SUPPLEMENTARY INFORMATION

Xiaobo Yu[1#*], Lusheng Song[2#*], Brianne Petritis[2], Xiaofang Bian [2], Haoyu Wang[2], Jennifer Viloria[2], Jin Park[2], Hoang Bui[2], Han Li[2], Jie Wang[2], Lei Liu[1], Liuhui Yang[1], Hu Duan[1],David N. McMurray[3], Jacqueline M. Achkar[4], Mitch Magee[2], Ji Qiu[2] and Joshua LaBaer[2*]

[1]State Key Laboratory of Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences (PHOENIX Center, Beijing), Beijing Institute of Radiation Medicine, Beijing, 102206, China

[2]The Virginia G. Piper Center for Personalized Diagnostics, Biodesign Institute, Arizona State University, Tempe, AZ 85287, USA

[3]Department of Microbial Pathogenesis and Immunology, College of Medicine, Texas A&M Health Science Center, College Station, TX 77843, USA.

[4]Department of Medicine, Albert Einstein College of Medicine, NY 10461, USA ; Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, NY 10461, USA.

[#]These authors contributed equally to this work.

*Correspondence should be addressed to X.Y. (yuxiaobo@mail.ncpsb.org) and J.L. (Joshua.Labaer@asu.edu)

Supplemental Methods

Screening for protein-specific antibodies with protein microarrays

We collected 45 antibodies targeting 32 proteins from a various sources (see also Materials and Methods) and tested their binding specificity on NAPPA and M-NAPPA arrays. We printed the gene plasmids encoding for those proteins on small NAPPA and M-NAPPA arrays. After protein expression using a human cell-free expression system, the arrays were incubated with antibodies diluted 1:200 or 1:300 for 1 hour at room temperature. After washing three times with PBST, antibodies that were bound to their target proteins on the array were detected by their corresponding fluorescein-labeled IgG secondary antibodies. The data were calculated using the ratio of fluorescent intensity of proteins divided by the negative control spot, which contained only the printing mixture with no plasmid cDNA (**Figure S1)**. Collectively, we found 8 antibodies {IA-2 (106 kDa), GAD2 (65 kDa), Clusterin (52 kDa), p53 (44 kDa), Fos (40 kDa), PP2A (36 kDa), SFN (28 kDa) and BCL2L2 (20 kDa)} that specifically recognized their target proteins on NAPPA and M-NAPPA arrays.

Supplemental Tables

Table S1. Percentage of hits identified from the HT screening in previously published reports.

Classification	Species	Method	Total proteins	Hits	Hit rate(%)	Verification	Year	PMID
Protein function	Arabidopsis	HaloTag-NAPPA human protein microarray	12,000	59	0.49	in vitro and vivo protein-protein interaction assays	2016	27357687
	Human	NAPPA human protein microarray	10,000	19	0.19	in vitro and vivo protein-protein interaction assays	2015	25739981
	E.coli	E.coli proteome microarray	4,256	4	0.09	in vitro and vivo acetylation assays	2015	26716769
	Human	Human protein microarray	16,368	360	2.20	in vitro and vivo functional assays	2015	26598702
	Human	Human protein microarray	4,191	58	1.38	in vitro kinase assays	2015	26109723
	Human	NAPPA human protein microarray	10,000	20	0.20	in vitro and vivo AMPylation assays	2014	25073739
	Human	NAPPA human protein microarray	10,000	11	0.11	in vitro and vivo protein-protein interaction assays	2014	24955142
	Human	Human protein microarray	2,236	29	1.30	in vitro kinase assays	2013	24023761
	Yeast	Yeast proteome microarray	5,800	91	1.57	in vitro activity assays	2009	19303850
	Human	Human protein microarray	8000	324	4.05	ELISA	2011	21244100
	Yeast	Yeast proteome microarray	4200	9	0.21	in vitro activity assays	2011	21460040
Autoantibody biomarker	Human	NAPPA human protein microarray	10,000	17	0.17	RAPID-ELISA	2016	26896032
	Human	NAPPA human protein microarray	10,000	6	0.06	RAPID-ELISA	2016	27690455
	Human	Human protein microarray	16,368	17	0.10	Small protein array	2016	26598640
	Human	NAPPA human protein microarray	10,000	748	7.48	NAPPA array	2015	26070530
	Human	NAPPA human protein microarray	10,000	741	7.41	RAPID-ELISA	2015	25365139
	Human	Human protein microarray	17,000	211	1.24	None	2015	26370624
	Human	Human protein microarray	17,000	137	0.81	Small protein array	2015	25954975
	Human	Human protein microarray	11,520	51	0.44	Bead based array	2013	23732997
	Human	NAPPA human protein microarray	4,388	71	1.62	Immunoblot	2012	22311593
	Human	NAPPA human protein microarray	4,988	761	15.26	NAPPA array	2011	20977275
	Human	Human protein microarray	10,846	14	0.13	Bead based array	2016	27206786
	Human	Human protein microarray	3,786	161	4.25	Bead based array	2016	27700103
	Human	Human protein microarray	3,072	131	4.26	Bead based array	2015	25608002

	Human	Human protein microarray	3,840	373	9.71	Bead based array	2014	25227461
	Human	Human protein microarray	17,000	23	0.14	Small protein array	2012	22647870
	Human	Human protein microarray	5056	38	0.75	ELISA	2011	21183621

Table S2. Hits identified from viral M-NAPPA arrays.

No	Name	Ratio[Rb1/Control]*	ring score
1	EBV-EBNA3A, H3N2-HA, HCMV-US34, HPV18-E7, Adenovirus-IX	2.7	5.0
2	EBV-BVLF1, SFV-E1, HCMV-US10, HCMV-UL36, HPV6-E7	2.7	3.0
3	EBV-BHLF1, H3N2-NS2, HCMV-UL88, HPV16-E7, SV40-Small t	2.5	5.0
4	EBV-BBLF1, H3N2-PA, HCMV-UL50, HPV16-E4, HPV33-E7	3.4	5.0
5	EBV-BARF0, H3N2-M1, HCMV-UL13, HPV18-E4, Adenovirus-E1A	3.1	5.0
6	EBV-BALF2, H1N1-M1, HCMV-UL141, HCMV-UL38, HPV11-E7	2.5	3.0

* The ratio was calculated by using the normalized signal of Rb1-HaloTag to the negative control expressing only HaloTag fusion protein. Proteins that were considered to be potential antigens of diseases had a ratio higher than 2 and a ring score higher than 1.

Table S3. Hits identified from viral NAPPA arrays.

No.	Name	Ratio[Rb1/Control]*	ring score
1	Adenovirus-E1A	2.8	5.0
2	HPV11-E7	4.1	3.0
3	HPV16-E7	2.3	5.0
4	HPV18-E7	2.4	5.0
5	HPV33-E7	2.8	5.0

* The ratio was calculated by using the normalized signal of Rb1-HaloTag to the negative control expressing only HaloTag fusion protein. Proteins that were considered to be potential antigens of diseases had a ratio higher than 2 and a ring score higher than 1.

Table S4. Coefficient of variation (CV) and standard deviation (SD) of M-NAPPA and NAPPA protein array fabrication.

	M-NAPPA				NAPPA			
	Spot number	**Gene number**	**Mean CV**	**SD**	**Spot number**	**Gene number**	**Mean CV**	**SD**
Spot-to-spot	32	80	3.64%	3.27%	160	80	7.63%	10.58%
Zone-to-zone	64	80	7.57%	3.41%	320	80	12.13%	7.56%
Slide-to-slide	128	80	7.27%	4.00%	640	80	13.25%	9.42%

Table S5. Coefficient of variation (CV) and standard deviation (SD) of protein-protein interaction assays.

	M-NAPPA				NAPPA			
	Spot number	**Interaction pair**	**Mean CV**	**SD**	**Spot number**	**Interaction pair**	**Mean CV**	SD
Spot-to-spot	10	5	2.55%	2.56%	10	5	5.65%	2.69%
Zone-to-zone	20	5	3.11%	3.46%	20	5	5.75%	3.86%

Supplemental figures

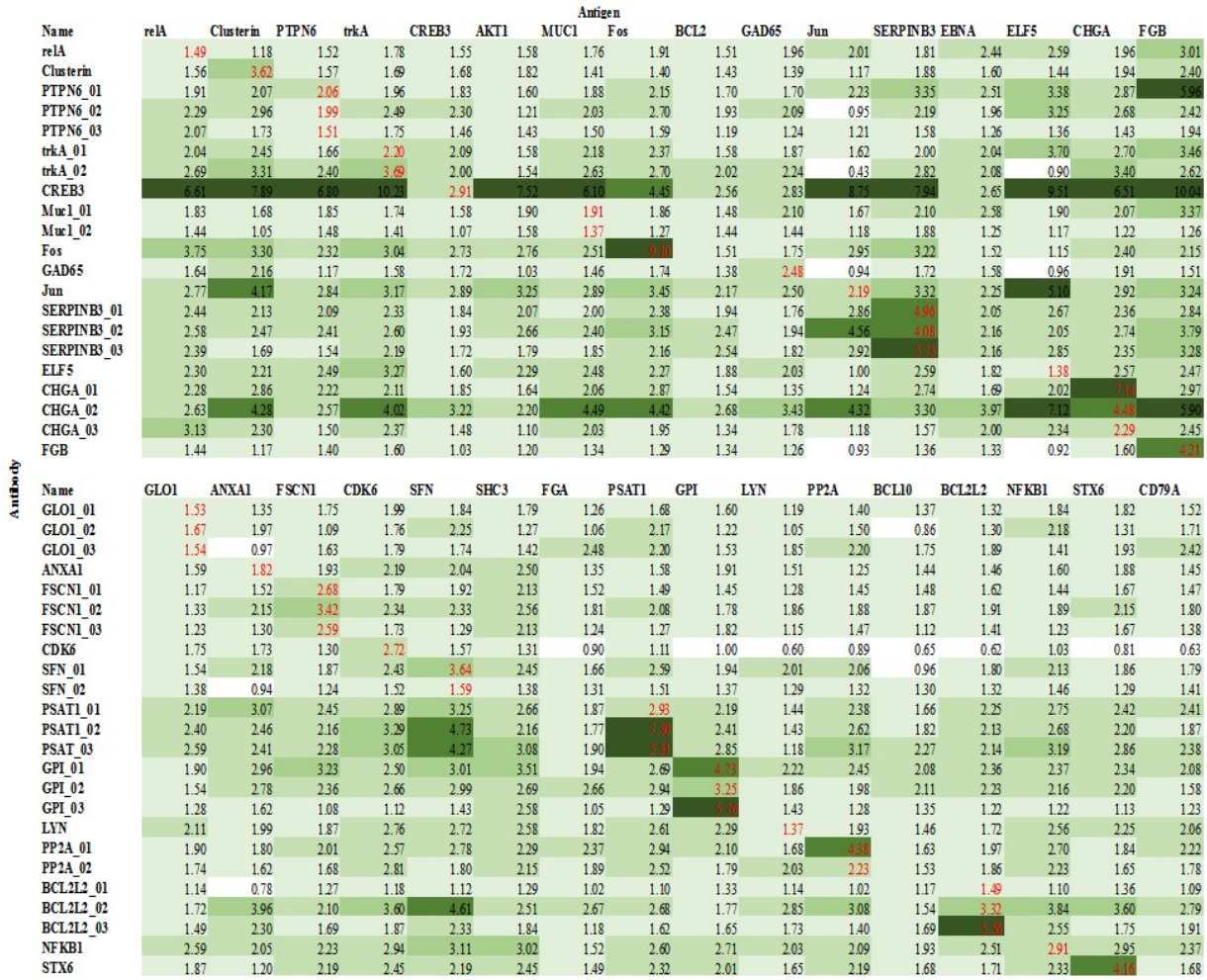


Figure S1. Screening for protein-specific antibodies with M-NAPPA. Binding between antibodies (y-axis) and their specific target antigens on the array (x-axis) was detected with an HRP-labeled goat anti-mouse or rabbit IgG secondary antibody as previously described. Numbers indicate the normalized signal intensity; those in red are the intensity when the antibody bound to its specific target antigen. Cell color corresponds to the level of normalized signal intensity, such that white reflects low normalized signal intensity and dark green reflects high normalized signal intensity.

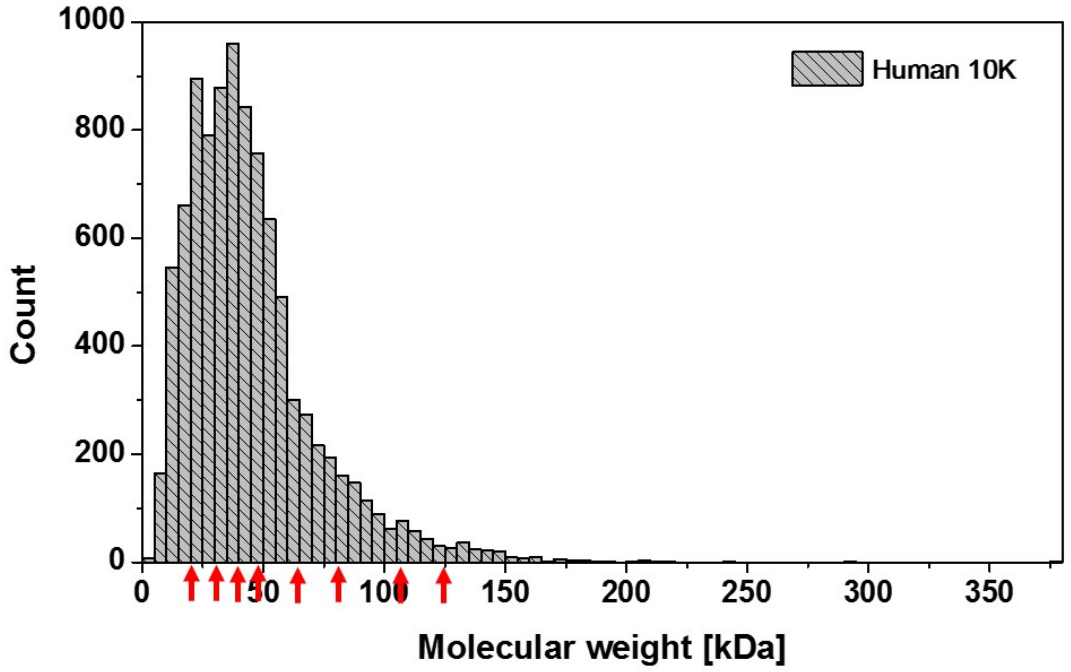


Figure S2. Size distribution of 10k human proteins expressed and displayed on M-NAPPA. The red arrows indicate proteins across a wide range of molecular weights that were further examined for protein expression and display.

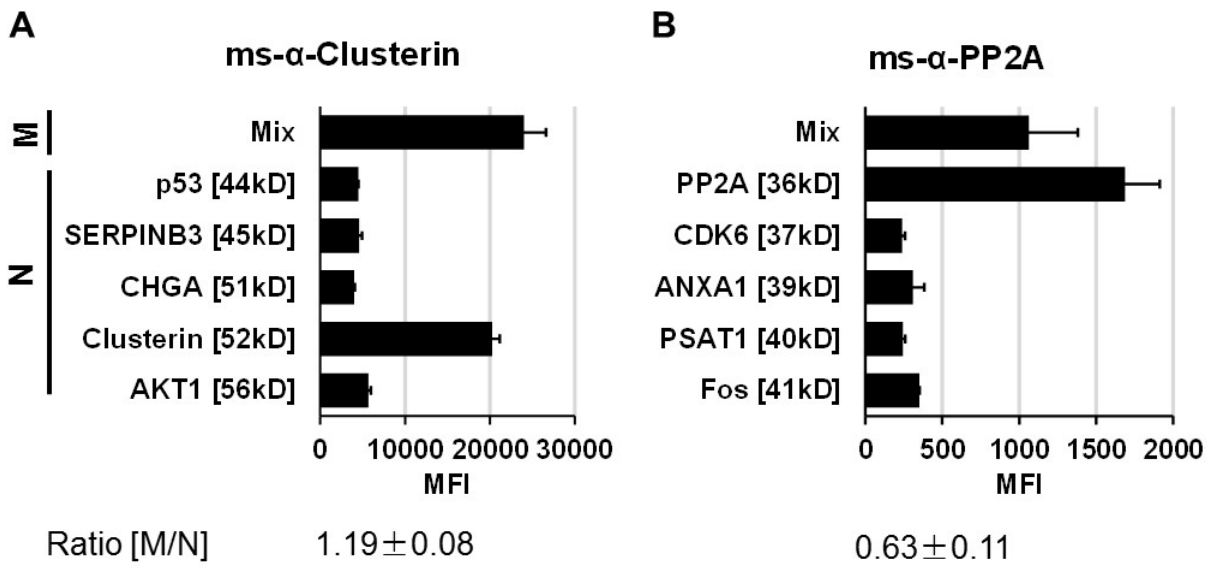


Figure S3. Protein-specific antibodies can detect their target antigens on NAPPA and M-NAPPA. A) An antibody to Clusterin probed multiplexed and non-multiplexed features displaying p53, SERPINB3, CHGA, Clusterin, and AKT1 on M-NAPPA and NAPPA, respectively. B) An antibody to PP2A probed multiplexed and non-multiplexed features displaying PP2A, CDK6, ANXA1, PSAT1, and Fos on M-NAPPA and NAPPA, respectively. MFI = mean of fluorescent intensity. ms = mouse. N = NAPPA with each spot displaying one protein. M = M-NAPPA with each spot displaying five different proteins. Error bars are the standard deviations of fluorescent signal intensity for triplicate spots on the arrays.

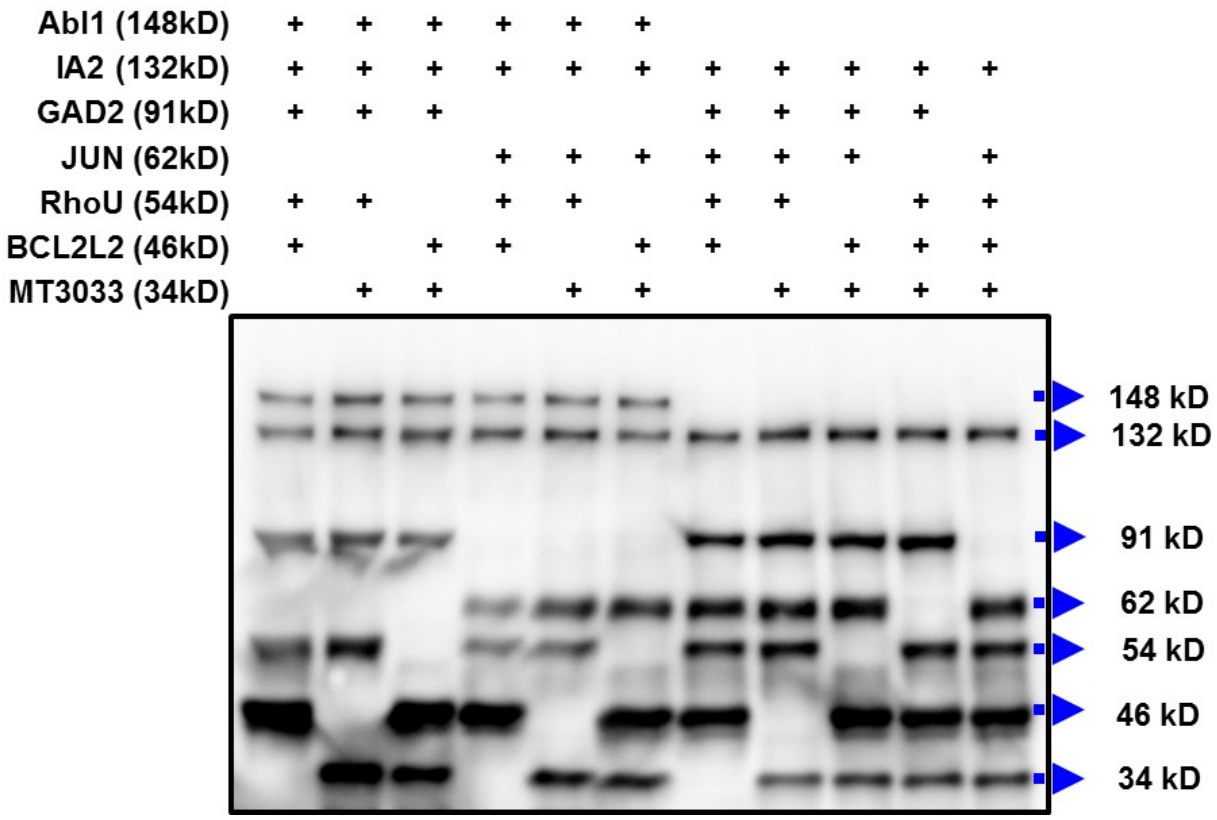


Figure S4. Western blot analyses of proteins with various molecular weights co-expressed in cell-free extract. All combinations of plasmid cDNA encoding for five proteins of varying molecular weights expressed in human cell-free lysate and at similar levels. GST-tagged proteins (Abl1, IA-2, GAD2, Jun, RhoU, BCL2L2 and MT3033) were co-expressed in human cell-free extract, and probed with a mouse anti-GST antibody and then an HRP-labeled goat anti-mouse IgG antibody.

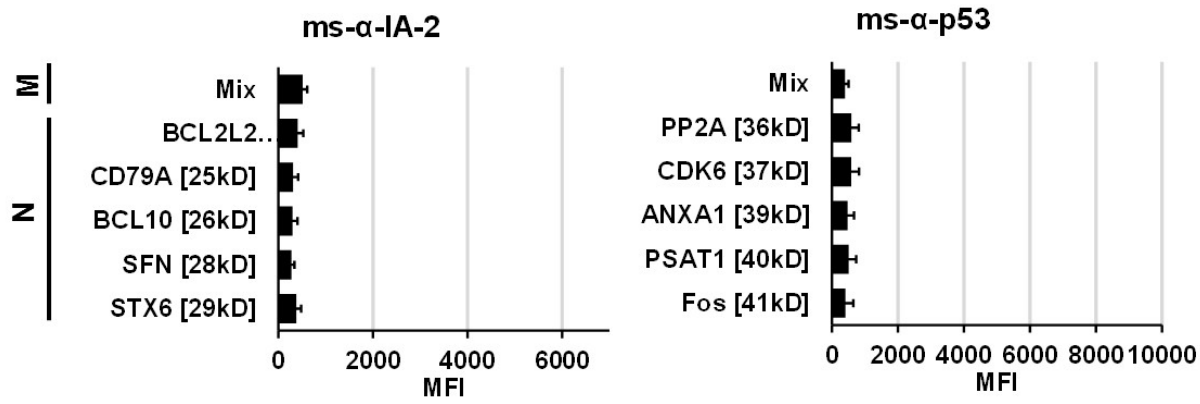


Figure S5. M-NAPPA and NAPPA have similar background signal. M-NAPPA and NAPPA that did not display IA-2 or p53 proteins were probed with anti-IA-2 and anti-p53 antibodies. Multiplexed and non-multiplexed spots had similar background signal from low, nonspecific antibody binding (see also Figure S3). MFI = mean of fluorescent intensity. ms = mouse. N = NAPPA with each spot displaying one protein. M = M-NAPPA with each spot displaying five different proteins. Error bars are the standard deviations of fluorescent signal intensity for triplicate spots on the arrays.

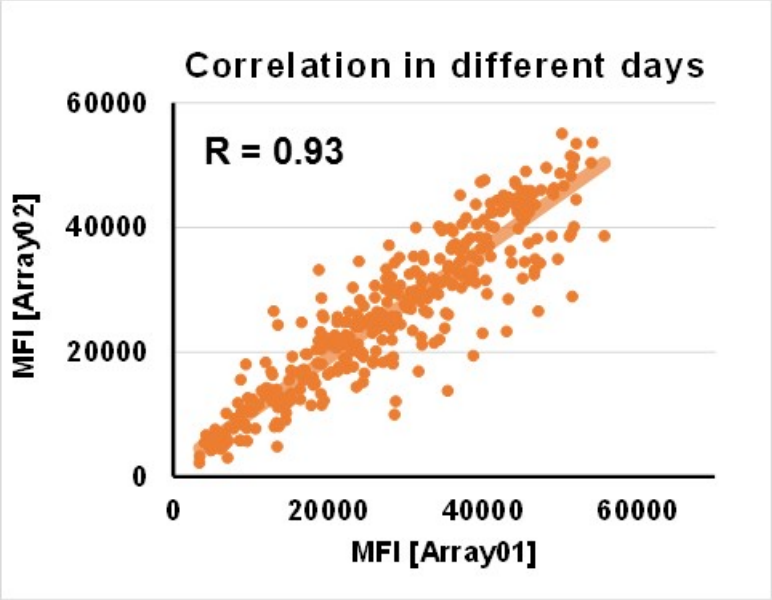


Figure S6. High correlation of 10k human proteins expressed and displayed on M-NAPPA viral arrays on two separate days. Displayed GST-tagged proteins were detected with a mouse anti-GST antibody and an Alexa Fluor 555 labeled anti-mouse secondary antibody. MFI = mean of fluorescent intensity.

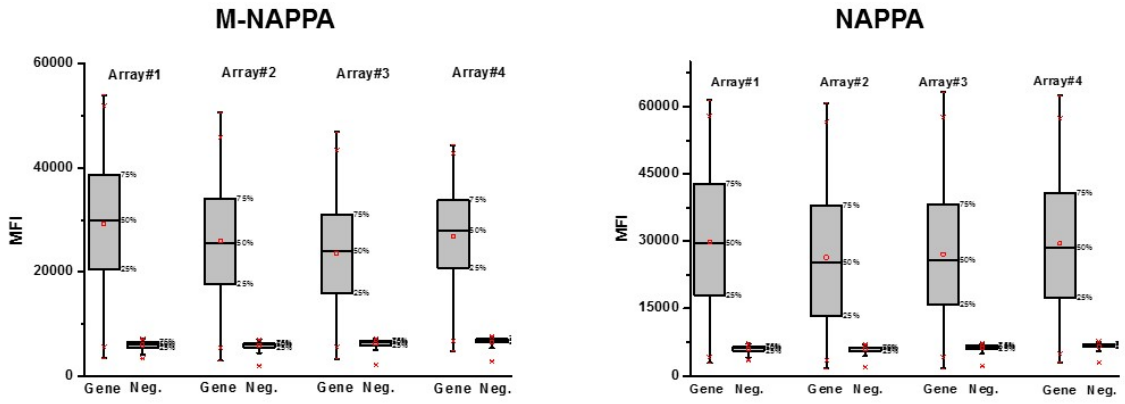


Figure S7. Boxplots of the MFI of GST-tagged proteins on M-NAPPA and NAPPA indicate that the majority of proteins are displayed on both array platforms.

Displayed GST-tagged proteins ("Gene") were detected with a mouse anti-GST antibody and an Alexa Fluor 555 labeled anti-mouse secondary antibody. Non-spots printed only with printing buffer ("Neg.") were used as negative controls. MFI = mean of fluorescent intensity.

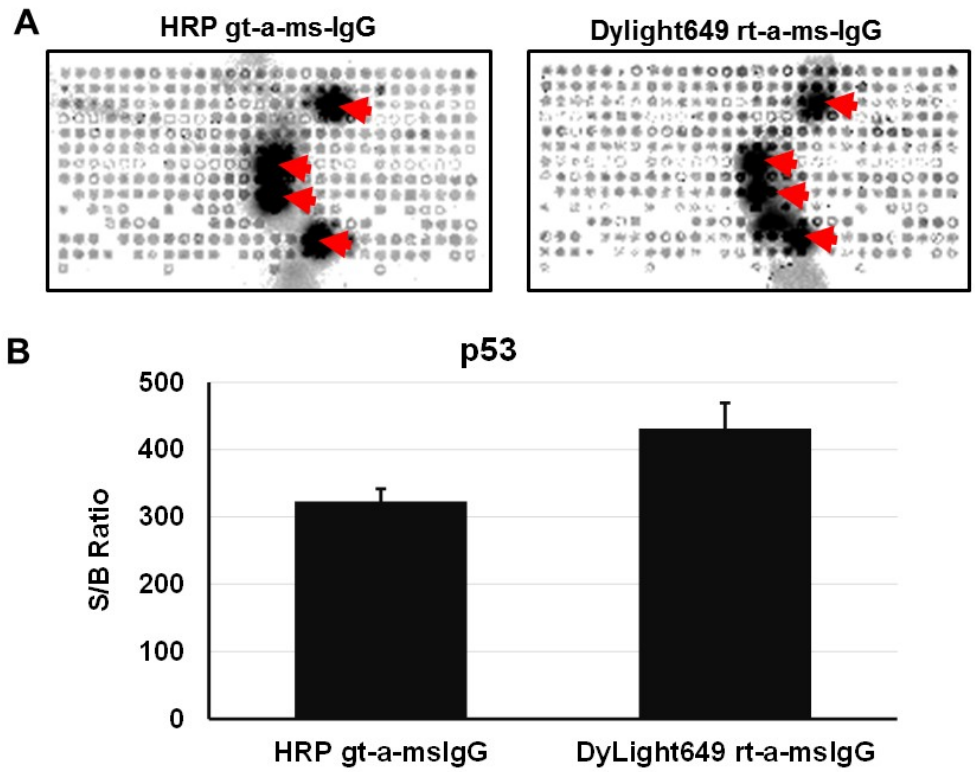


Figure S8. Direct fluorescent detection had a higher signal-to-background ratio than the TSA method on M-NAPPA. A) High contrast black-and-white images of M-NAPPA arrays probed with a mouse anti-p53 antibody and an HRP-conjugated secondary antibody ("HRP gt-a-ms-IgG") or fluorescently-labeled secondary antibody ("DyLight649 rt-a-msIgG"). Large black spots indicate antibody binding, which are the locations of the p53 antigen. B) Average S/B (signal to background) fluorescent mean intensity across four replicate p53 spots, where the mean signal of the p53 spots was divided by the background signal of the non-spots containing only the printing master mix. Error bars represent the range of S/B across quadruplet spots.

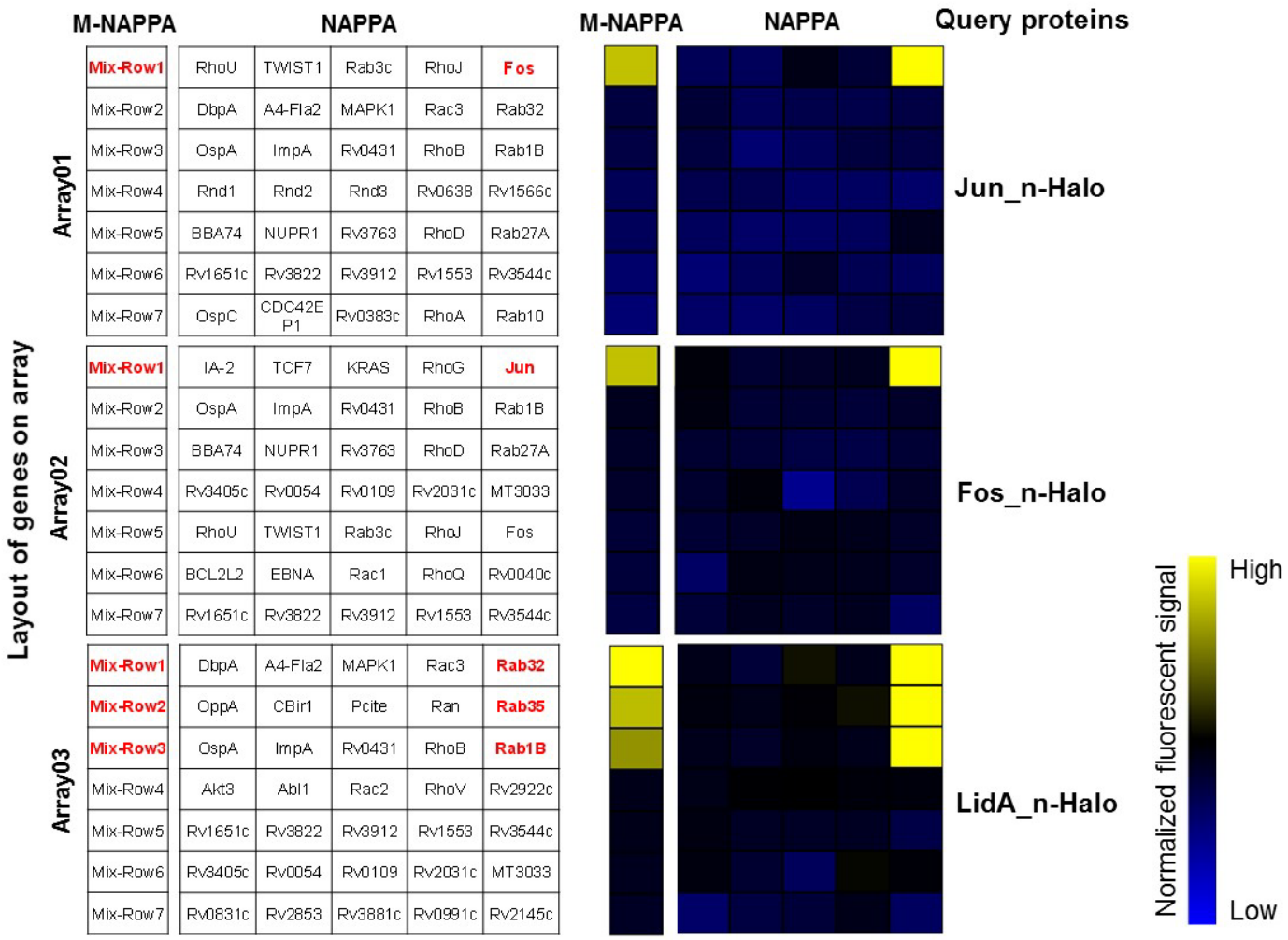


Figure S9. M-NAPPA can detect the same known protein-protein interactions as NAPPA arrays. 35 different GST-tagged proteins were expressed and displayed on NAPPA and M-NAPPA arrays (array layouts on left), and then probed with N-terminal Halo Tagged Jun, Fos and LidA proteins. Interactions between Jun, Fos, and LidA with their known binding partners on the arrays were detected with a chicken anti-HaloTag antibody and then an Alexa Fluor 555 goat anti-chicken secondary antibody. The false-colored blue and yellow images on right correspond to the normalized fluorescent signal from low to high, respectively. The signal of each feature was calculated from six

replicate spots and then normalized to the median signal of all microarray spots. Antigens that are known interactors of Jun, Fos, and LidA are indicated in red.

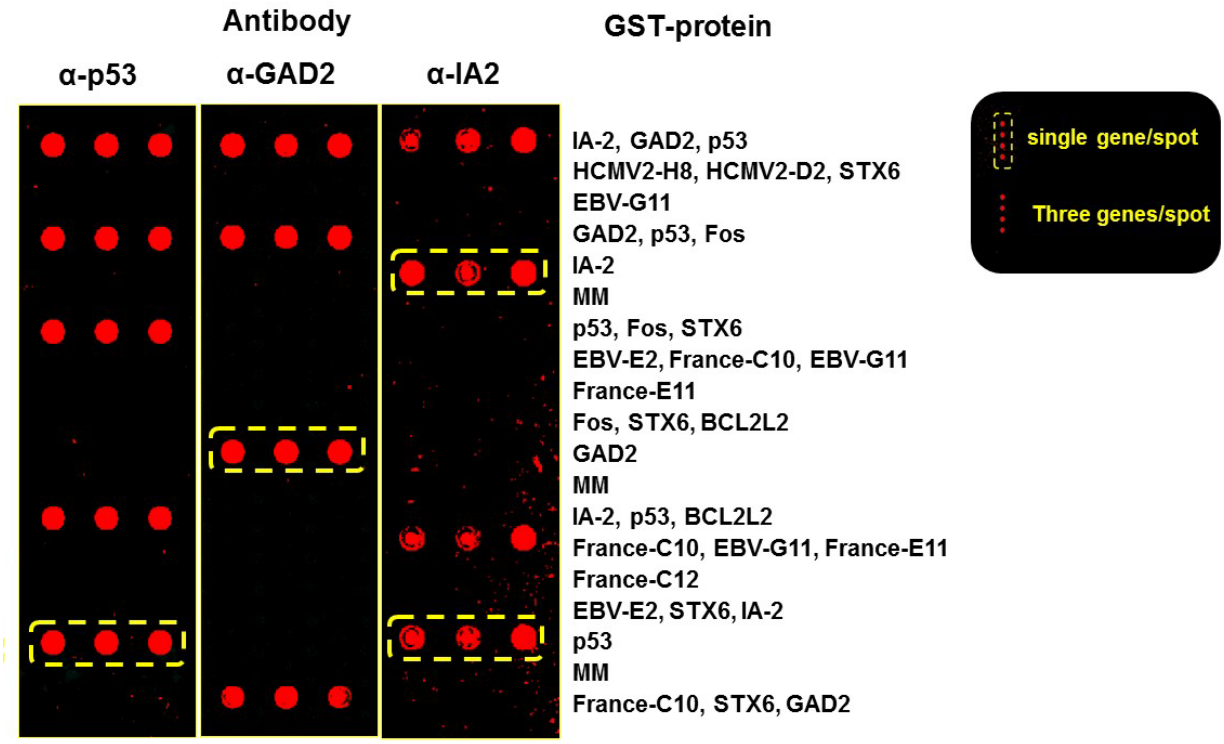


Figure S10. Multiplexed and non-multiplexed proteins displayed on HD-NAPPA were detected using their corresponding antibodies. Gene plasmids encoding for IA-2, GAD2, p53 and Fos were either mixed or spotted on the array individually.

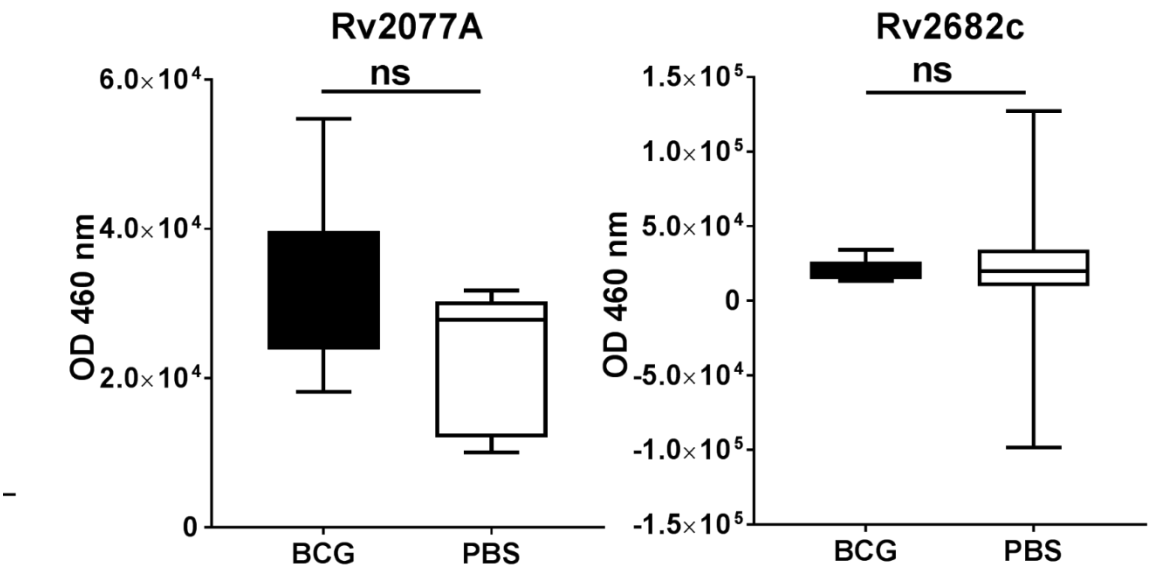
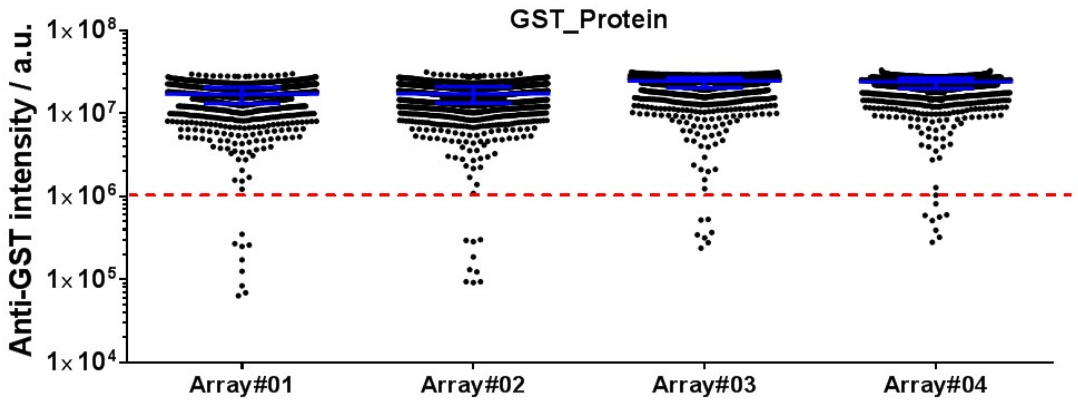


Figure S11. Detection of non-antigenic TB proteins using RAPID-ELISA in individual guinea pig sera samples. There is no significant difference in the serological antibody binding to Rv2077A or Rv2682c between guinea pigs vaccinated with BCG (n-4) or PBS (n=5). The p-value was calculated using an un-paired Man Whitney t-test. ns = nonsignifcance



Print time	0 month		6 month	
Display rate	99.41%	99.41%	99.22%	99.35%

Figure S12. Protein display levels on M-NAPPA arrays remain constant across six months. The M-NAPPA arrays can be stored for six months without a sacrifice in protein expression and display rate. The GST-protein display rate was calculated using the average signal intensity of “non-spots” plus two standard deviations using an anti-GST antibody.