**Antibody Profiling of Prostate Cancer Patients Reveals Differences in Antibody Signatures Between Disease Stages and Following Treatment**

Hemanth K. Potluri1, Tun Lee Ng2, Michael A. Newton2, Chris Maher3, Peter S. Nelson4, Douglas G. McNeel1,5

Running Title: Prostate cancer peptide microarray

1 University of Wisconsin Carbone Cancer Center, 1111 Highland Avenue, Madison, WI 53705

2 University of Wisconsin Department of Biostatistics and Medical Informatics, 1111 Highland Avenue, Madison, WI 53705

3 Washington University in St. Louis, St. Louis, MO

4 Fred Hutchinson Cancer Research Center, Seattle, WA

5 To whom correspondence should be addressed: 7007 Wisconsin Institutes for Medical Research, 1111 Highland Avenue, Madison, WI 53705. Tel: (608) 265-8131 Fax: (608) 265-0614 Email: [dm3@medicine.wisc.edu](mailto:dm3@medicine.wisc.edu)

Word Count: 4677

Abstract Word Count: 309

Keywords: Antibody, autoantibody, prostate cancer, peptide microarray, androgen deprivation, vaccine

List of Abbreviations:

**ABSTRACT**

**Background:** Previous studies of prostate cancer autoantibodies have largely focused on diagnostic applications. So far, there have been no reports attempting to more comprehensively profile the landscape of prostate cancer-associated antibodies. Specifically, it is unknown whether the quantity of antibody responses or the types of proteins recognized change with disease progression or treatment.

**Methods:** A peptide microarray spanning the amino acid sequences of the gene products of 1611 prostate cancer-associated genes was synthesized. Serum samples from healthy male volunteers (n=15) and prostate cancer patients (n=85) were used to probe the array. These samples included patients with various clinical stages of disease: newly diagnosed, castration-sensitive non-metastatic prostate cancer (nmCSPC), castration-resistant non-metastatic prostate cancer (nmCRPC), and castration-resistant metastatic disease (mCRPC). Serial sera samples from the individuals with nmCSPC, who received treatment with either standard androgen deprivation therapy (ADT) or a vaccine, were also used to probe the array.

**Results:** We demonstrated that this peptide array yields highly reproducible measurements of serum IgG levels. We found that the overall number of antibody responses does not increase with disease burden. However, the composition of antibody responses shifts with clinical stage of disease. Analysis revealed that nmCRPC patients recognized more proteins associated with chromosome organization and protein localization to the membrane than nmCSPC patients, while mCRPC patients recognized more proteins associated with nucleic acid binding and gene regulation than all other patients. Our longitudinal analysis showed that vaccine-treated patients developed increased responses to more proteins over the course of treatment than did ADT-treated patients, consistent with the development of antigen spread.

**Conclusions:** This study represents the largest survey of prostate-cancer associated antibodies to date. We have been able to better characterize the classes of proteins recognized by patients and how they change with disease burden. Our findings demonstrate the potential of this platform for measuring antigen spread and better studying response to immunomodulatory therapies.

**BACKGROUND**

It has been previously reported that patients with cancer develop antibodies to autologous proteins[1,2]. This phenomenon has been described across a wide variety of cancer types, including colon, melanoma, bladder, lung, and prostate[3–7]. These antibodies may arise due to overexpression of self antigens, inflammation, or tumor cell lysis[8]. Studies of serum antibodies may be particularly attractive for a variety of diagnostic applications because serum samples are relatively easy to obtain, antibodies can be present at early stages of disease, and antibodies can be present at high levels even when their target antigen is expressed at low levels. In contrast, monitoring serum proteins in cancer patients has been more challenging because they are often much less abundant and have more variable expression over time[9,10]. Antibody responses can also provide information about the relative immunogenicity of a given antigen. Many groups have used naturally existing antibody responses in cancer patients to identify targets for antibody therapies or vaccination strategies[11–13]. Another application of profiling antibody responses has been to detect antigen spread following immunotherapy[14]. Thus, further study of these antibody responses may have important implications for cancer diagnostics, biomarkers of response, and in guiding the design and targets of future therapies.

In the case of prostate cancer, several groups have developed methods to evaluate serum antibody responses[7,15,16]. Chinnaiyan and colleagues used phage display to screen patient serum for responses against many candidate prostate cancer-associated peptides, identifying 22 proteins against which antibody responses could distinguish prostate cancer patients and healthy individuals more reliably than detection of serum PSA protein. Taylor et al. and Ummanni et al. took similar approaches, probing prostate tumor lysates with patient serum and then performing mass spectrometry to identify the proteins that reacted more with cancer patient serum than control serum. Our group has also interrogated patient serum samples to discover prostate tumor-associated antibodies using ELISA for known prostate cancer tumor antigens and the serological identification of antigens by recombinant expression (SEREX) methodology to identify antibody targets from tissue expression libraries[17–19]. These previous studies of antibodies in prostate cancer patients focused primarily on diagnostic applications or on changes in responses. This approach has resulted in the discovery of small panels of shared antigens that may be useful for monitoring development of disease or response to treatment. However, to date, no studies have performed a more complete profile of the repertoires of prostate-cancer associated antibodies in individuals. In addition, data on whether the quantity or composition of antibody responses differ between patients with different disease severity are lacking.

Early studies were able to characterize antibody responses to small numbers of antigens, but advancements such as phage display and now microarray-based platforms have made it possible to develop more thorough profiles of antibodies in cancer patients. We sought to develop a microarray capable of detecting serum IgG responses against peptides using gene products from genes highly expressed in prostate cancer and predicted products of open reading frames (ORFs) from prostate cancer-associated lncRNAs. Our goal was to evaluate the number and character of proteins recognized by individuals with different clinical stages of disease, and whether a peptide microarray could be used to detect changes in antibody profiles following cancer treatment.

Here, we describe the use of the largest reported prostate cancer-specific peptide microarray. We demonstrate that the composition of antibody responses does change with stage, with the largest differences evident between patients with castration-resistant disease and castration-sensitive disease, but the overall number of proteins recognized by antibodies does not change with stage. We provide a detailed examination of the types of proteins that are recognized in patients with different clinical stages of prostate cancer and that have received treatment. Many more proteins have increased antibody recognition following treatment with vaccination than following androgen deprivation therapy, suggesting that the microarray platform could be used to measure prostate cancer-associated antigen spread.

**METHODS**

**Patient Populations**

Sera were previously collected from male volunteer blood donors without cancer (n=15, controls), or patients with prostate cancer (n=85). Sera from patients were grouped according to stage of disease (newly diagnosed (New Dx), castration-sensitive non-metastatic prostate cancer (nmCSPC), castration-resistant non-metastatic prostate cancer (nmCRPC), and castration-resistant metastatic disease (mCRPC)). Sera were also collected serially from the individuals with nmCSPC, who were enrolled on clinical trials in which they were treated with standard androgen deprivation therapy (ADT)[20] or an investigational vaccine[21]. All subjects gave IRB-approved written informed consent for their blood products to be used for immunological research. All samples were stored between -20° and -80°C until use for analysis.

**Antigen Selection**

Gene products from 1451 of the most highly expressed transcripts in prostate cancer and 160 predicted open reading frames from long non-coding RNAs that are highly expressed in prostate cancer were selected for inclusion on the array ([Supplemental Table 1) [refs]. Gene products included 125 antigens previously identified as recognized by IgG from patients with prostate cancer[22].

**Peptide Array Synthesis and Antibody Screening**

Peptide synthesis was performed as previously described, using a light-directed array synthesis in a Roche maskless array synthesizer.[23] Cycles of amino acid coupling were repeated until 16-mer peptides were synthesized on arrays containing 12 replicates of 177,604 peptides per subarray. Sera were diluted 1:100 with binding buffer (0.1M Tris, 1% alkali-soluble casein, 0.05% Tween-20), incubated overnight at 4°C, and washed. IgG was detected using an Alexa Fluor 647-labeled anti-human IgG secondary antibody (Jackson ImmunoResearch Labs). After final washing, arrays were dried and read using a Roche MS 200 microarray scanner, and signals were extracted using Roche internally developed software. Fluorescent signals were converted into arbitrary units (AU) with intensity plots ranging from 0 to 65,000 AU. All samples were evaluated in triplicate on separate arrays. Samples were considered positive for an antibody response at a given probe if the signal crossed 212 fluorescence units, with a sliding scale p value less than 0.05 in at least two of three technical replicates[24]. A binding buffer only control was also run to confirm the absence of signal above the 212 threshold. \*\*something about data produced via Roche protocol/normalization/binary calls vs quantitative scores\*\*

**Data Analysis**

Data analyses were performed in R version 3.6.2 [28] and RStudio [RStudio Team, 2020] using many extension packages and visualization tools available in these systems as well as custom scripts. To support reproducibility, we supply workflow details in a R markdown document and the rendered Statistical Supplement. Here is a link to the (currently private) github repository (better in the acknowledgements):

<https://github.com/wiscstatman/immunostat/tree/master/potluri/Project1>

*Array Reproducibility:* In a complementary analysis [Statistical Supplement, Section 2.3], we fit peptide-specific mixed-effects linear model to measure the relative size of technical variation to biological variation in this system. This utilized the R package lme4 [Bates et al. 2015] on log-transformed fluorescence intensity levels to compute variance components while adjusting for possible fixed effects of disease stage.

*Differences between clinical groups:* Fisher’s exact test was used to compare proportions of patients with antibody responses between groups \*\*Hemanth, is that right?\*\*. For protein recognition analysis, ANOVA with the Tukey’s Honest Significant Differences post-test was used to compare the numbers of proteins recognized among patients with different clinical stages. We also performed peptide-specific logistic regression testing for cancer-stage effects while controlling the false discovery rate (FDR) using the Benjamini-Hochberg (BH) method [Statistical Supplement, Section 2.4].

We reasoned that detectable antibody signatures between clinical groups may be present below the threshold of the stringent binary recognition calls. To test for such signals in the fluorescence intensity data, we applied peptide-specific analysis of variance according to the rank-based Kruskal Wallace (KW) procedure, first filtering peptides with significant clinical-group effects at 5% FDR by the BH method. Subject data were preprocessed to collapse replicate (usually triplicate) profiles per person to a single, consensus profile per person by using median per peptide [Statistical Supplement Section 2.5]. This rank-based procedure is robust to distributional anomalies and is expected to provide a conservative assessment of antibody-profile differences between the clinical groups [McDonald, 2014]. Peptides exhibiting sufficiently small KW p-value were examined for differences in various pairwise comparisons, which invoked both a median fold-change filter (at least two-fold difference) as well as a significance filter by two-sample Wilcoxon rank sum p-value, again with BH adjustment at 5% FDR [Statistical Supplement, Section 2.6].

*Temporal changes:* A mixed linear model was fit to each peptide to determine if there was an increase in signal over time, again using lme4; this allows a linear increase or decrease in mean log-transformed intensity over time per subject and per peptide. Patient-specific random effects allow for among subject variation in the temporal response, while a fixed time effect per peptide expresses the average response over subjects in that clinical group. Statistical significance was assessed using both the Kenward Roger (KR) and Satterthwaite approximate F tests as well as BH for FDR control [Statistical Supplement, Section 3]. \*\*Tun, please check\*\* Peptides with a coefficient of at least 0.33 and a p value less than 0.05 were considered to have increased antibody response over time. \*\*

: and also *allez* [Newton et al. 2007].  [for GOrilla] and a Bonferonni-corrected p-value threshold of 0.05 in *allez.* , or , and using waterfall plots in *allez*. These reveal dominant functional categories enriched in the protein list while accounting for set redundancies [e.g., Hao et al. 2013, Pleiman, 2018].

## extra citations:

Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.

*RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL <http://www.rstudio.com/>.*

Newton, M.A., Quintana, F.A., den Boon, J.A., Sengupta, S. and Ahlquist, P. (2007). Random-set methods identify distinct aspects of the enrichment signal in gene-set analysis. *Annals of Applied Statistics*, 1, 85-106. <https://github.com/wiscstatman/allez> [version 2.0.7]

Hao, L, He, Q, Wang, Z, Craven, M, Newton, MA, and Ahlquist, P (2013). [Limited agreement of independent RNAi screens for virus-required host genes owes more to false-negative than false-positive factors.](http://www.ploscompbiol.org/article/authors/info%3Adoi%2F10.1371%2Fjournal.pcbi.1003235)*PLoS Comput Biol*9(9): e1003235.

Pleiman, J.K., Irving, A.A., Wang, Z., Toraason, E., Clipson, L., Dove, W.F., Deming, D.A. and Newton, M.A., 2018. [The conserved protective cyclic AMP-phosphodiesterase function PDE4B is expressed in the adenoma and adjacent normal colonic epithelium of mammals and silenced in colorectal cancer.](https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1007611) *PLoS Genetics,* 14(9), p.e1007611.

McDonald, J.H. 2014. Handbook of Biological Statistics (3rd ed.). Sparky House Publishing, Baltimore, Maryland. (pages 157-164 on Kruskal Wallis)

**UniProt Analysis**

Proteins from the array were matched with UniProt IDs using UniGene IDs when available and protein names otherwise. Data were then retrieved from UniProt[27] on gene names, protein length, and subcellular location. UniProt may designate a protein with multiple subcellular localizations, in which case all localizations were kept in the analysis. This sometimes leads to percentages that add up to over 100%.

**Statistical Analysis**

**RESULTS**

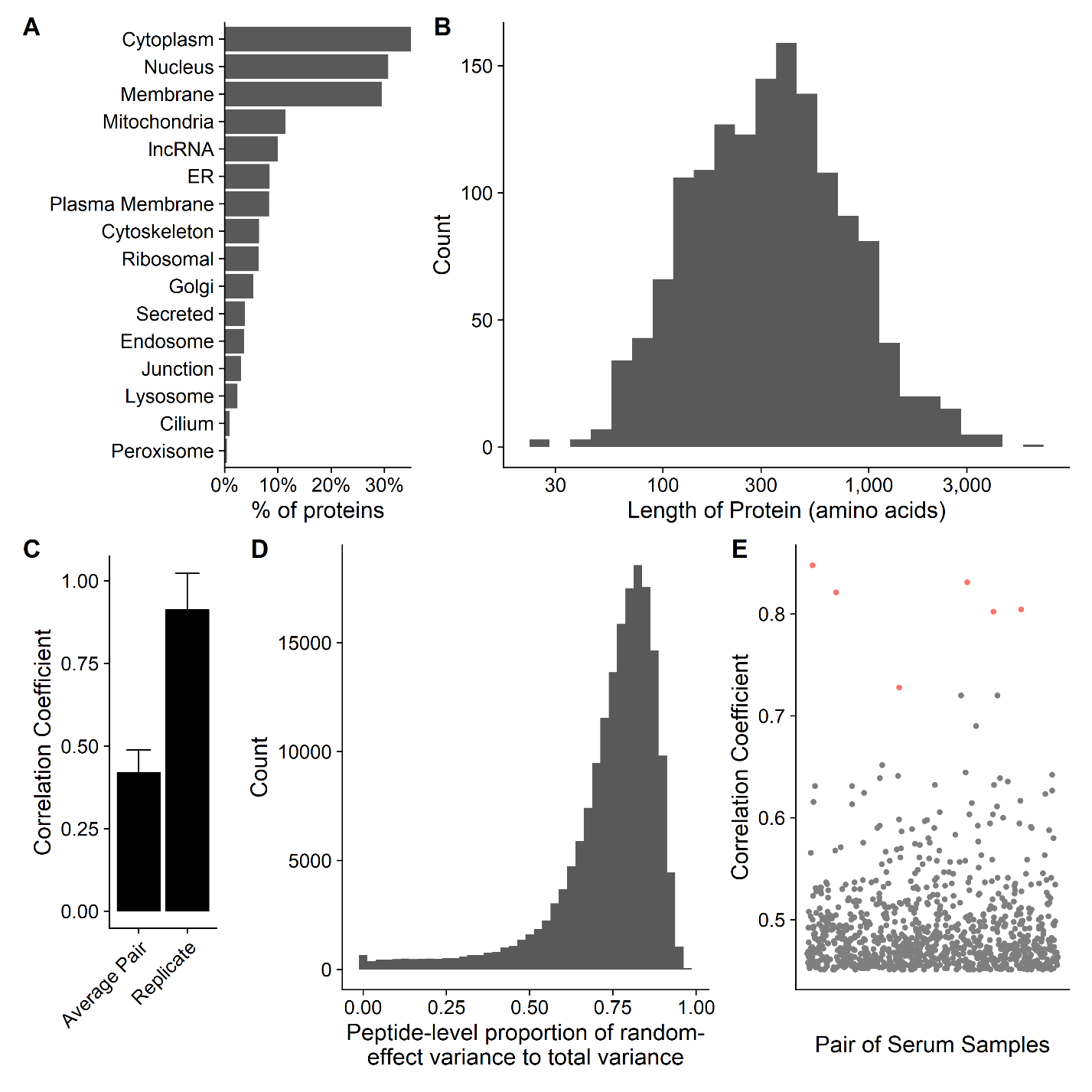
**A prostate cancer-specific peptide microarray was able to reproducibly measure antibody signatures from serum of healthy individuals and prostate cancer patients**

To characterize antibody responses to a wide variety of proteins in prostate cancer patients, we designed a peptide microarray able to be screened with patient sera. This array included peptides spanning the amino acid sequences of 1451 of the most abundantly expressed gene products in metastatic prostate cancer [citation], including 125 proteins identified in previous studies examining serum antibody responses in prostate cancer patients[22,29]. We also included a set of peptides spanning the predicted amino acid sequences of 160 potential open reading frames (ORFs) from long non-coding RNAs (lncRNAs) that have been shown to be highly expressed in prostate cancer[citation]. We included these given their strong association with prostate cancer. While most would likely serve as negative controls as they would not be expected to encode gene products, other groups have shown that some lncRNAs may be translated into unstable peptides or even functional proteins, especially with the dysregulation induced by cancer [30–32]. Hence we reasoned that a few might serve as antibody targets in prostate cancer patients.

16-mer peptides spanning the amino acid sequences of these 1611 gene products, and overlapping by 12 amino acids, were used to generate a microarray comprising 177,604 peptides. The complete list of probes and corresponding proteins is available in Supplementary Table 1. The manufacture of the array and synthesis of peptides was performed as previously described[33]. The characteristics of the proteins included in the array are summarized in Figure 1, using data retrieved from UniProt[27]. The majority of proteins included were those typically localized within the cytoplasm or nucleus, or that traffic between the two compartments (Fig 1A). Approximately 16% of the proteins were localized to the mitochondria or ribosomes. The median protein length was 483 amino acids (Fig 1B).

We obtained serum samples from 15 healthy male blood donors (Control), 15 patients with newly diagnosed prostate cancer (New Dx), 40 patients with castration-sensitive non-metastatic prostate cancer (nmCSPC), 15 patients with castration-resistant non-metastatic prostate cancer (nmCRPC), and 15 patients with castration-resistant metastatic disease (mCRPC). Each patient’s serum was assayed in triplicate for peptide-specific IgG responses using the microarray. Examples of the primary data are shown in Supplementary Fig 1 A and B. To assess the reproducibility of the assay, we calculated Pearson correlation coefficents between each pair of technical replicates and found high correlation on average among replicates (Fig 1C). To determine the degree of variability among serum samples, we calculated the mean correlation coefficient across all pairs of distinct serum samples. We observed low correlation between the average pair of serum samples (Fig 1C). In a complementary approach, we fit a linear mixed-effects model to estimate the amount of biological variation and technical variation across our triplicate data for each peptide. We found that in general, biological variation was far greater than the technical variation with an average ratio of biological variation to technical variation of 0.74 (Fig 1D).

Included in this study were 6 patients who had serum collected at two different time points, when they had an early stage of disease and again when they had a later stage of disease. Notably, these serum samples from the same patients had especially high correlation coefficients (Fig 1E). This suggests that while there is high variation among individuals, each particular individual has smaller variation in his antibody repertoire over time. These six patients had their first serum collection removed from further analysis to prevent inflating their impact on our results.

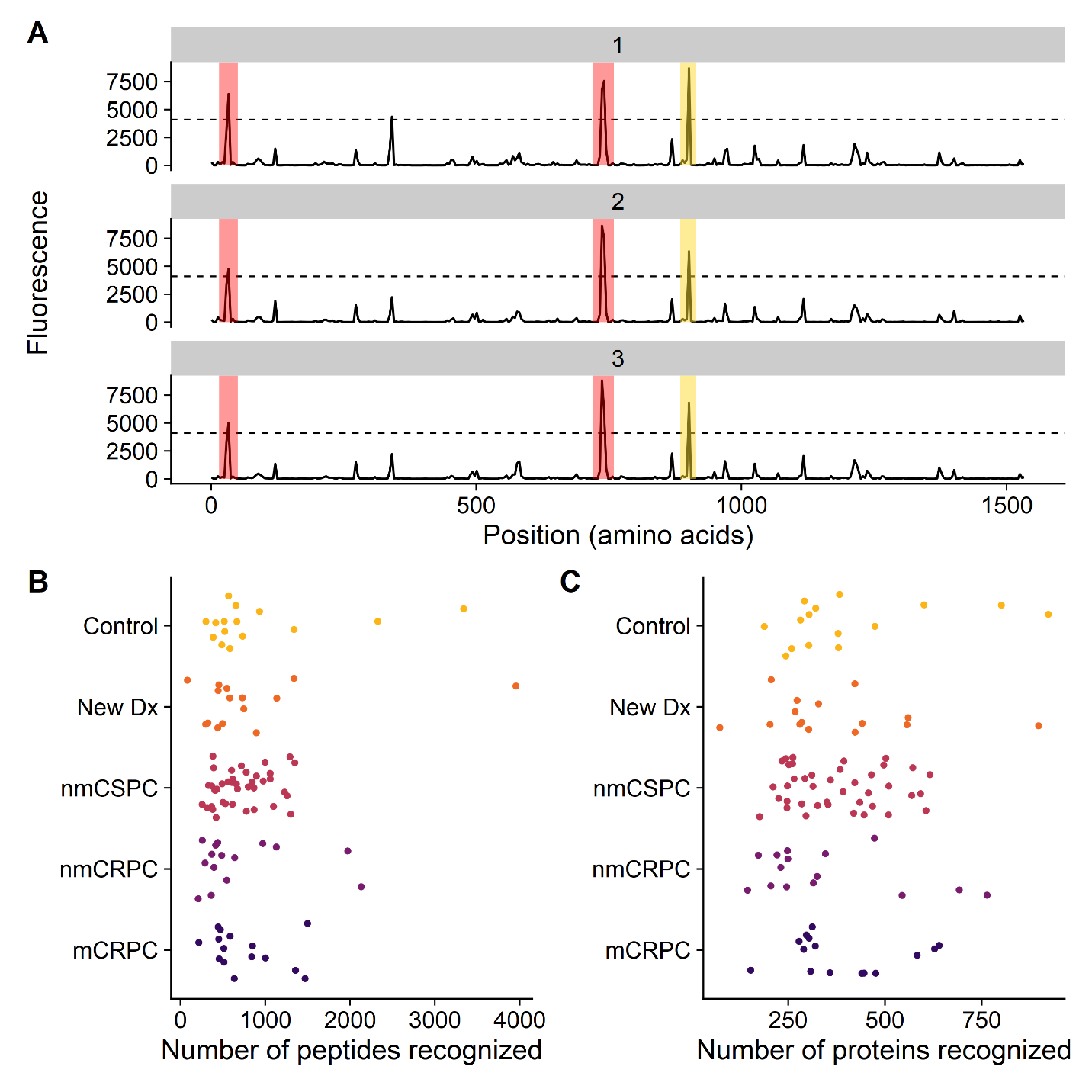


**Figure 1.** A prostate cancer-specific peptide microarray was able to reproducibly measure antibody signatures from serum of healthy individuals and prostate cancer patients. Summary of the **a)** subcellular localization and **b)** length in amino acids of all 1611 unique proteins on the array according to UniProt. **c)** The mean correlation coefficient among all pairs of different individuals (Average Pair) compared to the average correlation coefficient among all technical replicates (Replicate). Error bars represent standard deviation. **d)** Histogram depicting the ratio of the biological variation to the technical variation of the array data for each peptide as estimated by a linear mixed-effects model. **e)** Each point represents the correlation coefficient between antibody responses in two different serum samples. Points marked in red are instances when the same individual had serum collected at two different time points with different stages of disease.

To determine whether the array could detect IgG to common prostate antigens, we first defined a “positive” antibody response to individual peptides. We considered probes with fluorescence intensity of at least 212, and a sliding window p value less than 0.05 (indicating high signal in adjacent peptides), in at least 2 of the 3 technical replicates to be positive, as previously described[24]. Using binding buffer as a negative control, no peptides met these criteria (not shown). Two examples of positive responses are shown in Fig 2A. We specifically evaluated responses to peptides derived from prostate specific antigen (PSA), prostatic acid phosphatase (PAP), and the androgen receptor (AR), well defined prostate target antigens. 7.1% of prostate cancer patients (13.3% of patients with mCRPC) assayed on the array displayed antibody responses against peptides derived from PSA, while 6.7% of controls had PSA responses. 8.2% of prostate cancer patients (13.3% of patients with mCRPC) and 0% of controls had responses to PAP. Finally, 5.9% of prostate cancer patients (13.3% of patients with mCRPC) and 20.0% of controls had responses to peptides derived from the ligand-binding domain of AR. None of the antibody responses to these proteins were found significantly more often in cancer patients than controls by Fisher’s Exact Test.

**Frequency of protein recognition did not correlate with stage of disease**

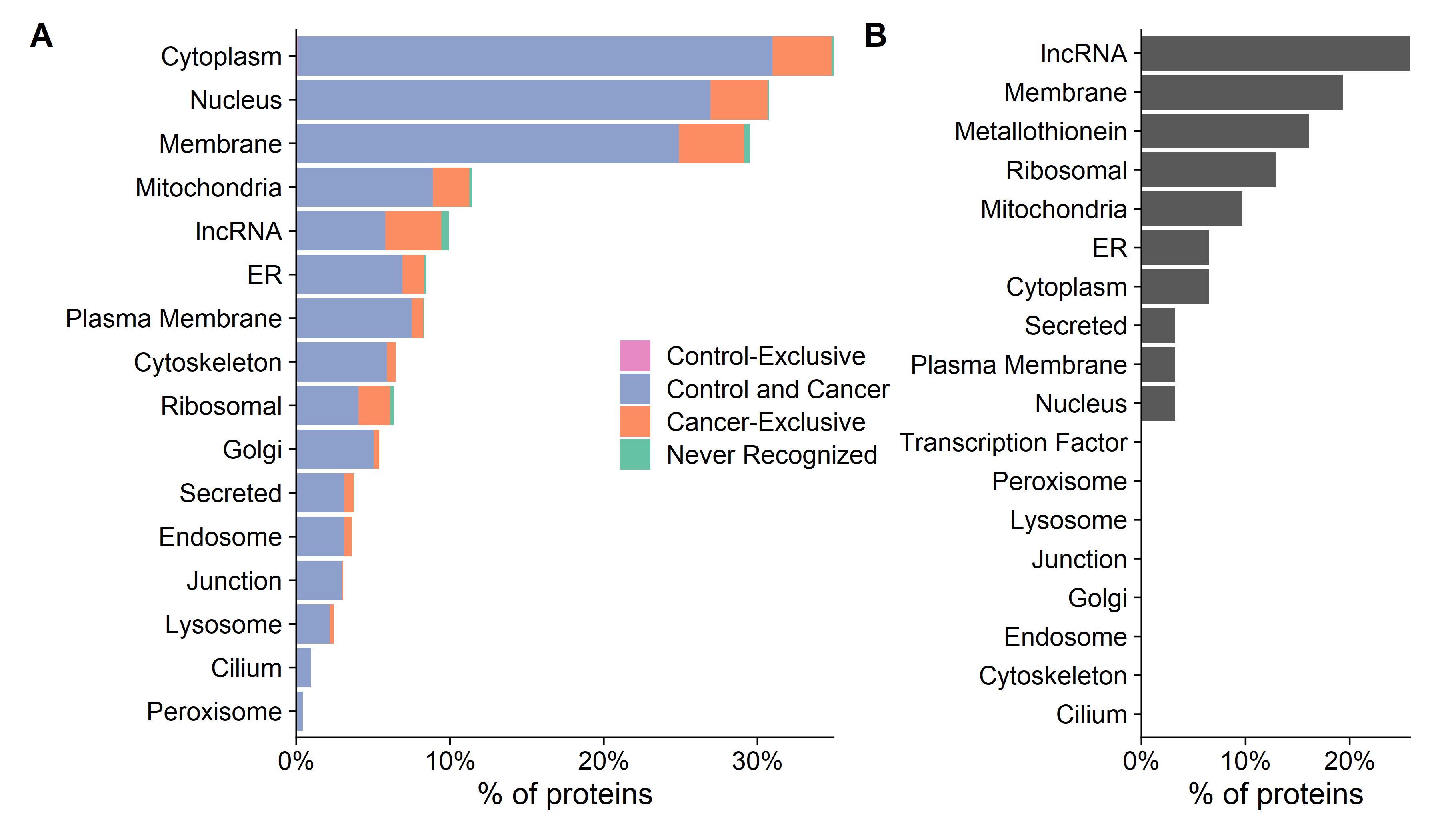
We next tested the hypothesis that patients with higher disease burden would recognize more peptides, potentially due to increased presentation of cancer-associated peptides[34]. We found no correlation between stage of disease and the number of probes recognized at either the peptide level or the protein level. The median number of proteins recognized were 321 for controls, 303 for New Dx, 353 for nmCSPC, 249 for nmCRPC, and 320 for mCRPC (Fig 2B, 2C). The median numbers of peptides were 919 for controls, 832 for New Dx, 712 for nmCSPC, 708 for nmCRPC, and 754 for mCRPC. We noted a substantial amount of heterogeneity in antibody responses among patients. For instance, the number of proteins recognized by controls ranged from 188 to 922.



**Figure 2**. Frequency of protein recognition did not correlate with stage of disease. **a)** Example microarray data for technical replicates of a single protein (ADT14) with the 212 signal threshold indicated by the dashed line. Positive calls are marked in red. In yellow is a negative call that did not meet the sliding window criterion. The number of **b)** peptides and **c)** proteins recognized by each patient, categorized by clinical stage of disease.

**Nearly all proteins on the array were recognized by prostate cancer patients**

Having established that there was a large diversity in antibody responses among patients, we next examined whether there were any broad trends in the types of proteins that were recognized. While only 0.4% of calls were positive overall, 20% of peptides were recognized by at least one subject. Nearly all proteins (1570 of 1611, 97%) had one or more peptides recognized by at least one subject. Conversely, there were no proteins that were recognized by all patients. Most proteins (1326 of 1611, 82%) were recognized by both controls and cancer patients (Fig 3A). As expected, the largest category of proteins that were never recognized were ORFs from lncRNAs (Fig 3B); however, contrary to our expectations, the majority of lncRNAs (145 of 160, 91%) were recognized by at least one patient (Fig 3A).

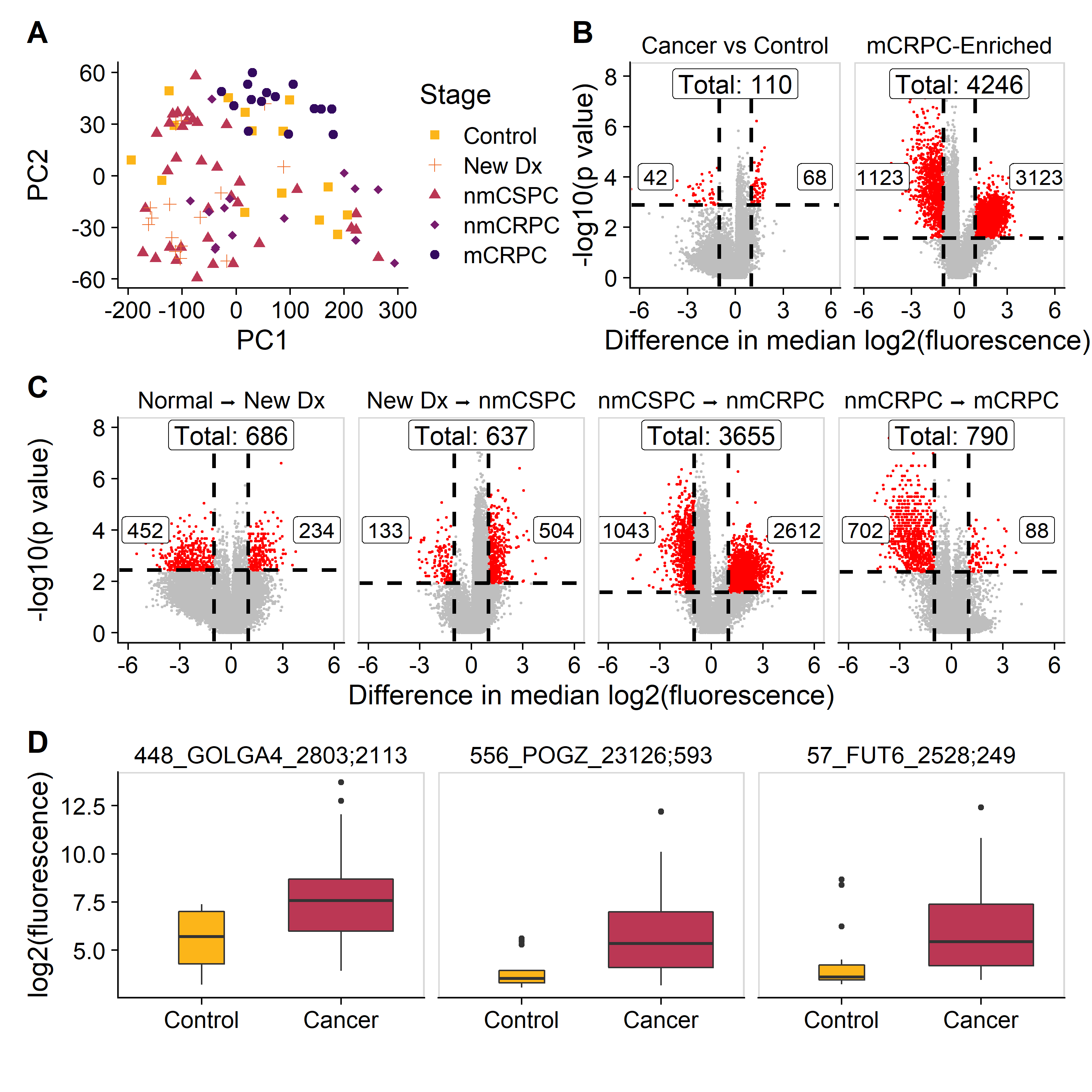
****

**Figure 3.** Nearly all proteins on the array were recognized by prostate cancer patients. **a)** Percentage of proteins that were recognized by only controls (*Control-Exclusive*), percentage of proteins recognized by at least one control and one cancer patient (*Control and Cancer*), percentage of proteins not recognized by any controls but recognized by at least one cancer patient (*Cancer-Exclusive*), and percentage not recognized at all (*Not Recognized*), categorized by subcellular localization. **b)** Characteristics of proteins that were never recognized by any controls or patients tested.

**The composition of patient antibody responses changes with clinical stage of disease**

We hypothesized that while the overall number of proteins recognized may not increase with burden of disease, the composition of proteins recognized may be different. We used a Kruskal-Wallis test to identify peptides that had significantly different fluorescence intensities across clinical stages and controls while controlling the false discovery rate (FDR) at 5% using the Benjamini-Hochberg (BH) method. This test identified 13279 significant peptides. We used PCA to visualize the residual fluorescence levels after subtracting the grand mean fluorescence level for each peptide and observed that patients tended to group with other patients with the same clinical stage of disease (Fig 4A). Patients with castration resistant tumors, and metastatic CRPC in particular, tended to cluster especially closely to one another. Notably, the controls did not exhibit this clustering. We were particularly interested in the subset of peptides that had significantly different fluorescence signals in cancer patients compared to controls. We identified these peptides by using a Wilcoxon-Rank-Sum test with a BH FDR of 5% and specifically focused on those that had differences in median fluorescence of at least two-fold in cancer patients compared to controls (Fig 5B, left; Supplementary Table 2). To discover which peptides were driving the especially strong clustering of mCRPC patients, we repeated this procedure to find peptides with significantly different fluorescence in mCRPC patients compared to all other patients (Fig 5B, right; Supplementary Table 3).

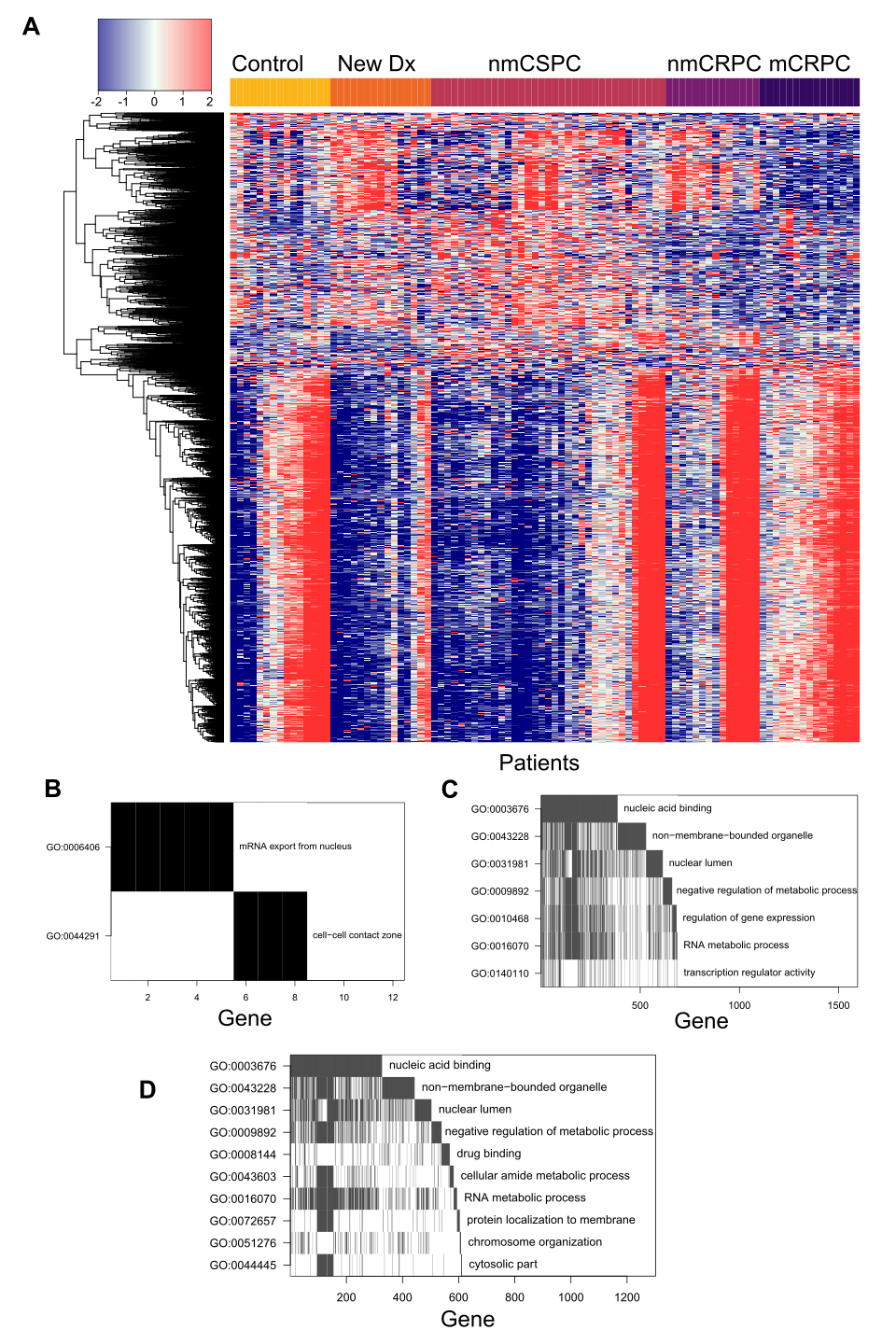
We applied this same approach to identify the number of peptides that had significantly higher or lower responses in patients in one clinical stage of disease compared to patients in the previous clinical stage. The largest change in number of recognized peptides occurred between the castration-sensitive (nmCSPC) and castration-resistant (nmCRPC) populations (Fig 4C; Supplementary Tables 4-7). Examples of the fluorescence signals of peptides that are detected by this strategy are shown in Fig 4D.



**Figure 4.** The composition of patient antibody responses changes with clinical stage of disease. **a)** PCA plot obtained by using the set of 13279 significantly changed peptides identified by the Kruskal-Wallis test then subtracting the grand mean of log2 fluorescence levels across patients for each peptide. Each point represents a patient, colored by clinical stage. **b)** Volcano plots depicting peptides that met the 5% BH FDR cutoff based on the Wilcoxon p values (horizontal lines) and had at least a two-fold difference in median log2 fluorescence values between the stages being compared (vertical lines). The number of significantly increased peptides is shown on the right of each plot, the number of significantly decreased peptides is shown on the left, and the overall number of significantly changed peptides is shown at the top. Significant peptides are colored red. The left plot indicates peptides that had significantly different signals in cancer patients compared to controls. The right plot indicates peptides that significantly different signals in mCRPC patients compared to all other groups. **c)** Volcano plots indicating peptides that had significantly different signals between patients with consecutive clinical stages of disease. **d)** Boxplots displaying fluorescence signals in cancer patients compared to controls in 3 example peptides that met both the two-fold signal change and p value criteria.

**Specific proteins are preferentially recognized in cancer patients and patients with mCRPC**

Between the six analyses shown in Fig 4B and 4C, we narrowed our list of peptides of interest to 6708. We visualized the residuals of these peptides in Fig 5A. As in Fig 4A, we observed high similarity in antibody responses between patients with the same stage of disease. We next more closely examined the sets of proteins we had identified earlier for common features and associations with cellular processes. Gene ontology analysis revealed that the genes corresponding to the 68 peptides that were recognized more heavily in cancer patients compared to controls were associated with mRNA export from the nucleus and the cell-cell contact zone (Fig 5B). GO analysis of the 3123 peptides that had particularly strong antibody responses in mCRPC patients showed an enrichment for proteins associated with nucleic acid binding, RNA metabolism, gene regulation, and down-regulation of metabolism (Fig 5C). One of the significant terms within the “non-membrane-bounded organelle” term was the cytosolic large ribosomal subunit. To investigate the large difference in antibody repertoires between patients with nmCSPC and nmCRPC, we performed GO analysis on the 2612 peptides with significantly higher signal in nmCRPC than nmCSPC. We identified differences in antibody responses to proteins associated with chromatin structure, amide metabolism, and protein localization to the membrane (Fig 5D).

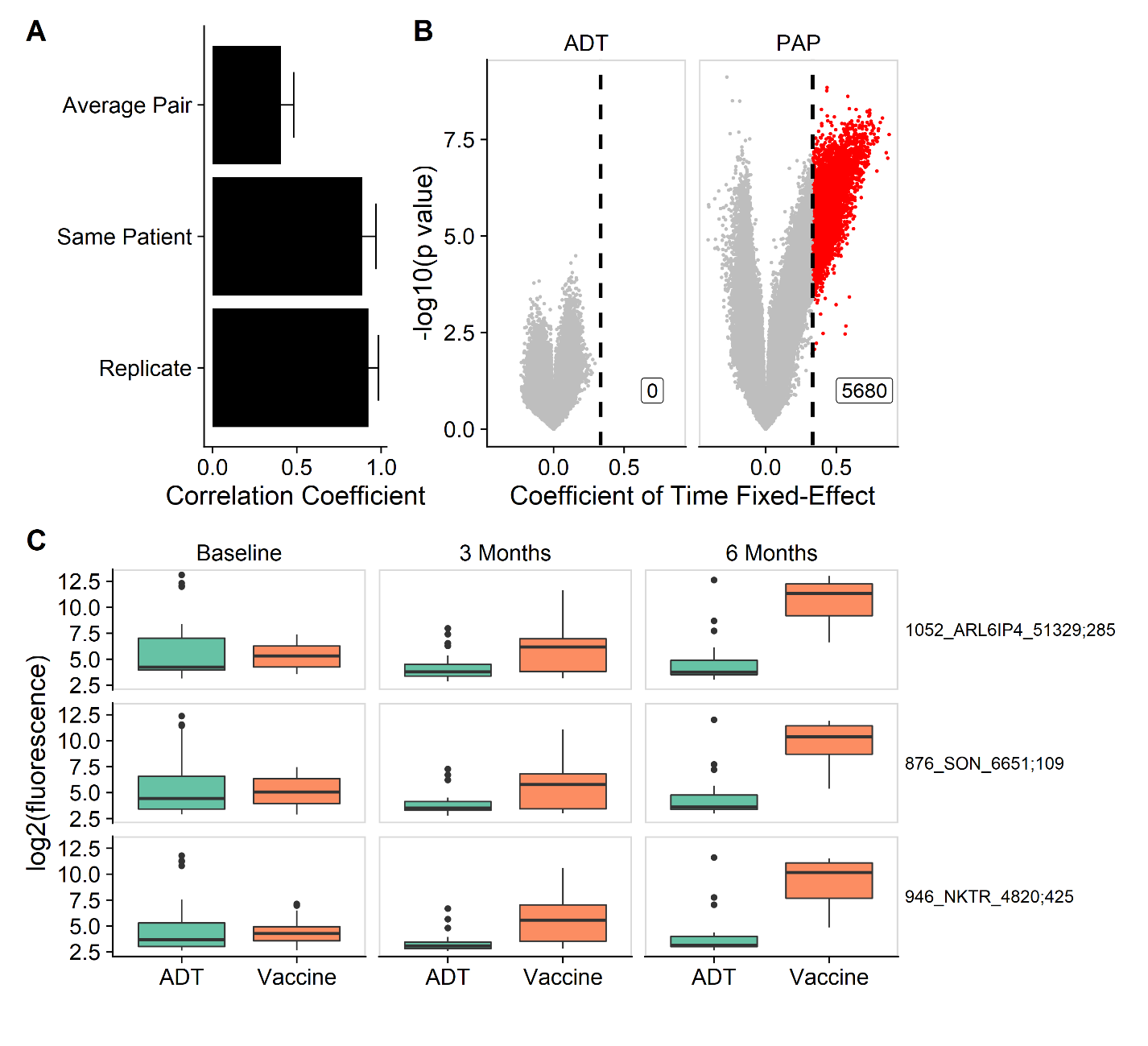


**Figure 5.** Specific proteins are preferentially recognized in cancer patients and patients with mCRPC **a)** Heatmap depicting the difference in log2 fluorescence levels between each peptide in each patient and its grand mean across patients, displaying only the set of 6708 peptides that met the secondary cutoff.[selection criteria] Patients are grouped by stage across the x-axis, while peptides are clustered along the y-axis. **b)** Waterfall plot depicting a gene ontology analysis of proteins that had significantly more antibody responses in cancer patients than controls. The top row indicates the GO term that encompasses the most genes corresponding to significant peptides. For the second row, these genes are then removed from the list and the GO term that encompasses the most genes in the remainder of the list is chosen. Genes identified by this process are counted along the x-axis to visualize overlapping GO terms. Waterfall plots depicting gene ontology analysis of proteins that had significantly increased antibody responses in **c)** mCRPC patients compared to all other patients or **d)** nmCRPC patients compared to nmCSPC patients.

**Vaccination elicits increased antibody responses over time, unlike androgen deprivation therapy**

Based on our finding that individual patients tend to have relatively small variation in their antibody responses over time, we hypothesized that this could make the microarray particularly sensitive for detecting changes induced by treatment in a longitudinal analysis. To address this, we used serum samples available from the 40 patients with nmCSPC who were treated with one of two therapies. 20 patients received standard androgen deprivation therapy (ADT; GnRh analogue given every 3 months)[20], and 20 patients received a DNA vaccine encoding prostatic-acid phosphatase (PAP; pTVG-HP given every 14 days for 6 administrations)[21]. Blood samples were collected from these patients immediately prior to treatment, and at 3 months and 6 months following initiation of treatment. Consistent with our observations in Fig 1E, we found high correlation between samples from an individual patient over time (Fig 6A).

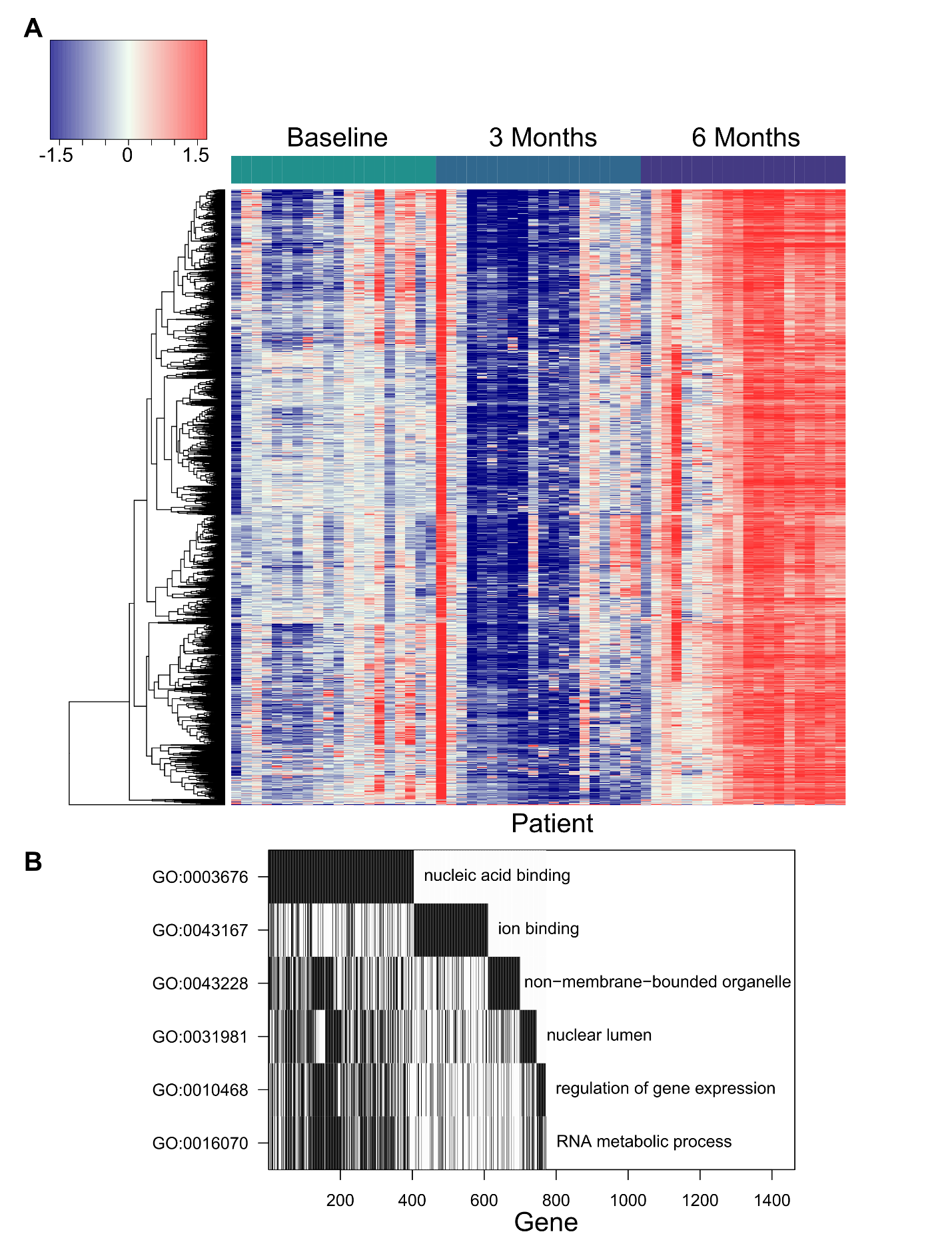
We next fit a linear model to [mixed effects] determine if there were any peptides which had increased signal over time [controlling FDR]. In the vaccine-treated patients, we found 5680 significant peptides that had a coefficient of time fixed-effect of at least 0.3333, indicating a two-fold increase in signal every three months (Supplementary Table 8). We were unable to detect any peptides that ADT patients developed increasing response to over time using this procedure (Fig 6B). Examples of the fluorescence levels of 3 peptides over time in ADT-treated and vaccine-treated are shown in Fig 6C.

****

**Figure 6.** Vaccination elicits increased antibody responses over time, unlike androgen deprivation therapy. **a)** The mean correlation coefficient among all pairs of different individuals (*Average Pair*) compared to the average correlation coefficient among all technical replicates (*Replicate*) and the average correlation among samples collected from the same patient at different time points (*Same patient*). Error bars represent standard deviation. **b)** Volcano plots depicting peptides that increased in signal following treatment with ADT or vaccine by at least 2-fold every three months, corresponding to a coefficient of time fixed-effect of 0.333 (vertical line), and met the 5% BH FDR cut off using both KR and Satterthwaite methods. Significant peptides are colored red. **c)** Example boxplots displaying log2 fluorescence levels for 3 peptides at baseline, 3 months, and 6 months, in patients treated with ADT or vaccine.

**PAP vaccination causes consistent increases in antibodies against proteins associated with nucleic acid binding and gene regulation across patients**

We visualized the changes in peptide recognition over time in vaccine-treated patients by plotting the residuals of the null model in the heatmap in Fig 7A. This further demonstrates that vaccine-treated patients had robust increases in antibody responses to these 5680 peptides. To characterize these peptides, we performed GO analysis. We found that a significantly enriched set of these antibodies were specific to nucleic acid binding proteins. There were also more antibodies against proteins associated with RNA metabolism, ion binding, and ribosomal or nucleolar cellular components than would be expected by chance (Fig 7B).

****

**Figure 7.** PAP vaccination causes consistent increases in antibodies against proteins associated with nucleic acid binding and gene regulation across patients. **a)** Heatmap of the fluorescence residuals from the null model for each of the 5680 peptides that were significantly increased in vaccine-treated patients. Samples from vaccine-treated patients at each collection time point (baseline, 3 months, and 6 months) are grouped together along the x-axis, while peptides are clustered along the y-axis. **b)** Waterfall plot of GO analysis of proteins recognized more following vaccine.

**DISCUSSION**

The purpose of this study was to perform a comprehensive survey of serum antibody responses in prostate cancer patients and to determine whether responses changed with disease progression or treatment. Previous examinations of serum antibodies in patients with prostate cancer focused mainly on diagnostic applications; thus, a more complete picture of patient antibody repertoires has been lacking. We addressed this by designing the largest reported prostate-cancer specific peptide microarray, capable of measuring IgG responses to over 177,000 peptides. Our major findings were 1) the microarray data are highly reproducible, 2) the overall number of peptides recognized is not greater in patients with more advanced disease, 3) the composition of patient antibody repertoires change with disease progression, 4) each individual has their own antibody signature, and 5) this approach can be used to track changes in individuals elicited by therapy.

Here we have shown that this novel prostate cancer-specific peptide microarray yields highly robust and reproducible measurements of serum IgG levels. We found that technical replicates were highly correlated and background fluorescence signal was negligible. The microarray’s measurements also exhibit generally strong concordance with existing literature on serum antibodies in prostate cancer patients. A previous study using ELISA detected anti-PSA antibodies in 11% of mCRPC patients[35]. Similarly, the microarray detected PSA responses in 13.3% of mCRPC patients. Looking at PAP, ELISA detected antibody responses in 5.5% of patients, while the microarray detected antibody responses in 8.2% of prostate cancer patients. On the other hand, ELISA detected antibodies specific for the AR ligand-binding domain in 17.1%, whereas the microarray detected antibody responses in 5.9% of patients[36].

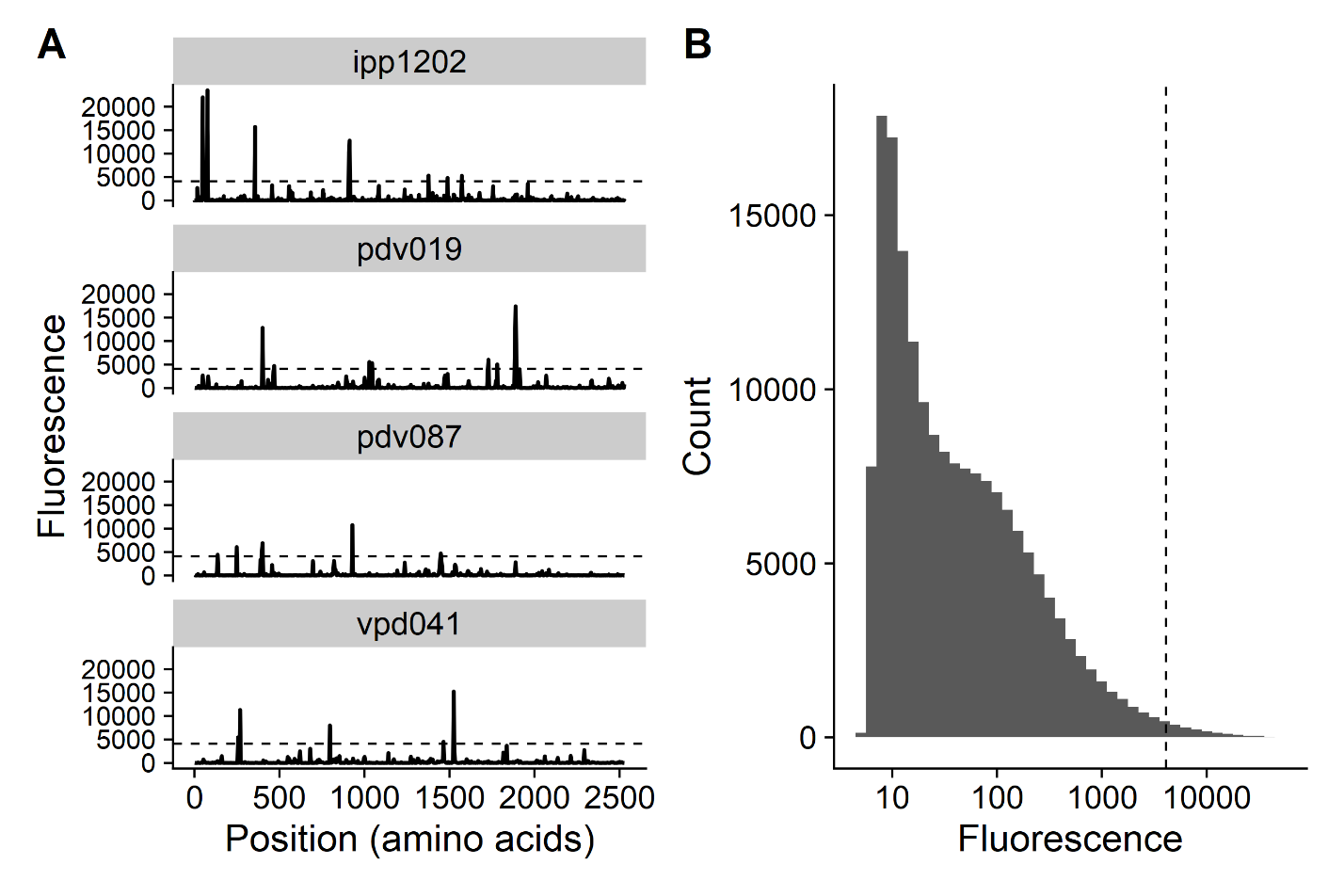
Based on reports that individual proteins like PSA and PAP are more recognized in patients than controls, we hypothesized that patients with more advanced disease would have antibodies against more proteins. Previous studies have focused on the use of antibody profiling as possible diagnostic tool to identify prostate cancer. They have used phage display or whole tissue fractionation followed by mass spectrometry to specifically discover proteins that are recognized more in prostate cancer patients than controls[15,16,37]. Because previous studies focused only on antibodies that are enriched in prostate cancer patients, they were unable to address this question of whether the overall size of the antibody repertoire changes with clinical stage of disease. One of the additional strengths of our microarray approach was that it also allowed us to examine the classes of proteins recognized by patients in each clinical stage.

Contrary to our expectations, we did not observe an increase in the number of peptides recognized with more advanced disease. While the overall number of antibody responses did not appear to increase, we found that the composition of proteins recognized appears to change. Interestingly, we discovered that the vast majority of predicted lncRNA ORF gene products were recognized by at least one subject, with a large proportion recognized exclusively in cancer patients. This could be the result of unstable peptides being translated from lncRNAs at higher rates due to the dysregulation induced by prostate cancer. We also found that the cancer-exclusive proteins were significantly enriched for ribosomal proteins. It is possible that the upregulation of the translational machinery required to support rapid cell division in cancer leads to a greater abundance in ribosomal proteins. This lends further credence to observations made by Wang et al. that 2 of the 5 coding proteins they identified in their screen for prostate cancer-specific antibodies were ribosomal and the majority of the other proteins they identified came from untranslated regions[37]. We demonstrated that mCRPC patients recognize more lncRNAs, ribosomal proteins, mitochondrial proteins, and proteins involved in nuclear division than other patients and controls.

Despite the power of this approach, we were limited to observing antibody responses to 1611 proteins that are all highly expressed in prostate cancer and it is possible that there are humoral responses to other targets that may be expressed at lower levels that we did not capture. In addition, only 15 of our 100 samples were from subjects without cancer, making us more likely to detect antibody responses that were specific to cancer patients than to controls. Due to practical concerns, we took a cross-sectional approach to identifying antibody responses associated with each stage of disease rather than following individual patients across the many year natural history of prostate cancer, which prevented us from observing changes in individual patients. However, we were able to obtain longitudinal data from a subset of patients for a period of six months.

Our longitudinal analysis of 40 patients with sample collections at baseline, 3 months, and 6 months revealed that each individual appears to have their own unique antibody signature or fingerprint which remains stable over time. Others have demonstrated this phenomenon, observing that healthy individuals have largely unchanged responses over time to a panel of self antigens[38,39], though we are the first to our knowledge to observe it with an array of this size and the first to study it in the setting of prostate cancer. Despite this individual signature, we did observe common recognized proteins among patients with the same clinical stage of disease. Due to the lack of large random fluctuations in antibody responses over time, this platform appears particularly suited to identify changes in individuals over time induced by disease or treatment.

Most strikingly, we have shown that treatments can modulate a patient’s antibody repertoire, at least during a 6-month study period. We found that antigen-specific vaccination elicited greater increases in off-target antibody responses over time than did traditional targeted therapy, showing that this may be a method of quantifying antigen spread caused by treatment. Our data are consistent with previous findings examining off-target antibody responses to Sipuleucel-T, but we were able to additionally study a greater number of prostate cancer-associated peptides and compare the effects of immunotherapy to the effects of androgen deprivation therapy[14]. These specific proteins to which patients receiving the PAP vaccine developed increasing responses may be useful as biomarkers of response to therapy. We also found that the proteins with increasing responses in ADT-treated patients are the [same/different] as those that were significantly higher in nmCRPC than in nmCSPC. This indicates that ADT is likely driving the dramatic shift in antibody profiles in patients with these two stages of disease. Future applications may focus on using this platform to identify antibody signature changes that are specific for various types of immunotherapies, quantitating antigen spread and immunomodulation induced by therapy, and predicting response to prostate cancer immunotherapy.



**Supplementary Figure 1.** Fluorescence signals vary between individuals and across peptides. **a)** Example traces from 5 patients of the same protein (ANKHD1). Dashed line indicates 2^12 threshold. **b)** Histogram of the distribution of fluorescence signals for all 177,604 peptides for patient ADT115. Dashed line represents 2^12 threshold.

**DECLARATIONS**

**Ethics approval and consent to participate:** Study protocols that permitted collection and use of human blood samples were reviewed and approved the University of Wisconsin Human Subjects’ Review Board (IRB). All patients gave written informed consent for use of blood products for research.

**Consent for publication:** Not applicable

**Availability of data and material:** The data generated and/or analyzed during this study are available from the corresponding author on reasonable request.

**Competing interests:** Douglas G. McNeel has ownership interest, has received research support, and serves as consultant to Madison Vaccines, Inc. which has licensed intellectual property related to this content. None of the other authors have relevant potential conflicts of interest.

**Funding:** This work was supported by the National Institutes of Health NIH R01 CA219154 and P30 CA014520. The research presented was supported under NIH award TL1TR002375. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**Authors’ contributions:** HKP wrote the manuscript and performed data analysis; DGM designed the microarray and obtained serum samples; TLN, MAN, and KL performed statistical analysis; all authors contributed to the writing and approval of the final manuscript.

**Acknowledgements:** Ken Lo

**Authors’ information (optional)**

**REFERENCES**

1. von Kleist S, Burtin P. On the specificity of autoantibodies present in colon cancer patients. Immunology. 1966;10:507–15.

2. Gold JM, Freedman SO, Gold P. Human anti-CEA antibodies detected by radioimmunoelectrophoresis. Nature New Biol. 1972;239:60–2.

3. Scanlan MJ, Chen YT, Williamson B, Gure AO, Stockert E, Gordan JD, et al. Characterization of human colon cancer antigens recognized by autologous antibodies. Int J Cancer. 1998;76:652–8.

4. Cai X, Garen A. Anti-melanoma antibodies from melanoma patients immunized with genetically modified autologous tumor cells: selection of specific antibodies from single-chain Fv fusion phage libraries. Proc Natl Acad Sci U S A. 1995;92:6537–41.

5. Grossman HB, Wedemeyer G, Stein J. Autologous antibodies to human bladder cancer. Cancer Immunology Immunotherapy [Internet]. 1988 [cited 2020 May 5];26. Available from: http://link.springer.com/10.1007/BF00199940

6. Atakan S, Bayiz H, Sak S, Poyraz A, Vural B, Yildirim AS, et al. Autologous anti-SOX2 antibody responses reflect intensity but not frequency of antigen expression in small cell lung cancer. BMC Clinical Pathology. 2014;14:24.

7. Wang X. Autoantibody Biomarkers in Prostate Cancer. Lab Med. Oxford Academic; 2008;39:165–71.

8. Zaenker P, Gray ES, Ziman MR. Autoantibody Production in Cancer—The Humoral Immune Response toward Autologous Antigens in Cancer Patients. Autoimmunity Reviews. 2016;15:477–83.

9. Wandall HH, Blixt O, Tarp MA, Pedersen JW, Bennett EP, Mandel U, et al. Cancer Biomarkers Defined by Autoantibody Signatures to Aberrant O-Glycopeptide Epitopes. Cancer Res. American Association for Cancer Research; 2010;70:1306–13.

10. Wahrenbrock MG, Varki A. Multiple hepatic receptors cooperate to eliminate secretory mucins aberrantly entering the bloodstream: are circulating cancer mucins the “tip of the iceberg”? Cancer Res. 2006;66:2433–41.

11. Olson BM, McNeel DG. Antibody and T-cell responses specific for the androgen receptor in patients with prostate cancer. The Prostate. 2007;67:1729–39.

12. Silva WA, Gnjatic S, Ritter E, Chua R, Cohen T, Hsu M, et al. PLAC1, a trophoblast-specific cell surface protein, is expressed in a range of human tumors and elicits spontaneous antibody responses. Cancer Immun. 2007;7:18.

13. Chen YT, Scanlan MJ, Sahin U, Türeci O, Gure AO, Tsang S, et al. A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. Proc Natl Acad Sci USA. 1997;94:1914–8.

14. GuhaThakurta D, Sheikh NA, Fan L-Q, Kandadi H, Meagher TC, Hall SJ, et al. Humoral Immune Response against Nontargeted Tumor Antigens after Treatment with Sipuleucel-T and Its Association with Improved Clinical Outcome. Clin Cancer Res. 2015;21:3619–30.

15. Taylor BS, Pal M, Yu J, Laxman B, Kalyana-Sundaram S, Zhao R, et al. Humoral response profiling reveals pathways to prostate cancer progression. Mol Cell Proteomics. 2008;7:600–11.

16. Ummanni R, Duscharla D, Barett C, Venz S, Schlomm T, Heinzer H, et al. Prostate cancer-associated autoantibodies in serum against tumor-associated antigens as potential new biomarkers. Journal of Proteomics. 2015;119:218–29.

17. Smith HA, Maricque BB, Eberhardt J, Petersen B, Gulley JL, Schlom J, et al. IgG Responses to Tissue-Associated Antigens as Biomarkers of Immunological Treatment Efficacy. J Biomed Biotechnol [Internet]. 2011 [cited 2020 Jun 8];2011. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3010827/

18. Morse MD, McNeel DG. Prostate cancer patients on androgen deprivation therapy develop persistent changes in adaptive immune responses. Human Immunology. 2010;71:496–504.

19. Dunphy EJ, Eickhoff JC, Muller CH, Berger RE, McNeel DG. Identification of Antigen-Specific IgG in Sera from Patients with Chronic Prostatitis. J Clin Immunol. 2004;24:492–502.

20. Lang JM, Wallace M, Becker JT, Eickhoff JC, Buehring B, Binkley N, et al. A Randomized Phase II Trial Evaluating Different Schedules of Zoledronic Acid on Bone Mineral Density in Patients With Prostate Cancer Beginning Androgen Deprivation Therapy. Clin Genitourin Cancer [Internet]. 2013 [cited 2020 Mar 18];11. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3836858/

21. McNeel DG, Dunphy EJ, Davies JG, Frye TP, Johnson LE, Staab MJ, et al. Safety and Immunological Efficacy of a DNA Vaccine Encoding Prostatic Acid Phosphatase in Patients With Stage D0 Prostate Cancer. J Clin Oncol. 2009;27:4047–54.

22. Maricque BB, Eickhoff JC, McNeel DG. Antibody responses to prostate-associated antigens in patients with prostatitis and prostate cancer. Prostate. 2011;71:134–46.

23. Mishra N, Caciula A, Price A, Thakkar R, Ng J, Chauhan LV, et al. Diagnosis of Zika Virus Infection by Peptide Array and Enzyme-Linked Immunosorbent Assay. MBio. 2018;9.

24. Lo KC, Sullivan E, Bannen RM, Jin H, Rowe M, Li H, et al. Comprehensive Profiling of the Rheumatoid Arthritis Antibody Repertoire. Arthritis & Rheumatology. 2020;72:242–50.

25. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009;10:48.

26. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. PLOS ONE. Public Library of Science; 2011;6:e21800.

27. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. Oxford Academic; 2019;47:D506–15.

28. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2019. Available from: https://www.R-project.org/

29. Ha S, Bb M, J E, B P, Jl G, J S, et al. IgG responses to tissue-associated antigens as biomarkers of immunological treatment efficacy. J Biomed Biotechnol. 2010;2011:454861–454861.

30. Guo Z-W, Meng Y, Zhai X-M, Xie C, Zhao N, Li M, et al. Translated Long Non-Coding Ribonucleic Acid ZFAS1 Promotes Cancer Cell Migration by Elevating Reactive Oxygen Species Production in Hepatocellular Carcinoma. Front Genet [Internet]. 2019 [cited 2020 Mar 3];10. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6861293/

31. Ji Z, Song R, Regev A, Struhl K. Many lncRNAs, 5’UTRs, and pseudogenes are translated and some are likely to express functional proteins. eLife [Internet]. [cited 2020 Mar 3];4. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4739776/

32. Ingolia NT, Lareau LF, Weissman JS. Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity of Mammalian Proteomes. Cell. 2011;147:789–802.

33. Yan Y, Sun N, Wang H, Kobayashi M, Ladd JJ, Long JP, et al. Whole Genome-Derived Tiled Peptide Arrays Detect Prediagnostic Autoantibody Signatures in Non-Small-Cell Lung Cancer. Cancer Res. 2019;79:1549–57.

34. McNeel DG, Nguyen LD, Storer BE, Vessella R, Lange PH, Disis ML. Antibody immunity to prostate cancer associated antigens can be detected in the serum of patients with prostate cancer. J Urol. 2000;164:1825–9.

35. McNEEL DOUGLAS G., Nguyen Lan D., Storer Barry E., Vessella Robert, Lange Paul H., Disis Mary L. Antibody immunity to prostate cancer associated antigens can be detected in the serum of patients with prostate cancer. Journal of Urology. 2000;164:1825–9.

36. Olson BM, McNeel DG. Antibody and T-cell responses specific for the androgen receptor in patients with prostate cancer. Prostate. 2007;67:1729–39.

37. Wang X, Yu J, Sreekumar A, Varambally S, Shen R, Giacherio D, et al. Autoantibody signatures in prostate cancer. N Engl J Med. 2005;353:1224–35.

38. Francoeur A-M. Antibody Fingerprinting: A Novel Method for Identifying Individual People and Animals. Nat Biotechnol. Nature Publishing Group; 1988;6:822–5.

39. Neiman M, Hellström C, Just D, Mattsson C, Fagerberg L, Schuppe-Koistinen I, et al. Individual and stable autoantibody repertoires in healthy individuals. Autoimmunity. Taylor & Francis; 2019;52:1–11.