Human I-TAC / CXCL11 Protein	10876-HNAE	19	32	41	148	0.0288020
DCBLD2 Q96PD2	JHU10673.B5C28R78	16	25	44	155	0.0293330
SAMSN1 Q9NSI8	JHU04385.B1C27R68	6	5	54	175	0.0309730
FAM81A Q8TBF8	JHU00804.B3C26R16	6	5	54	175	0.0309730
BAP1 Q92560	JHU08650.B6C6R46	6	5	54	175	0.0309730
PYCR1 P32322	JHU14776.B9C4R50	6	5	54	175	0.0309730
GAPDH5 O14556	JHU03103.B3C2R52	8	8	52	172	0.0313040
OMPI P47874	JHU10138.B8C14R72	8	8	52	172	0.0313040
ACSBG1 Q96GR2	JHU14267.B12C14R46	8	8	52	172	0.0313040
GPR152 Q8TDT2	JHU29771.B18C16R30	8	8	52	172	0.0313040

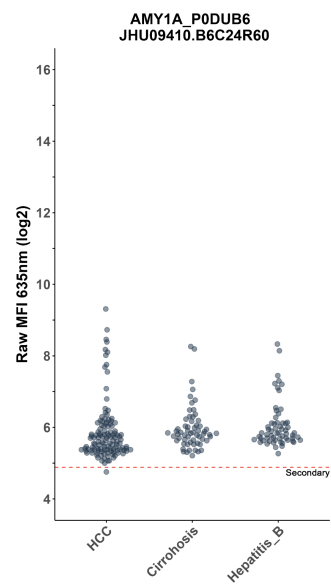
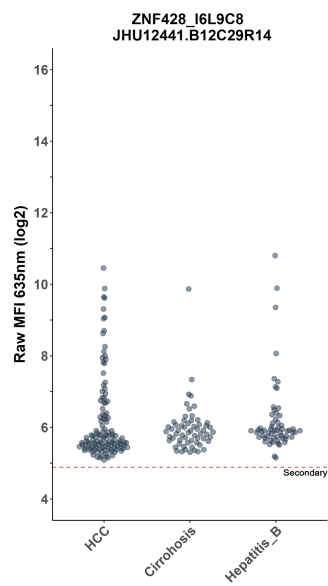
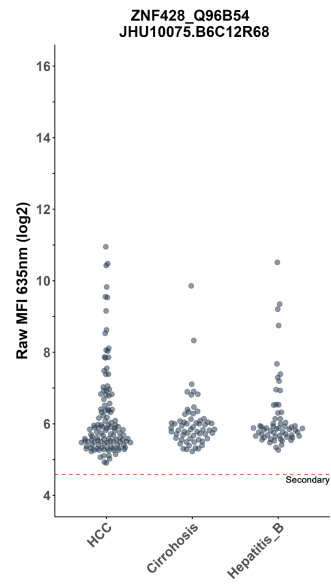
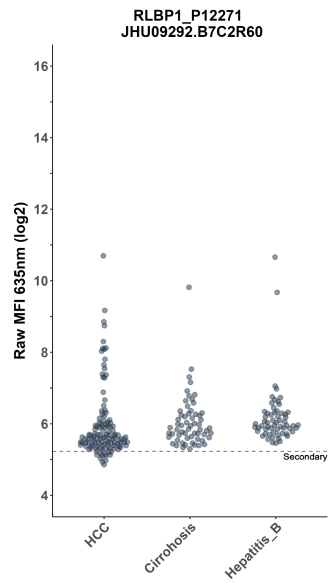
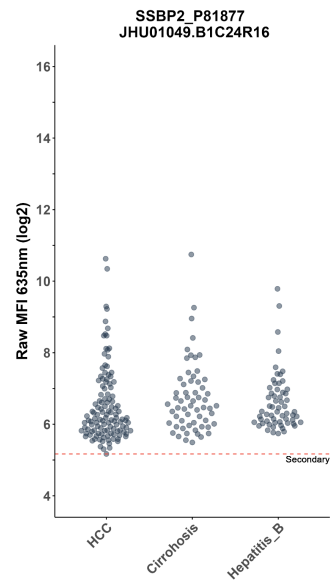
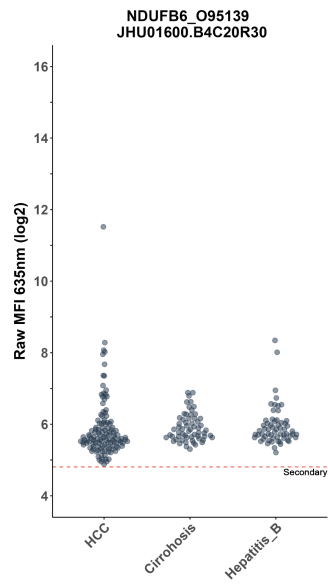
fisher_exact_tests/GC_IgG_IgA_Cirrohosis_vs_cohort_fisher_exact_tests_635nm_red.csv



GC IgG IgA HCC vs cohort 635nm

Name	ID	Hits Case	Hits Cohort	No Hits Case	No Hits Cohort	p value
NDUFB6 O95139	JHU01600.B4C20R30	10	0	110	120	0.0016066
S5BP2 P81877	JHU01049.B1C24R16	19	5	101	115	0.0042736
RLBP1 P12271	JHU09292.B7C2R60	13	2	107	118	0.0057522
ZNF428 Q96B54	JHU10075.B6C12R68	21	7	99	113	0.0079948
ZNF428 I6L9C8	JHU12441.B12C29R14	21	7	99	113	0.0079948
AMY1A P0DUB6	JHU09410.B6C24R60	9	1	111	119	0.0189760
OPTN Q96CV9	JHU14029.B12C19R38	11	2	109	118	0.0190680
TGM1 P22735	JHU08437.B7C10R42	26	12	94	108	0.0206320
USF1 P22415-2	JHU11605.B9C3R6	14	4	106	116	0.0249470
ZNF385A Q96PM9-1	JHU05180.B3C22R82	6	0	114	120	0.0293130
LIN28A Q9H9Z2	JHU03223.B4C14R50	6	0	114	120	0.0293130
ENDOV Q8N8Q3-3	JHU09724.B13C11R68	6	0	114	120	0.0293130
TSPY10 P0CV98	JHU10609.B8C16R74	12	3	108	117	0.0297260
SNCG O76070	JHU03439.B1C28R50	8	1	112	119	0.0355850
NAP1L5 Q96NT1	JHU11855.B9C28R2	16	6	104	114	0.0420140
HERC3 Q15034-2	JHU13058.B12C15R30	16	6	104	114	0.0420140
TMOD1 P28289	JHU13681.B11C16R32	20	9	100	111	0.0461840
SLC52A1 Q9NWF4	JHU09819.B5C25R66	11	3	109	117	0.0503120
BEGAIN Q9BUH8	JHU08593.B8C8R48	17	7	103	113	0.0509410
BOD1 Q96IK1	JHU09627.B6C5R62	21	10	99	110	0.0529260
RFX4 Q33E94-3	JHU04105.B2C6R66	9	2	111	118	0.0594160
BC033035.2 NA	JHU04733.B4C27R78	9	2	111	118	0.0594160
RCVRN P35243	JHU00272.B1C25R6	5	0	115	120	0.0599010
SERPINB2 P05120	JHU04773.B1C17R76	5	0	115	120	0.0599010
CCDC146 Q8NZ26	JHU05043.B2C3R82	5	0	115	120	0.0599010
FAF2 Q96CS3	JHU09496.B5C11R58	5	0	115	120	0.0599010
ZFP36L2 P47974	JHU14705.B10C29R54	5	0	115	120	0.0599010
GPS1 Q13098-7	JHU09442.B7C4R60	18	8	102	112	0.0599220
ERMN Q8TAM6-2	JHU10779.B7C26R84	15	6	105	114	0.0654690
DLX2 Q07687	JHU19565.B15C3R4	28	16	92	104	0.0656770
IFIH1 Q9BYX4-2	JHU06000.B6C25R4	7	1	113	119	0.0657660
BEAN1 Q3B7T3-2	JHU10757.B6C12R84	7	1	113	119	0.0657660
DDX43 Q9NXZ2	JHU09623.B7C15R64	7	1	113	119	0.0657660
IRF2BP1 Q8IU81	JHU12393.B9C27R16	7	1	113	119	0.0657660
EIF4G3 A0A004HJV5	JHU05497.B15C10R78	7	1	113	119	0.0657660
UBE2Q1 Q7Z7E8	JHU25696.B18C24R2	7	1	113	119	0.0657660
RAD23B P54727	JHU04765.B4C17R76	12	4	108	116	0.0671240
PRMT6 Q96LA8	JHU16320.B11C4R74	12	4	108	116	0.0671240
SHD Q96IW2	JHU02186.B4C4R32	30	18	90	102	0.0751680
FUCA2 Q9BTY2	JHU04638.B4C26R74	10	3	110	117	0.0835440
GCLC P48506	JHU13052.B11C12R30	10	3	110	117	0.0835440
KLC4 Q9NSK0	JHU10782.B8C15R80	21	11	99	109	0.0862600
PRKAR1A P10644	JHU05633.B1C4R90	8	2	112	118	0.1020200
PRKCSH A0A024R7F1	JHU09087.B7C6R54	8	2	112	118	0.1020200
KMT5B Q4FZB7-2	JHU13571.B10C9R32	8	2	112	118	0.1020200
DDX53 Q86TM3	JHU16271.B11C24R74	8	2	112	118	0.1020200
LIN28B Q6ZN17	JHU30004.B19C3R34	8	2	112	118	0.1020200
FFAR2 O15552	JHU07902.B6C11R32	11	4	109	116	0.1069200
XPA P23025	JHU09882.B13C1R8	11	4	109	116	0.1069200
TMOD1 P28289	JHU02485.B4C9R40	19	10	101	110	0.1119100
TCOF1 Q13428-5	JHU08626.B5C29R48	19	10	101	110	0.1119100
BAG6 P46379-2	JHU03468.B3C15R56	6	1	114	119	0.1195100
TMEM219 Q86XT9	JHU09361.B6C9R60	6	1	114	119	0.1195100
COTL1 Q14019	JHU08082.B7C5R42	6	1	114	119	0.1195100
WASH6P A8K222	JHU16278.B9C25R78	6	1	114	119	0.1195100
BORCS7 Q96B45	JHU03269.B13C3R14	6	1	114	119	0.1195100
TFPT P0C126	JHU04884.B15C14R18	6	1	114	119	0.1195100
ZC3H12A Q5D1E8	JHU29864.B18C18R28	6	1	114	119	0.1195100
GAGE12F Q76087	JHU05218.B3C4R84	4	0	116	120	0.1218700
RETRG3 Q86VR2	JHU05798.B5C23R2	4	0	116	120	0.1218700
TCP11 Q8WWU5-2	JHU08046.B6C13R32	4	0	116	120	0.1218700
PUM2 Q8TB72-3	JHU11293.B6C29R86	4	0	116	120	0.1218700
MMADHC Q9H3L0	JHU06245.B7C29R12	4	0	116	120	0.1218700
ANKRD13A Q8IZ07	JHU07110.B7C11R24	4	0	116	120	0.1218700
S100G P29377	JHU10343.B7C31R68	4	0	116	120	0.1218700
GALK2 Q01415-2	JHU13731.B9C30R32	4	0	116	120	0.1218700
SCGN O76038	JHU14046.B10C24R42	4	0	116	120	0.1218700
GRIPAP1 Q4V328	JHU14834.B10C30R50	4	0	116	120	0.1218700
DHFR P00374	JHU16402.B10C15R82	4	0	116	120	0.1218700
FAM234B A2RU67	JHU14309.B11C22R44	4	0	116	120	0.1218700
RTN2 O75298-3	JHU14595.B12C26R52	4	0	116	120	0.1218700
GRB7 Q14451	JHU23817.B13C16R88	4	0	116	120	0.1218700
CALR P27797	JHU02029.B14C32R10	4	0	116	120	0.1218700
MAGI1 Q96QZ7-3	JHU18003.B14C25R64	4	0	116	120	0.1218700
GRB7 Q14451	JHU14304.B14C32R86	4	0	116	120	0.1218700
Q6ir13 NA	JHU19587.B17C3R14	4	0	116	120	0.1218700
SLC20A2 Q08357	JHU22033.B18C1R14	4	0	116	120	0.1218700
ZRANB1 Q9UGI0	JHU24528.B19C17R4	4	0	116	120	0.1218700
Human XRCC6 Protein	P70277PA	4	0	116	120	0.1218700
Human IL33 / Interleukin-33 / NF-HEV Protein	10368-HNAE	4	0	116	120	0.1218700
Varicella-zoster virus (VZV) / HHV-3 gE / glycoprotein E Protein (ECD, His Tag)	40907-V08H	120	116	0	4	0.1218700
TEX44 Q53QW1	JHU04272.B3C15R70	20	11	100	109	0.1225900
EIF4EBP2 Q13542	JHU04543.B4C28R70	12	5	108	115	0.1291000
PAIP2 Q49AE6	JHU15547.B9C16R62	12	5	108	115	0.1291000
MARCKS P29966	JHU18576.B12C16R60	12	5	108	115	0.1291000
LAMTOR5 A0A0C4DGV4	JHU09922.B14C29R82	12	5	108	115	0.1291000
IL13 P35225	JHU14472.B15C25R84	12	5	108	115	0.1291000
AMIGO3 Q86WK7	JHU10561.B8C4R78	16	8	104	112	0.1306800
ARL2BP Q9Y2Y0-2	JHU15098.B11C24R60	27	17	93	103	0.1326400
XRCC4 Q13426	JHU03452.B16C4R10	21	12	99	108	0.1328400
EP400P1 Q6ZTU2-2	JHU10388.B6C15R74	9	3	111	117	0.1358200
KANK4 Q5T7N3	JHU12158.B9C18R8	9	3	111	117	0.1358200
KLHL40 A8K5H9	JHU19744.B17C4R12	9	3	111	117	0.1358200
WASHC2C Q9Y4E1-6	JHU04256.B1C5R68	10	4	110	116	0.1666200
Human VEGF-D / VEGFD / FIGF Protein (His Tag)	10557-H08H	10	4	110	116	0.1666200
ITGB1BP2 Q9UKP3	JHU10218.B5C19R72	14	7	106	113	0.1693100
NOSTRIN Q8IV19	JHU15153.B10C25R60	14	7	106	113	0.1693100
SIGMAR1 Q99720	JHU03610.B1C27R60	7	2	113	118	0.1714000
RG519 P49795	JHU06319.B5C22R12	7	2	113	118	0.1714000
COL14A1 Q05707-3	JHU08658.B6C16R44	7	2	113	118	0.1714000

fisher_exact_tests/GC_IgG_IgA_HCC_vs_cohort_fisher_exact_tests_635nm_red.csv

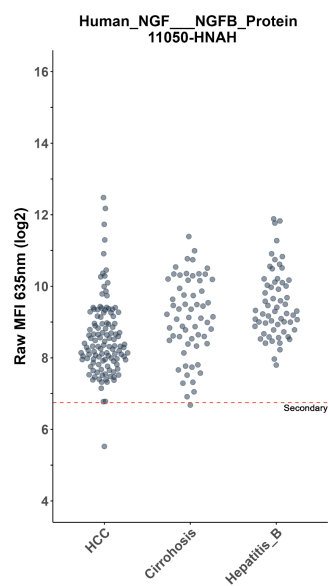
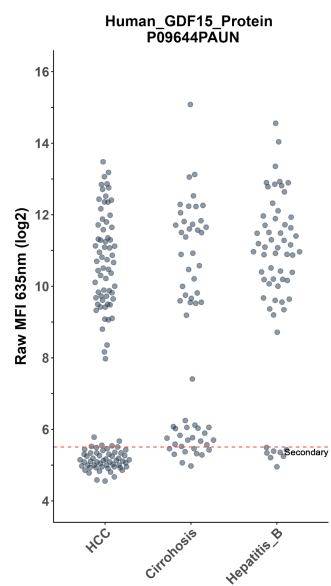
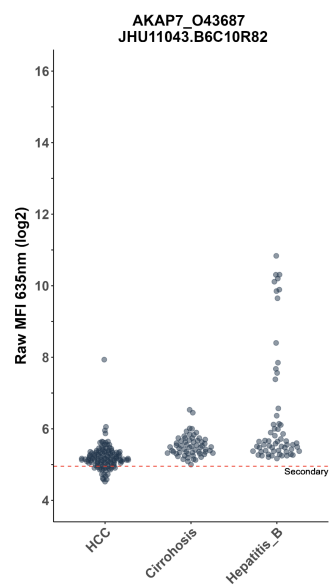
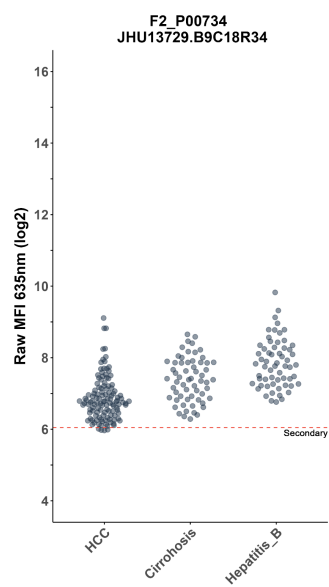
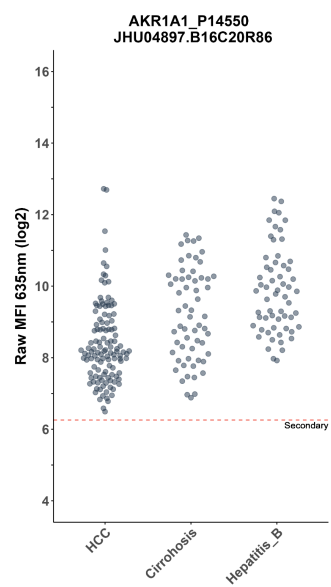


GC IgG IgA Hepatitis B vs cohort 635nm

Name	ID	Hits Case	Hits Cohort	No Hits Case	No Hits Cohort	p value
AKR1A1 P14550	JHU04897.B16C20R86	55	95	5	85	1.1986e-08
F2 P00734	JHU13729.B9C18R34	21	10	39	170	6.9072e-08
AKAP7 Q43687	JHU11043.B6C10R82	11	1	49	179	1.0942e-06
Histone (non recombinant; bovine) antigen	A31100	33	38	27	142	2.6207e-06
Human GDF15 Protein	P09644PAUN	52	97	8	83	2.8858e-06
Human NGF / NGFB Protein	11050-HNAH	46	79	14	101	1.1471e-05
Human IL12A / NKSF1 Protein (His Tag)	10021-H08H	53	110	7	70	5.3428e-05
BLVRB P30043	JHU10852.B7C8R82	36	56	24	124	0.00010752
Human PBEF / Visfatin / NAMPT Protein (His & GST Tag)	10990-H20B	57	133	3	47	0.00020459
SIRT5 Q9NXA8	JHU04487.B1C25R68	15	12	45	168	0.0002952
AKR7A2 Q43488	JHU09221.B7C26R56	53	116	7	64	0.00029716
SPDL1 Q96EA4	JHU08273.B13C32R84	24	31	36	149	0.00061263
MYPN Q86TC9	JHU25487.B20C24R6	5	0	55	180	0.00085832
LPCAT4 Q643R3	JHU16483.B9C29R84	14	12	46	168	0.0011237
RAP1GAP P47736	JHU24931.B14C31R86	6	1	54	179	0.001128
ECHDC1 Q9NNTX5-2	JHU25034.B14C21R90	17	18	43	162	0.00117
KAT5 Q92993	JHU09730.B7C17R66	19	23	41	157	0.0015466
BTN2A2 Q8WVV5-4	JHU11049.B5C10R82	9	5	51	175	0.0015584
RPUSD2 Q8I273	JHU10054.B5C5R68	16	17	44	163	0.0018799
Human FAM3B Protein	P02252PA	40	78	20	102	0.0026853
ZNF358 Q9NW07	JHU05563.B3C8R90	7	3	53	177	0.0028592
PRRG2 Q14669	JHU11101.B7C30R84	7	3	53	177	0.0028592
NME4 Q00746	JHU15632.B10C13R62	7	3	53	177	0.0028592
METTL21A Q8WXB1	JHU03779.B4C31R56	4	0	56	180	0.0036172
DMRTB1 J6L9A0	JHU18841.B13C20R68	4	0	56	180	0.0036172
ZBTB78 Q15156-2	JHU25320.B13C31R90	12	11	48	169	0.0038665
DNAJB8 Q8NH50	JHU09045.B7C11R52	15	17	45	163	0.0038822
SPG7 Q9UQ90	JHU13673.B10C6R34	25	40	35	140	0.0044211
DPP3 G3V180	JHU08467.B5C5R48	8	5	52	175	0.0044629
NAV1 Q8NEY1-3	JHU25871.B17C31R6	22	33	38	147	0.0046685
Human TNFSF13 Protein	P02055PCSN	48	107	12	73	0.0047571
HGS Q14964	JHU08677.B6C1R46	40	82	20	98	0.0048224
Lupus La	Auto-antigen.B20C4R38	40	82	20	98	0.0048224
SOX5 P35711-2	JHU13672.B11C6R34	32	58	28	122	0.0053023
CPLX2 Q6PUV4	JHU11920.B9C23R10	33	62	27	118	0.0060343
TH P07101-3	JHU30247.B17C10R32	11	10	49	170	0.0061473
KIAA1191 Q96A73	JHU07922.B7C6R36	7	4	53	176	0.0063001
R3HDM2 Q9Y2K5	JHU15555.B10C4R62	7	4	53	176	0.0063001
STYK1 Q6J9G0	JHU07184.B13C25R72	24	39	36	141	0.0067959
Human CCL1 Protein	P05549PCUN	57	145	3	35	0.0072448
OR13J1 Q8NGT2	JHU30061.B20C6R32	22	34	38	146	0.0077011
WASHC3 Q9Y3C0	JHU04431.B3C25R70	11	11	49	169	0.0083165
GBP6 Q6ZN66	JHU10214.B13C23R90	17	23	43	157	0.0085525
WARS2 Q9UGM6-2	JHU11316.B16C22R78	22	35	38	145	0.0087089
DYDC1 Q8WWB3	JHU01758.B2C24R28	13	15	47	165	0.0094153
INTU Q9ULD6-2	JHU05509.B2C29R86	5	2	55	178	0.011689
TRIM24 Q15164	JHU11599.B9C23R4	5	2	55	178	0.011689
AJ315540.1 NA	JHU12499.B10C4R14	5	2	55	178	0.011689
OCM P0CE72	JHU17076.B16C5R36	10	10	50	170	0.01287
MMP8 P22894	JHU11474.B6C10R88	12	14	48	166	0.014628
ANAPC15 P60006	JHU03470.B1C15R56	4	1	56	179	0.014653
ZNF207 Q43670-3	JHU02206.B2C4R34	4	1	56	179	0.014653
MAGEA8 P43361	JHU02742.B3C20R46	4	1	56	179	0.014653
ERIC2 A1L162	JHU10895.B8C22R82	3	0	57	180	0.01504
DGKZ Q13574-4	JHU14637.B9C29R52	3	0	57	180	0.01504
MTMR3 Q13615	JHU12975.B10C30R20	3	0	57	180	0.01504
RPA3 P35244	JHU13284.B11C11R30	3	0	57	180	0.01504
ARHGEF1 Q92888-2	JHU13985.B11C6R40	3	0	57	180	0.01504
EIF4A2 Q14240-2	JHU05113.B3C8R84	9	8	51	172	0.01585
ETFB P38117	JHU11163.B7C2R90	9	8	51	172	0.01585
Human IL15 Protein	P01004PB	37	78	23	102	0.016884
MLF2 Q15773	JHU02644.B1C7R42	6	4	54	176	0.017515
ZNF207 Q43670-2	JHU12824.B12C22R24	6	4	54	176	0.017515
SNX4 Q95219	JHU01424.B16C10R82	10	11	50	169	0.017736
NUDT15 Q9NV35	JHU29945.B20C5R32	12	15	48	165	0.018425
HSP90AA1 P07900	JHU29409.B17C19R20	24	42	36	138	0.018745
SNX29 Q8TEQ0	JHU18130.B13C25R78	29	56	31	124	0.019419
ACO1 P21399	JHU03940.B15C15R84	9	9	51	171	0.020008
C8orf58 Q8NAV2	JHU13609.B11C6R36	11	13	49	167	0.022702
ATP5F1B P06576	JHU07875.B8C1R36	22	38	38	142	0.024371
NME2 P22392	JHU00058.B1C5R2	42	96	18	84	0.024658
ARHGEF16 Q5VV41	JHU02405.B3C9R40	35	74	25	106	0.024707
PCP2 Q8IVA1-2	JHU05244.B1C16R84	5	3	55	177	0.025136
ADH5 P11766	JHU04710.B3C17R74	5	3	55	177	0.025136
PIK3R1 P27986-3	JHU11097.B6C13R82	5	3	55	177	0.025136
DLG3 Q92796	JHU15594.B11C31R66	5	3	55	177	0.025136
LCNL1 B9EHA8	JHU14742.B19C6R18	5	3	55	177	0.025136
Human BCA-1 / CXCL13 Protein	10621-HNAE	5	3	55	177	0.025136
CDKN1B P46527	JHU08459.B7C2R46	9	10	51	170	0.026855
FGFR2 P21802-3	JHU12013.B13C24R14	11	14	49	166	0.02787
HNRNPCL1 Q06012	JHU30033.B19C15R36	29	58	31	122	0.030007
MARVELD2 Q8N459	JHU05049.B2C14R68	6	5	54	175	0.030973
KCNK10 P57789-4	JHU10221.B15C26R84	6	5	54	175	0.030973
EEF1DP3 Q658K8	JHU03313.B4C6R52	8	8	52	172	0.031304
DPP3 Q9NY33	JHU07031.B6C32R22	8	8	52	172	0.031304
UCN3 Q969E3	JHU15086.B11C32R56	8	8	52	172	0.031304
VCL P18206-2	JHU18520.B14C30R46	8	8	52	172	0.031304
C18orf25 Q96B23-2	JHU11431.B14C9R84	8	8	52	172	0.031304
PTP4A2 Q12974	JHU07268.B13C6R68	12	16	48	164	0.034042
Human C1QA Protein	P00094PA	32	67	28	113	0.034085
CDC123 Q13519	JHU02321.B2C5R38	4	2	56	178	0.035667
PNOG Q13519	JHU03912.B2C23R62	4	2	56	178	0.035667
MXD4 NA	JHU03903.B2C15R64	4	2	56	178	0.035667
HK1 P19367	JHU08679.B6C16R46	4	2	56	178	0.035667
RHOJ Q9H4E5	JHU11489.B6C7R86	4	2	56	178	0.035667
MED20 Q9H944	JHU06039.B7C24R4	4	2	56	178	0.035667
HOMER1 Q5U5K4	JHU16593.B11C16R80	4	2	56	178	0.035667
NUTM2G Q5VZR2-2	JHU19285.B16C6R78	4	2	56	178	0.035667
APBA2 Q99767-2	JHU30292.B19C31R34	4	2	56	178	0.035667
Human CNTF Protein (His Tag)	11841-H07E	4	2	56	178	0.035667

fisher_exact_tests/GC_IgG_IgA_Hepatitis_B_vs_cohort_fisher_exact_tests_635nm_red.csv

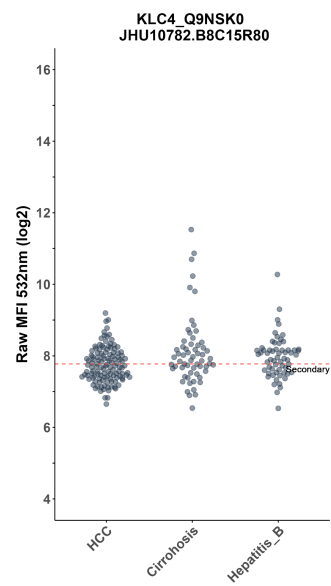
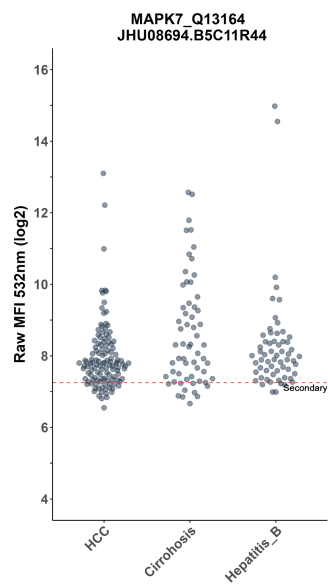
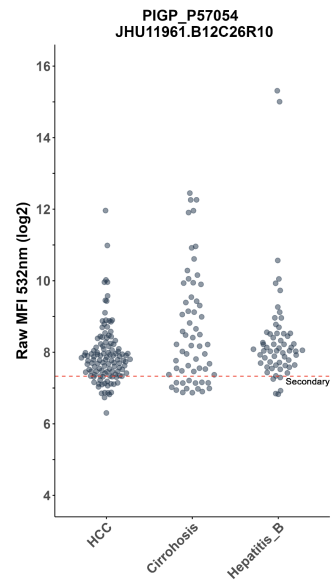
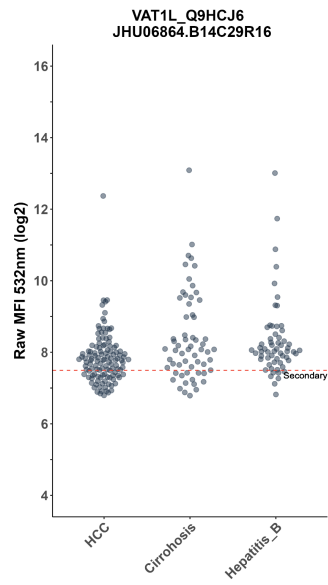
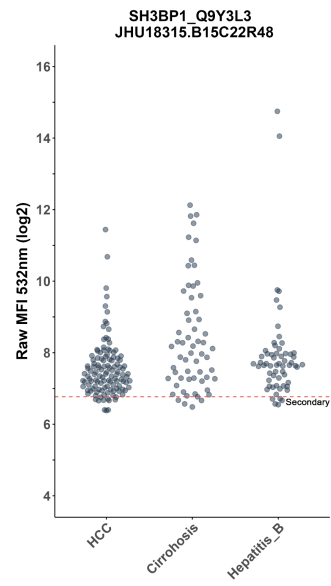
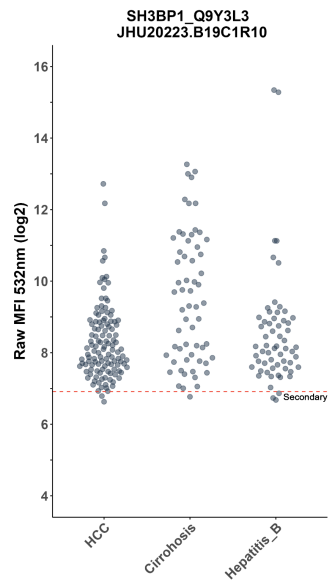
Warning in FUN(X[[i]], ...): File not found: protein_plots/Histone_non_recombinant_bovine_



GC IgG IgA Cirrohosis vs cohort 532nm

Name	ID	Hits Case	Hits Cohort	No Hits Case	No Hits Cohort	p value
SH3BP1 Q9Y3L3	JHU20223.B19C1R10	35	40	25	140	6.8326e-07
SH3BP1 Q9Y3L3	JHU18315.B15C22R48	19	13	41	167	8.0935e-06
VAT1L Q9HJC6	JHU06864.B14C29R16	13	6	47	174	3.8125e-05
PIGP P57054	JHU11961.B12C26R10	17	13	43	167	6.984e-05
MAPK7 Q13164	JHU08694.B5C11R44	18	15	42	165	0.00011352
KLC4 Q9NSK0	JHU10782.B8C15R80	6	0	54	180	0.00020088
MYL1 P05976	JHU11955.B12C23R10	14	10	46	170	0.00024454
RNP-Sm FREE	ATR04-02	5	0	55	180	0.00085832
MYL4 P12829	JHU00633.B3C19R8	12	9	48	171	0.00099891
ACSBG1 Q96GR2	JHU14267.B12C14R46	7	2	53	178	0.0010707
STAT4 Q14765	JHU14690.B12C9R54	10	6	50	174	0.001108
TIGD1 Q96MW7	JHU09780.B5C22R64	6	1	54	179	0.001128
SPP1 P10451-3	JHU16134.B11C3R78	6	1	54	179	0.001128
TTC32 Q510X7	JHU10819.B7C20R80	25	37	35	143	0.0020016
SYTL2 Q9HCH5-2	JHU00086.B14C13R16	8	4	52	176	0.0021444
CRTAC1 Q9NQ79	JHU03192.B4C32R50	10	7	50	173	0.0021539
TGM1 P22735	JHU08437.B7C10R42	7	3	53	177	0.0028592
CCDC181 B2R917	JHU14181.B9C5R44	7	3	53	177	0.0028592
RBM12 Q9NTZ6	JHU08611.B8C17R48	9	6	51	174	0.0031177
RUVBL1 Q9Y265	JHU02091.B14C23R14	9	6	51	174	0.0031177
Human MMP3 Protein	P02109PA	9	6	51	174	0.0031177
JO-1	ATJ01-02	4	0	56	180	0.0036172
PTP4A1 Q93096	JHU02949.B1C24R46	6	2	54	178	0.003621
PTGES3L-AARSD1 Q9BTE6-2	JHU08928.B8C27R54	10	8	50	172	0.0038781
SPAG7 Q75391	JHU05073.B1C23R80	5	1	55	179	0.0041455
TRIM33 Q9UPN9	JHU19543.B8C11R66	5	1	55	179	0.0041455
APIP Q96GX9	JHU04801.B15C4R8	5	1	55	179	0.0041455
WASF2 Q9Y6W5	JHU05846.B5C12R4	8	5	52	175	0.0044629
STARDB8 Q92502	JHU08042.B5C30R36	8	5	52	175	0.0044629
PTPN12 Q05209	JHU10628.B8C19R74	8	5	52	175	0.0044629
HSPA13 P48723	JHU13676.B10C19R32	8	5	52	175	0.0044629
PGK1 P00558	JHU16547.B11C29R82	8	5	52	175	0.0044629
CFDP1 Q9UEE9	JHU12568.B18C25R16	8	5	52	175	0.0044629
CKM P06732	JHU00018.B16C12R14	9	7	51	173	0.0057018
CAVIN1 Q6NZI2	JHU09958.B16C32R84	9	7	51	173	0.0057018
ARMH4 Q86TY3	JHU15103.B12C1R56	11	10	49	170	0.0061473
LMOD3 Q0VAK6	JHU12210.B9C7R12	15	18	45	162	0.0080479
RALGDS Q12967-6	JHU08031.B6C13R36	26	44	34	136	0.0082663
TIGD1 Q96MW7	JHU09976.B8C31R62	8	6	52	174	0.0083355
ACSBG1 Q96GR2	JHU08259.B5C24R42	6	3	54	177	0.0087215
TYRO3 Q06418	JHU08055.B6C30R36	6	3	54	177	0.0087215
PLEKHO2 Q8TD55-2	JHU09471.B6C25R58	6	3	54	177	0.0087215
TEX33 Q043247-2	JHU07207.B9C32R90	6	3	54	177	0.0087215
EPN1 Q9Y6I3-3	JHU12367.B12C8R18	6	3	54	177	0.0087215
BTF3L4 Q96K17	JHU01739.B15C20R16	6	3	54	177	0.0087215
Human SNRPD3 Protein	P00442PCSN	6	3	54	177	0.0087215
Human CSF3 Protein	P00197PDSN	6	3	54	177	0.0087215
Human CCL15 Protein	P09941PCUN	15	19	45	161	0.0093359
CASTOR2 A6NHX0	JHU10701.B15C20R70	10	9	50	171	0.0098337
CCL13	10475-H08Y	43	94	17	86	0.010247
SLC4A1AP Q9BWU0	JHU10547.B5C8R74	5	2	55	178	0.011689
ESD P10768	JHU12753.B9C14R22	5	2	55	178	0.011689
RBBP5 Q15291-2	JHU13371.B10C18R26	5	2	55	178	0.011689
Human EXOSC9 Protein	P65611PA	5	2	55	178	0.011689
Human CCL17 Protein	P05485QBUN	5	2	55	178	0.011689
Human CCL5 Protein	P05487QBSN	5	2	55	178	0.011689
Human AMPH Protein	P69786PC	11	12	49	168	0.011729
BCL2L1 Q043521-5	JHU10570.B5C11R78	7	5	53	175	0.012118
SORBS1 Q9BX66-7	JHU07575.B6C24R28	7	5	53	175	0.012118
ERICH2 A1L162	JHU10895.B8C22R82	7	5	53	175	0.012118
SPA17 Q15506	JHU13004.B11C7R24	7	5	53	175	0.012118
Human CCL26 Protein	P05552QBUN	7	5	53	175	0.012118
STMN4 Q9H169-2	JHU13294.B11C25R26	13	16	47	164	0.012154
YAP1 P46937	JHU13111.B10C2R30	12	13	48	167	0.012354
Human CCL16 Protein	P08898PBUN	14	17	46	163	0.012742
NME2 P22392	JHU00058.B1C5R2	10	10	50	170	0.01287
MT3 P25713	JHU04849.B2C16R74	10	10	50	170	0.01287
Human TRIM21 Protein	P00282PA	30	57	30	123	0.013102
ACAA1 P09110	JHU08544.B7C8R48	18	27	42	153	0.013177
ACY1 Q03154	JHU04708.B1C14R76	8	7	52	173	0.014299
STARDB5 Q9NSY2	JHU09680.B13C2R86	12	14	48	166	0.014628
NINJ1 Q92982	JHU02068.B3C8R36	4	1	56	179	0.014653
EEF1D P29692	JHU04251.B3C5R68	4	1	56	179	0.014653
BDH2 Q9BUT1	JHU02026.B4C29R36	4	1	56	179	0.014653
VPS35L Q7Z3J2	JHU07638.B5C26R26	4	1	56	179	0.014653
PLEKHF2 Q9H8W4	JHU09851.B5C1R66	4	1	56	179	0.014653
KRT36 Q76013-2	JHU05796.B6C5R2	4	1	56	179	0.014653
UBALD1 Q8TB05	JHU08290.B7C6R40	4	1	56	179	0.014653
MALL Q13021	JHU09363.B7C9R58	4	1	56	179	0.014653
MRPL1 Q9BYD6	JHU08312.B8C16R42	4	1	56	179	0.014653
IRF3 Q14653	JHU11362.B8C25R88	4	1	56	179	0.014653
CGRRF1 Q99675	JHU13896.B10C22R38	4	1	56	179	0.014653
ALDH9A1 P49189-3	JHU16249.B11C27R36	4	1	56	179	0.014653
RND1 Q92730	JHU14334.B11C29R48	4	1	56	179	0.014653
MAL P21145	JHU12021.B12C8R8	4	1	56	179	0.014653
GSTT2B POCG30	JHU00424.B14C25R80	4	1	56	179	0.014653
TACC1 Q075410-7	JHU17245.B15C11R74	4	1	56	179	0.014653
MDA5 antigen	A30000	4	1	56	179	0.014653
SPP1 P10451-5	JHU04492.B1C11R68	3	0	57	180	0.01504
HSBP1 Q75506	JHU02734.B3C3R48	3	0	57	180	0.01504
NID2 Q8IV28	JHU07642.B5C23R26	3	0	57	180	0.01504
TRAP1 Q12931	JHU08244.B5C14R42	3	0	57	180	0.01504
DCAF8 Q5TAQ9-2	JHU08154.B6C3R42	3	0	57	180	0.01504
TRAP1 Q12931	JHU08631.B6C29R48	3	0	57	180	0.01504
PALS1 Q8N3R9	JHU10617.B6C11R76	3	0	57	180	0.01504
NPM3 Q75607	JHU10230.B7C1R72	3	0	57	180	0.01504
KIAA0930 Q6ICG6	JHU10576.B8C1R78	3	0	57	180	0.01504
FAM104B Q5XKR9-2	JHU12283.B9C2R14	3	0	57	180	0.01504
AQP1 P29972	JHU14810.B9C30R54	3	0	57	180	0.01504
ZXDC Q2QGD7	JHU12054.B10C8R10	3	0	57	180	0.01504

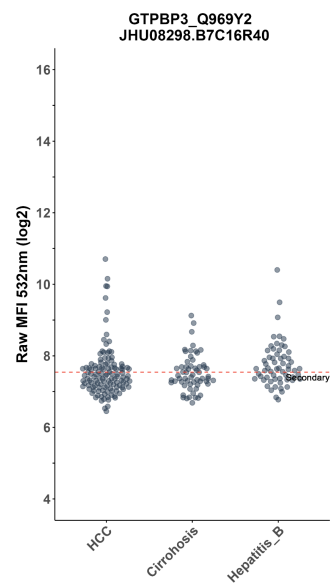
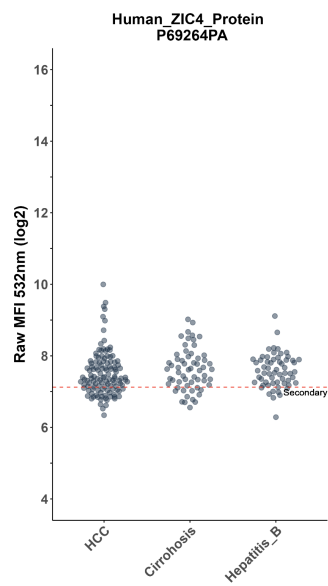
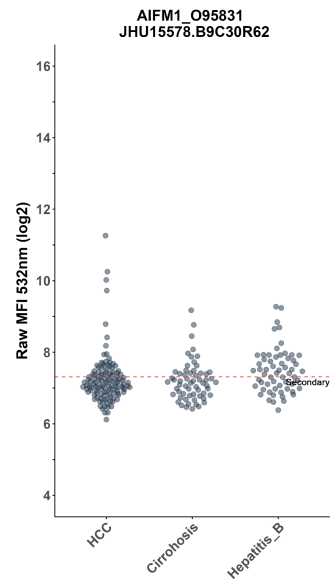
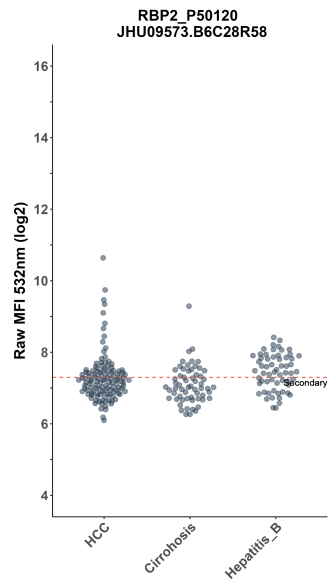
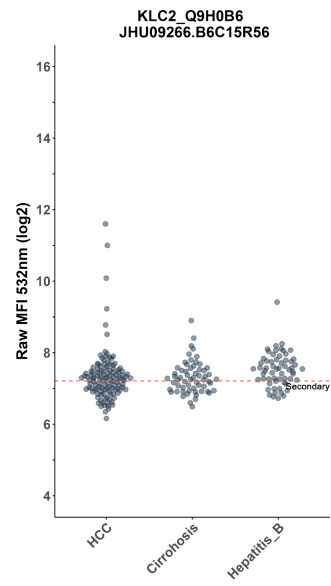
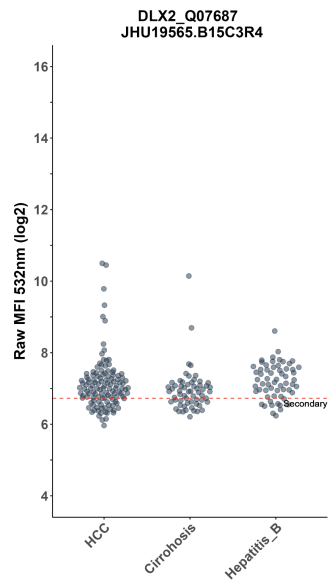
fisher_exact_tests/GC_IgG_IgA_Cirrohosis_vs_cohort_fisher_exact_tests_532nm_green.csv



GC IgG IgA HCC vs cohort 532nm

Name	ID	Hits Case	Hits Cohort	No Hits Case	No Hits Cohort	p value
DLX2 Q07687	JHU19565.B15C3R4	6	1	114	119	0.11951
KLC2 Q9H086	JHU09266.B6C15R56	4	0	116	120	0.12187
RBP2 P50120	JHU09573.B6C28R58	4	0	116	120	0.12187
AIFM1 Q95831	JHU15578.B9C30R62	4	0	116	120	0.12187
Human ZIC4 Protein	P69264PA	4	0	116	120	0.12187
GTPBP3 Q969Y2	JHU08298.B7C16R40	5	1	115	119	0.21284
GPRC5A Q8NFJ5	JHU02917.B1C1R44	3	0	117	120	0.24686
FAH P16930	JHU08579.B5C32R46	3	0	117	120	0.24686
DNMT3A Q9Y6K1-3	JHU05986.B7C25R2	3	0	117	120	0.24686
CCDC32 Q9BV29	JHU11437.B7C21R90	3	0	117	120	0.24686
ANKS6 Q68DC2	JHU09698.B8C27R64	3	0	117	120	0.24686
NUB1 Q9Y5A7-2	JHU13363.B9C18R28	3	0	117	120	0.24686
RAB11FIP1 Q6WKZ4-1	JHU15069.B9C32R58	3	0	117	120	0.24686
ATOX1 O00244	JHU11523.B10C26R4	3	0	117	120	0.24686
POLB P06746	JHU11865.B10C21R6	3	0	117	120	0.24686
AUP1 Q9Y679	JHU14175.B11C2R46	3	0	117	120	0.24686
HSPB11 Q9Y547	JHU04810.B16C8R80	3	0	117	120	0.24686
Human IL37 / IL1F7 / IL-1H4 Protein	10155-HNAE	3	0	117	120	0.24686
CLUAP1 Q96AJ1-2	JHU14927.B10C20R56	6	2	114	118	0.28074
DMTN Q08495	JHU04343.B4C8R70	4	1	116	119	0.36974
TCOF1 Q13428-5	JHU08626.B5C29R48	4	1	116	119	0.36974
PKNOX1 P55347	JHU08417.B7C30R38	4	1	116	119	0.36974
Human IGF1 / IGF-1 Protein	10598-HNAE	9	5	111	115	0.40973
OGFR Q9NZT2-2	JHU08317.B5C1R42	5	2	115	118	0.44617
HMGNS P82970	JHU01985.B15C13R80	5	2	115	118	0.44617
RWDD1 Q9H446	JHU01218.B4C15R20	11	7	109	113	0.46318
ATRX Q86U63	JHU02695.B1C15R46	2	0	118	120	0.49791
TBCA O75347	JHU04596.B1C10R70	2	0	118	120	0.49791
CCDC124 Q96CT7	JHU05007.B1C3R80	2	0	118	120	0.49791
PAIP1 Q9H074-2	JHU02075.B3C8R34	2	0	118	120	0.49791
NECTIN3 Q9NQ53	JHU03238.B3C8R54	2	0	118	120	0.49791
SMU1 Q2TAY7	JHU02377.B4C15R40	2	0	118	120	0.49791
CD2BP2 Q95400	JHU02601.B4C13R42	2	0	118	120	0.49791
CTBP2 P56545-3	JHU05493.B4C32R90	2	0	118	120	0.49791
TPM2 P07951-2	JHU06700.B5C22R16	2	0	118	120	0.49791
PIKFYVE Q9Y217-2	JHU08222.B6C22R40	2	0	118	120	0.49791
PPM1B O75688-2	JHU09378.B6C14R56	2	0	118	120	0.49791
KDELR3 Q43731	JHU06269.B7C27R12	2	0	118	120	0.49791
SLC47A1 Q96FL8-2	JHU06455.B7C28R12	2	0	118	120	0.49791
MX1 P20591	JHU08023.B7C10R34	2	0	118	120	0.49791
S1PR5 Q9H228	JHU08088.B7C5R40	2	0	118	120	0.49791
CYP4F11 Q9HBI6	JHU08567.B7C29R44	2	0	118	120	0.49791
NEDD9 Q14511	JHU10419.B7C12R76	2	0	118	120	0.49791
GPR61 Q9BZJ8	JHU07335.B8C13R30	2	0	118	120	0.49791
PTGES3 Q15185	JHU09474.B8C1R58	2	0	118	120	0.49791
PGBD1 Q96J53	JHU10799.B8C12R80	2	0	118	120	0.49791
TLE3 Q04726-6	JHU12338.B10C23R18	2	0	118	120	0.49791
FN3K Q9H479	JHU14546.B10C3R54	2	0	118	120	0.49791
ZFP36L2 P47974	JHU14705.B10C29R54	2	0	118	120	0.49791
PKM NA	JHU16124.B10C20R78	2	0	118	120	0.49791
S100A2 P29034	JHU11683.B11C32R4	2	0	118	120	0.49791
MAGEA6 P43360	JHU13255.B11C16R26	2	0	118	120	0.49791
ZNF410 Q86VK4	JHU13588.B11C22R36	2	0	118	120	0.49791
RBM39 E1P552	JHU11871.B12C13R6	2	0	118	120	0.49791
IDH1 Q075874	JHU12958.B12C31R22	2	0	118	120	0.49791
SRF P11831	JHU13293.B12C4R26	2	0	118	120	0.49791
PVALB P20472	JHU16554.B12C29R84	2	0	118	120	0.49791
ADIPOQ Q15848	JHU18972.B13C26R20	2	0	118	120	0.49791
FABP5 Q01469	JHU00992.B13C1R76	2	0	118	120	0.49791
MOSPD2 Q8NHP6	JHU18478.B14C31R46	2	0	118	120	0.49791
SLC16A2 P36021	JHU18127.B14C31R64	2	0	118	120	0.49791
HDAC3 O15379	JHU00426.B14C1R84	2	0	118	120	0.49791
RSPH6A Q9H0K4	JHU17380.B15C28R40	2	0	118	120	0.49791
PELO Q9BRX2	JHU02265.B15C21R84	2	0	118	120	0.49791
TCHHL1 Q5QJ38	JHU14890.B16C3R84	2	0	118	120	0.49791
ZHX1 Q9UKY1	JHU14358.B18C18R12	2	0	118	120	0.49791
IFNB1	10704-HNAS	2	0	118	120	0.49791
Human DBT Protein	P65514PA	2	0	118	120	0.49791
Human CNTF Protein (His Tag)	11841-H07E	2	0	118	120	0.49791
Human IL5 Protein	P00729QBFN	2	0	118	120	0.49791
BIN1 O00499-6	JHU10279.B16C9R80	6	3	114	117	0.49942
PNMA2 Q5U523	JHU15552.B10C6R62	7	4	113	116	0.53915
TPM1 P09493-10	JHU14060.B12C25R42	9	6	111	114	0.59528
SEC22C Q9BRL7	JHU08999.B2C14R64	10	7	110	113	0.61607
PNMA8A Q86V59	JHU03389.B1C21R50	3	1	117	119	0.62184
EIF2B5 Q13144	JHU08287.B6C11R38	3	1	117	119	0.62184
NUDT1 P36639-2	JHU09465.B6C16R60	3	1	117	119	0.62184
COL9A3 Q14050	JHU08659.B8C4R44	3	1	117	119	0.62184
IRF2BP1 Q8IU81	JHU12393.B9C27R16	3	1	117	119	0.62184
HEMGN Q9BXL5	JHU13532.B9C22R36	3	1	117	119	0.62184
ACSL4 O60488-2	JHU12255.B10C26R18	3	1	117	119	0.62184
FCHO1 O14526	JHU12286.B12C23R16	3	1	117	119	0.62184
DUSP12 Q9UN16	JHU00123.B14C24R80	3	1	117	119	0.62184
Capn15 O75808	JHU19944.B16C26R12	3	1	117	119	0.62184
TEX33 Q43247	JHU20055.B17C9R4	3	1	117	119	0.62184
USP13 Q92995	JHU21064.B17C11R18	3	1	117	119	0.62184
Human C1QA Protein	P00094PA	3	1	117	119	0.62184
MAB21L1 Q13394	JHU04362.B2C17R68	4	2	116	118	0.68353
CNKSR1 Q969H4	JHU08559.B6C8R48	4	2	116	118	0.68353
C5orf22 Q49AR2	JHU13991.B15C1R86	4	2	116	118	0.68353
RAD23A P54725-3	JHU16026.B9C31R68	5	3	115	117	0.72190
STAT5A P42229	JHU14691.B10C32R52	5	3	115	117	0.72190
MTURN Q8N3F0	JHU24542.B19C6R6	5	3	115	117	0.72190
KHDC4 Q727F0	JHU11460.B5C7R86	6	4	114	116	0.74862
ABTB1 Q969K4	JHU08641.B8C16R48	6	4	114	116	0.74862
NPM2 Q865E8	JHU13834.B10C5R40	6	4	114	116	0.74862
NCL P19338	JHU14856.B12C7R50	6	4	114	116	0.74862
HCLS1 P14317	JHU14116.B12C18R42	7	5	113	115	0.76857
HMGNS3 Q15651-2	JHU02816.B4C14R46	9	7	111	113	0.79674
NOL3 Q60936	JHU09079.B5C16R50	10	8	110	112	0.80717

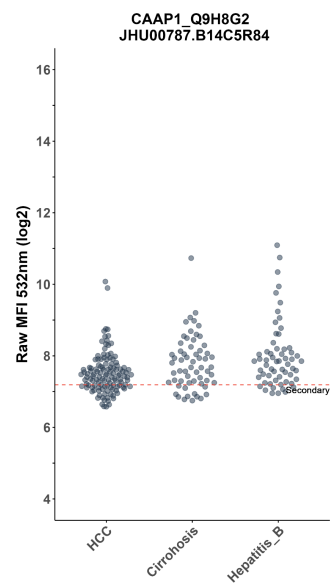
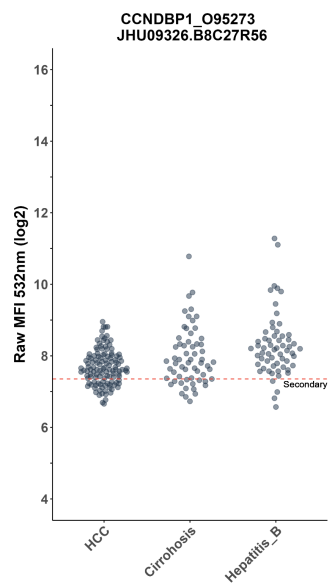
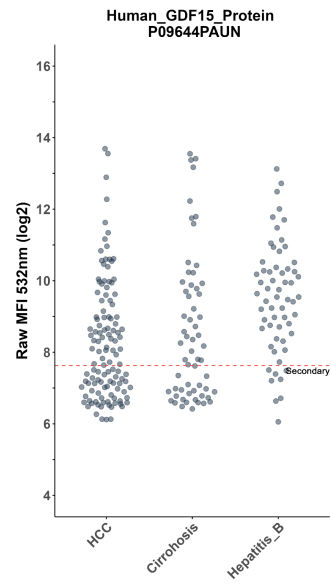
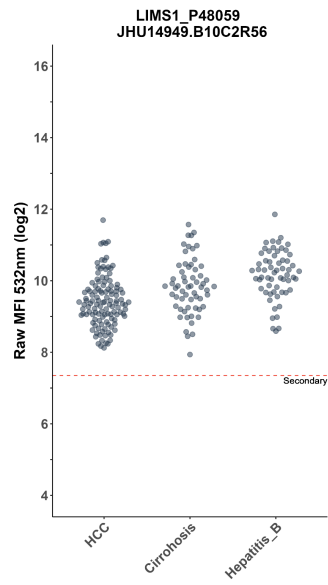
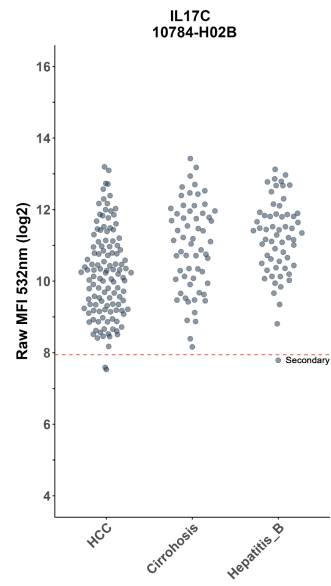
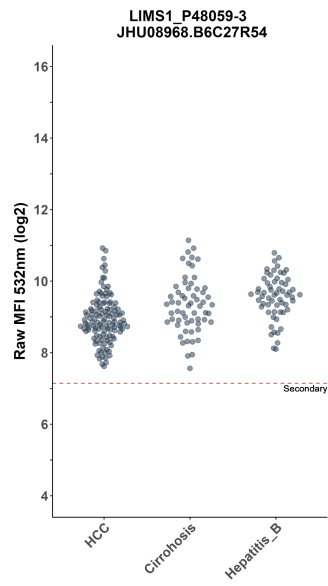
fisher_exact_tests/GC_IgG_IgA_HCC_vs_cohort_fisher_exact_tests_532nm_green.csv



GC IgG IgA Hepatitis B vs cohort 532nm

Name	ID	Hits Case	Hits Cohort	No Hits Case	No Hits Cohort	p value
LIMS1 P48059-3	JHU08968.B6C27R54	45	73	15	107	5.4028e-06
IL17C	10784-H02B	55	113	5	67	8.1876e-06
LIMS1 P48059	JHU14949.B10C2R56	52	106	8	74	6.6529e-05
Human GDF15 Protein	P09644PAUN	28	42	32	138	0.00094588
CCNDBP1 Q95273	JHU09326.B8C27R56	7	3	53	177	0.0028592
CAAP1 Q9H8G2	JHU00787.B14C5R84	7	3	53	177	0.0028592
NDRG2 Q9UN36	JHU20304.B20C10R12	7	3	53	177	0.0028592
ILRUN Q9H6K1	JHU02417.B1C9R38	4	0	56	180	0.0036172
CALD1 Q05682-4	JHU11915.B9C3R12	4	0	56	180	0.0036172
EID1 Q9Y6B2	JHU02047.B14C14R10	4	0	56	180	0.0036172
ABRACL Q9P1F3	JHU04812.B16C25R12	4	0	56	180	0.0036172
KASH5 Q8N6L0	JHU09052.B5C16R52	5	1	55	179	0.0041455
RBMS3 Q6XE24	JHU29173.B20C17R24	11	11	49	169	0.0083165
EIF4A2 Q14240-2	JHU05113.B3C8R84	5	2	55	178	0.011689
Human TNFSF11 Protein	P01905PFUN	30	56	30	124	0.012435
AQP5 P55064	JHU03649.B3C1R58	4	1	56	179	0.014653
MYBPH Q13203	JHU07744.B8C5R32	4	1	56	179	0.014653
GIMAP6 Q6P9H5	JHU07229.B5C18R22	3	0	57	180	0.01504
E2F2 Q14209	JHU07129.B7C1R22	3	0	57	180	0.01504
CTBP1 Q13363-2	JHU07120.B7C6R22	3	0	57	180	0.01504
SSU72 Q9NP77	JHU09485.B8C4R56	3	0	57	180	0.01504
SNX19 Q92543-2	JHU10446.B8C2R76	3	0	57	180	0.01504
LASP1 Q14847	JHU14849.B10C30R54	3	0	57	180	0.01504
SDE2 Q6IQ49	JHU16160.B10C22R78	3	0	57	180	0.01504
SCYL1 Q96KG9	JHU16467.B10C5R82	3	0	57	180	0.01504
MARCKSL1 P49006	JHU16451.B12C15R84	3	0	57	180	0.01504
EIF3F Q00303	JHU09253.B14C1R12	3	0	57	180	0.01504
SNX19 Q92543-2	JHU10446.B19C5R24	3	0	57	180	0.01504
SYAP1 Q96A49	JHU07852.B6C27R32	6	4	54	176	0.017515
CLIC3 Q95833	JHU12560.B9C18R14	6	4	54	176	0.017515
FAM114A2 Q9NRY5	JHU07409.B5C14R30	5	3	55	177	0.025136
PCMTD1 Q96MG8	JHU09086.B5C6R52	5	3	55	177	0.025136
PUS7L Q9H0K6	JHU08610.B7C14R46	10	12	50	168	0.035199
LZIC Q8WZA0	JHU10127.B8C9R68	4	2	56	178	0.035667
NACA2 Q9H009	JHU12216.B9C21R12	4	2	56	178	0.035667
TRIM16 Q95361	JHU13780.B12C13R34	4	2	56	178	0.035667
Human EXOSC10 Protein	P00461PA	4	2	56	178	0.035667
TMOD4 Q9NZQ9	JHU05172.B4C17R82	5	4	55	176	0.045654
C1orf94 Q6P1W5-2	JHU13411.B9C23R32	5	4	55	176	0.045654
SPRR3 Q9UBC9	JHU14973.B9C12R58	5	4	55	176	0.045654
MAP2 P11137	JHU25645.B20C26R12	5	4	55	176	0.045654
Human SLURP1 Protein	P04528PCUN	30	63	30	117	0.046838
ASNSD1 Q9NWL6	JHU03465.B2C20R58	3	1	57	179	0.049308
BAG6 P46379-2	JHU03468.B3C15R56	3	1	57	179	0.049308
POLR1D P0DPB6	JHU04579.B3C7R72	3	1	57	179	0.049308
PHPT1 Q9NRX4	JHU00544.B4C32R10	3	1	57	179	0.049308
STAT5B P51692	JHU04306.B4C23R70	3	1	57	179	0.049308
VP511 Q9H270	JHU07863.B5C32R32	3	1	57	179	0.049308
TBPL1 P62380	JHU09874.B5C16R64	3	1	57	179	0.049308
SNX19 Q92543-2	JHU10446.B5C20R76	3	1	57	179	0.049308
C2orf74 C9JB1	JHU10699.B6C13R78	3	1	57	179	0.049308
NME8 Q8N427	JHU08151.B7C23R38	3	1	57	179	0.049308
CTBP1 Q13363	JHU09138.B7C28R50	3	1	57	179	0.049308
CCR1 P32246	JHU07596.B8C15R28	3	1	57	179	0.049308
CLIC4 Q9Y696	JHU08556.B8C22R44	3	1	57	179	0.049308
CHCHD4 Q8N4Q1	JHU09244.B8C15R56	3	1	57	179	0.049308
CPLX2 Q6PUV4	JHU11920.B9C23R10	3	1	57	179	0.049308
SYAP1 Q96A49	JHU13677.B9C1R32	3	1	57	179	0.049308
CTBP1 Q4KMQ8	JHU15973.B9C21R72	3	1	57	179	0.049308
SH3D19 Q5HYK7-4	JHU16130.B10C2R74	3	1	57	179	0.049308
SULF1 JNA	JHU12043.B11C27R8	3	1	57	179	0.049308
RAB3IP Q96QF0-2	JHU13558.B12C9R36	3	1	57	179	0.049308
VCL P18206-2	JHU18520.B14C30R46	3	1	57	179	0.049308
CDR2L Q86X02	JHU19739.B16C19R2	3	1	57	179	0.049308
NUDT5 Q9UKK9	JHU02262.B19C12R16	3	1	57	179	0.049308
RSRP1 A0A024RAD9	JHU02215.B19C26R16	3	1	57	179	0.049308
SRRT Q9BXP5	JHU29258.B19C22R24	3	1	57	179	0.049308
TCEANC2 Q96MNS	JHU07308.B13C32R14	9	11	51	169	0.054506
PLS3 P13797	JHU00446.B1C3R8	2	0	58	180	0.061715
PMP2 P02689	JHU05248.B1C19R84	2	0	58	180	0.061715
KEAP1 Q14145	JHU02347.B2C26R40	2	0	58	180	0.061715
AFAP1L2 Q8N4X5-2	JHU00965.B3C16R16	2	0	58	180	0.061715
HMBS P08397	JHU02731.B3C2R44	2	0	58	180	0.061715
TMEM61 Q8N0U2	JHU03064.B3C28R46	2	0	58	180	0.061715
NT5C3A Q9H0P0-2	JHU04285.B3C20R68	2	0	58	180	0.061715
TXNDC17 Q9BRA2	JHU04604.B3C28R70	2	0	58	180	0.061715
HDGF P51858	JHU05224.B3C24R82	2	0	58	180	0.061715
KPNA2 P52292	JHU05045.B3C15R84	2	0	58	180	0.061715
SNCB Q16143	JHU02378.B4C3R40	2	0	58	180	0.061715
UBQLN4 D3DVA8	JHU02783.B4C23R48	2	0	58	180	0.061715
ASB9 Q96DX5-3	JHU04518.B4C13R68	2	0	58	180	0.061715
EIF2S1 P05198	JHU04439.B4C6R70	2	0	58	180	0.061715
UGP2 Q16851-2	JHU04697.B4C12R76	2	0	58	180	0.061715
CLEC1A Q8NC01	JHU05673.B4C30R90	2	0	58	180	0.061715
PKIB Q9C010	JHU06315.B5C9R10	2	0	58	180	0.061715
HDAC4 P56524-2	JHU07041.B5C9R22	2	0	58	180	0.061715
DPYSL2 Q16555	JHU07127.B5C1R24	2	0	58	180	0.061715
BC025318 JNA	JHU09833.B5C1R64	2	0	58	180	0.061715
ADPRHL1 Q8NDY3	JHU10752.B5C23R84	2	0	58	180	0.061715
FAM114A1 Q8IWE2	JHU05990.B6C25R6	2	0	58	180	0.061715
BCL10 Q95999	JHU06828.B6C21R14	2	0	58	180	0.061715
B3GALT4 Q96024	JHU07973.B6C21R36	2	0	58	180	0.061715
POLR2M POCAP2	JHU09061.B6C16R50	2	0	58	180	0.061715
ANKRD10 Q9NXR5-2	JHU08931.B6C17R52	2	0	58	180	0.061715
BLOC1S4 Q9NUP1	JHU10093.B6C32R68	2	0	58	180	0.061715
PIK3R1 P27986-3	JHU11097.B6C13R82	2	0	58	180	0.061715
DGCR6 Q14129	JHU05984.B7C25R4	2	0	58	180	0.061715
PSAT1 Q9Y617	JHU07936.B7C25R34	2	0	58	180	0.061715
RPRD1B Q9NQG5	JHU08173.B7C29R42	2	0	58	180	0.061715
DNAJB8 Q8NH50	JHU09045.B7C11R52	2	0	58	180	0.061715

fisher_exact_tests/GC_IgG_IgA_Hepatitis_B_vs_cohort_fisher_exact_tests_532nm_green.csv



15 Summary Heatmaps: Top Public

Unbiased summary heatmaps of the top public reactive proteins from the Top_Public_635nm_red.csv and Top_Public_532nm_green.csv output rows. Full resolution figures are provided in the **summary_heatmaps** folder. These figures are often good at independently identifying and clustering individual patient replicates via observation of the study's highest frequency hits. Note: most public epitopes are unrelated with conditions of interest and this plot rarely clusters disease groups.

All visualized hits are log2 Z-scores from the **Rawlog2MFI_ZScore_Hits** semi-continuous hits data. Non-passing values are set to zero.

summary_heatmaps/Top_Public_channel_unclustered_heatmap.png

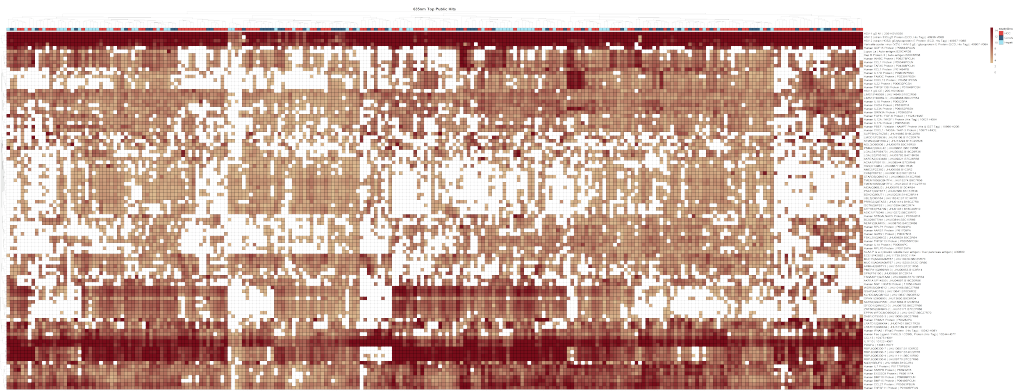
Created in R with the *pheatmap* package and no clustering. Samples and groups are arranged according to the *studyManifest*.

summary_heatmaps/Top_Public_channel_clustered_heatmap.png

Created in R with the *pheatmap* package with default row and column clustering settings (Euclidean distance as the similarity measure and clustering samples based on the 'complete' method).

summary_heatmaps/Top_Public_channel_row_clustered_heatmap.png

Created in R with the *pheatmap* package with default row clustering settings (Euclidean distance as the similarity measure). Samples and groups are unclustered and remain arranged according to the *studyManifest*.



summary_heatmaps/Top_Public_635nm_clustered_heatmap.png

16 Summary Heatmaps: Fisher Exact Tests

Summary heatmaps of each of the top 100 protein-level Fisher exact test outliers shown on the preceding pages. Full resolution figures are provided in the **summary_heatmaps** folder. These figures are all directional for top p-values associated with each test and attempt to visualize co-occurring autoantibody signature subgroups.

All visualized hits are log2 Z-scores from the **Rawlog2MFI_ZScore_Hits** semi-continuous hits data. Non-passing values are set to zero.

summary_heatmaps/filename_unclustered_heatmap.png - One created for each directional protein-level test in the **fisher_exact_tests** folder. Created in R with the *pheatmap* package and no clustering. Samples and groups are arranged according to the *studyManifest*.

summary_heatmaps/filename_clustered_heatmap.png - One created for each directional protein-level test in the **fisher_exact_tests** folder. Created in R with the *pheatmap* package with default row and column clustering settings (Euclidean distance as the similarity measure and clustering samples based on the 'complete' method).

summary_heatmaps/filename_row_clustered_heatmap.png - One created for each directional protein-level test in the **fisher_exact_tests** folder. Created in R with the *pheatmap* package with default row clustering settings (Euclidean distance as the similarity measure). Samples and groups are unclustered and remain arranged according to the *studyManifest*.

17 UniProt BLAST Annotations

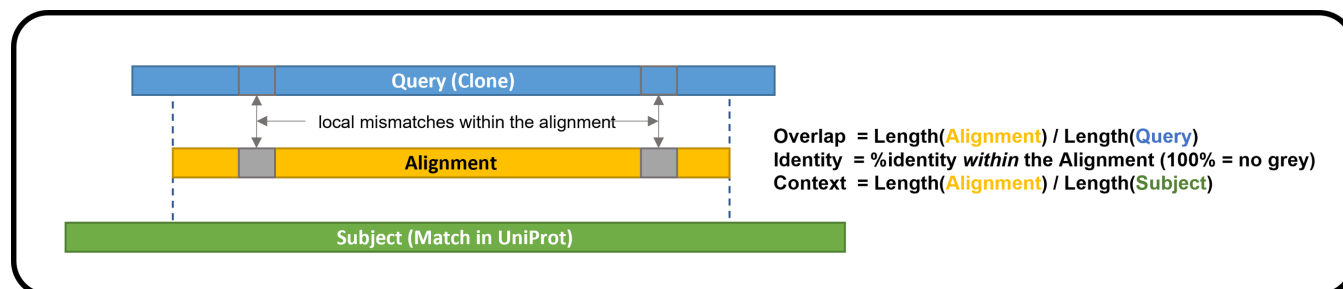
The most trustworthy part of a HuProt collection annotation is its JHU (Johns Hopkins University) identifier and protein sequence. These protein sequence validations and the initial gene name annotations used in *.gpr* microarray data files were performed and reported as-described in 2018.⁴ While parent *.gpr* files retain this original naming structure, databases such as UniProt now include improved names for some protein sequences.

In order to modernize the labels for HuProt sequences, BLAST versus UniProt was performed in 2022. All deliverables and output data described in this report use these updated labels; protein sequences and JHU identifiers remain unchanged. Summary data from this renaming process, including 'best mouse proteome labels' for the HuProt collection are provided in the **annotations** folder as **HuProt_BLAST_Nomenclature.xlsx**.

To update collection names, reference protein sequences were downloaded from UniProt. Reference sequences encompassed the manually reviewed SWISS-PROT and computer-annotated TrEMBL databases across both human and mouse proteomes (downloaded on 2/15/22).

https://ftp.uniprot.org/pub/databases/uniprot/previous_releases/release-2021_04/

To compare JHU sequences from the HuProt collection to these reference sequence databases, we used NCBI BLAST suite (Protein-Protein BLAST 2.12.0+; downloaded on 2/15/22) to compute *Overlap*, *Identity*, *Context*, and *E-Value* scores versus each of four libraries (human SWISS-PROT, human TrEMBL, mouse SWISS-PROT, mouse TrEMBL). An overview of the full alignment, ranking, and preferred match selected is described below:



Gene: Gene Name for best match.

Entrez: Entrez Gene ID for best match.

UniProt: Accession of the best match (TrEMBL or SP).

Perfect? Is there a match where the full sequences agree?

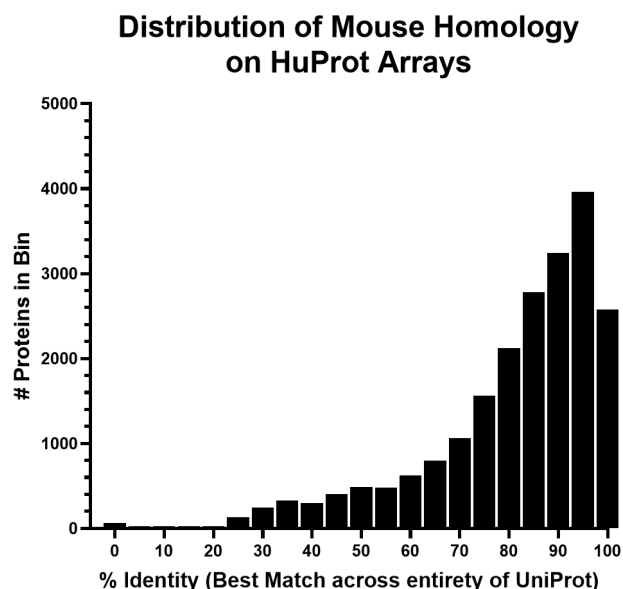
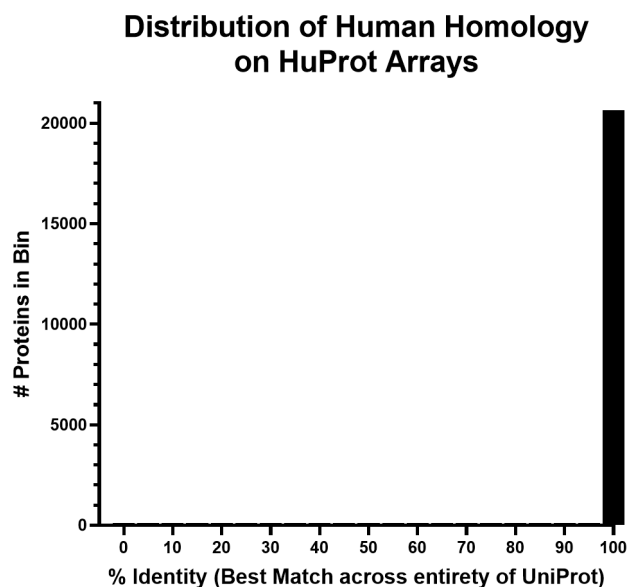
Overlap: The percentage of the JHU clone that participates in the overlap.

Identity: The percentage of matching JHU clone identity in the overlapping region.

Context The percentage of the UniProt subject sequence covered by the overlap.

SWISS-PROT entries represent a high quality manually annotated and non-redundant protein sequence database and were preferred for final annotations. Where only poor matches were available, original 2018 clone labels were retained. Best HuProt names are recorded under the section *Preferred Naming Scheme* for Gene Name, UniProt, and Entrez identifiers. Official Gene Name nomenclature for human and mouse was obtained from NCBI (downloaded on 2/18/22).

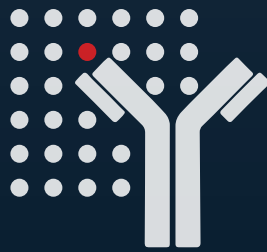
The collection was confirmed to be primarily human proteins, isoforms, and protein fragments with the *Identity* of most sequences aligning to the human UniProt libraries; high homology was also observed versus the mouse UniProt libraries. Low *Identity* scores can be used to identify occasional poorly-aligning proteins without strong homolog in UniProt via the **HuProt_BLAST_Nomenclature.xlsx** file.



The collection was confirmed to be primarily composed of full-length human proteins and isoforms with 19,330 / 21,197 (91.2%) sequences having a >98% simultaneous match for *Overlap*, *Identity*, and *Context* across UniProt. This indicates most HuProt proteins are well-matched to public database entries and are untruncated. Notable exceptions include exceptionally long sequences such as Titin (JHU15762; TTN), where a 5.4% *Context* score indicates we only create a small fragment of this protein; low *Context* scores can be used to identify other truncated proteins via the **HuProt_BLAST_Nomenclature.xlsx** file.

18 References

1. H. Zhu, M. Bilgin, R. Bangham, D. Hall, A. Casamayor, P. Bertone, N. Lan, R. Jansen, S. Bidlingmaier, T. Houfek, T. Mitchell, P. Miller, R.A. Dean, M. Gerstein, M. Snyder, Global Analysis of Protein Activities Using Proteome Chips, *Science*. 293 (2001) 2101–2105. <https://doi.org/10.1126/science.1062191>.
2. J.S. Jeong, L. Jiang, E. Albino, J. Marrero, H.S. Rho, J. Hu, S. Hu, C. Vera, D. Bayron-Poueymiroy, Z.A. Rivera-Pacheco, L. Ramos, C. Torres-Castro, J. Qian, J. Bonaventura, J.D. Boeke, W.Y. Yap, I. Pino, D.J. Eichinger, H. Zhu, S. Blackshaw, Rapid Identification of Monospecific Monoclonal Antibodies Using a Human Proteome Microarray, *Molecular & Cellular Proteomics*. 11 (2012) O111.016253. <https://doi.org/10.1074/mcp.O111.016253>.
3. S. Hu, Z. Xie, A. Onishi, X. Yu, L. Jiang, J. Lin, H. Rho, C. Woodard, H. Wang, J.-S. Jeong, S. Long, X. He, H. Wade, S. Blackshaw, J. Qian, H. Zhu, Profiling the Human Protein-DNA Interactome Reveals ERK2 as a Transcriptional Repressor of Interferon Signaling, *Cell*. 139 (2009) 610–622. <https://doi.org/10.1016/j.cell.2009.08.037>.
4. A. Venkataraman, K. Yang, J. Irizarry, M. Mackiewicz, P. Mita, Z. Kuang, L. Xue, D. Ghosh, S. Liu, P. Ramos, S. Hu, D. Bayron Kain, S. Keegan, R. Saul, S. Colantonio, H. Zhang, F.P. Behn, G. Song, E. Albino, L. Asencio, L. Ramos, L. Lugo, G. Morell, J. Rivera, K. Ruiz, R. Almodovar, L. Nazario, K. Murphy, I. Vargas, Z.A. Rivera-Pacheco, C. Rosa, M. Vargas, J. McDade, B.S. Clark, S. Yoo, S.G. Khambadkone, J. De Melo, M. Stevanovic, L. Jiang, Y. Li, W.Y. Yap, B. Jones, A. Tandon, E. Campbell, G.T. Montelione, S. Anderson, R.M. Myers, J.D. Boeke, D. Fenyö, G. Whiteley, J.S. Bader, I. Pino, D.J. Eichinger, H. Zhu, S. Blackshaw, A toolbox of immunoprecipitation-grade monoclonal antibodies to human transcription factors, *Nat Methods*. 15 (2018) 330–338. <https://doi.org/10.1038/nmeth.4632>.



THE AUTOANTIBODY SEROMICS COMPANY

CDI Labs

12 W Mendez Vigo Avenue
2nd Floor
Mayagüez, PR 00680
USA

Sales

Email: info@cdilabs.com

Support

Email: support@cdilabs.com

FOR RESEARCH USE ONLY. Not for use in diagnostic procedures.

© **CDI Labs**. All rights reserved.

CDI Labs and HuProt™ are trademarks of CDI Labs.

