Antibody Responses to Different Proteins in Prostate Cancer Patients

Tun Lee Ng and Michael A. Newton

May 05, 2020

Contents

1 Introduction

This project aims to characterize antibody responses to a wide variety of proteins in prostate cancer patients at different stages of the disease. 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. In this study, there were healthy subjects and patients with different stages of prostate cancer

- `new_dx`: newly diagnosed,
- `nmCSPC`: non-metastatic castration-sensitive
- `mCSPC`: metastatic castration-sensitive,
- `nmCRPC`: non-metastatic castration-resistant,
- `mCRPC`: metastatic castration-resistant

Each patient's serum was assayed in a number of replicates, `rep`, which were 1, 2, or 3, for peptide-specific IgG responses using the microarray. We remove patients with `rep` = 1. The criterion for a positive call on a peptide for a patient was that they had to meet the signal (fluorescence level) threshold in at least two of the technical replicates. Since this is not possible for patients without technical replicates, we exclude them for consistency.

Note that these 94 patient-records do not represent distinct patients, as certain patients were measured at two different stages of prostate cancer. There are 6 patients who had replicates for both stages. We removed their earlier-stage records, and finally arrive at 94 distinct patients.
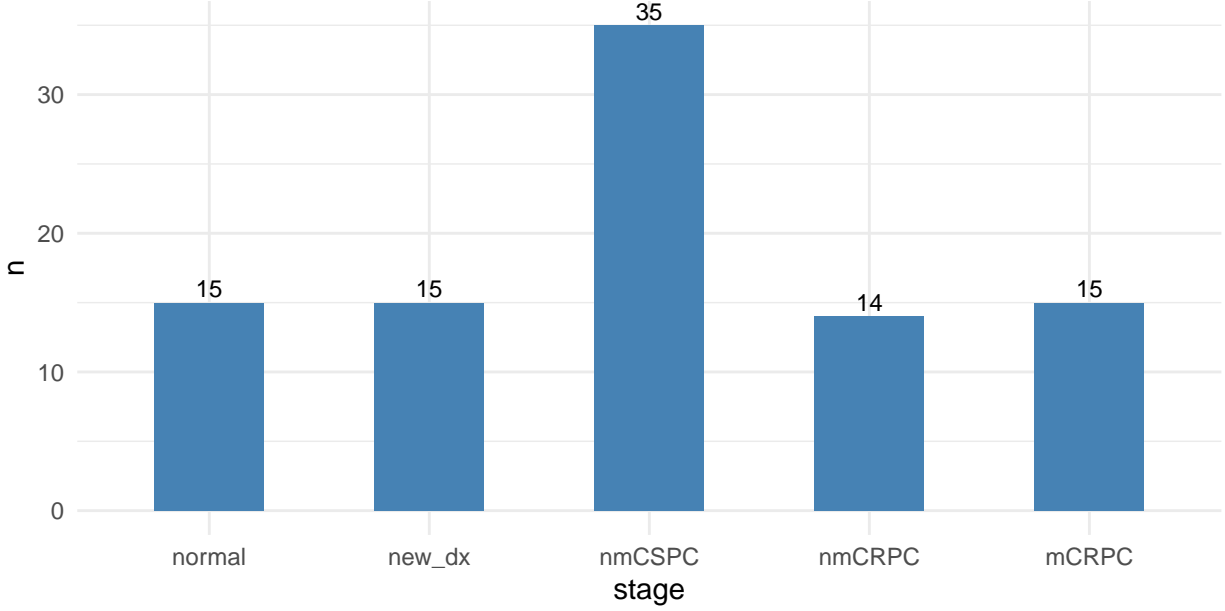
We will utilize both fluorescence levels data and patients' binary calls data to **investigate if patients at different stages of prostate cancer exhibit different antibody responses to certain peptide chains or proteins**. We take $\log_2$ transformation on the fluorescence levels prior to subsequent steps in our analysis.

2 Reproducibility of Replicates

We have assessed the issue of replicate reproducibility by looking at (Pearson and Spearman's Rank) correlation coefficients between patients' fluorescence levels (both with and without $\log_2$ transformation). Another approach is to measure how much variation the technical replicates are contributing to the overall variation in the data. If the replicates are reproducible, ie. they "largely agree with one another", then the technical variation in the dataset should be minimal.

Everytime when the fluorescence levels were measured (with replicates) for patient's stage effects, there are two sources of random variation at play, namely

- patient/subject `random effect`: this reflects the biological variation of a patient (as opposed to the `fixed effect` term, which would be the cancer stage effect in this experiment)
- (residual) random error: measuring replicates of a patient is itself a source of technical variation.

Specifically,

$$y_{ijk} = \mu + \beta_i + b_j + \epsilon_{ijk},$$

where

- y_{ijk} denotes the $\log_2$ fluorescence level of a replicate,
- μ denotes the grand mean/intercept,
- β_i denotes the fixed effect term, ie. cancer stage, with i indexing the patients' cancer stage,
- b_j denotes the random effect term, ie. individual patient, with j indexing the patients,
- ϵ_{ijk} denotes the (residual) random error of the model, with k indexing the replicates.

This is the linear mixed-effects model, which we deploy using the R package `lme4`. The model estimates the two sources of variation: $\hat{\sigma}_b^2$ (biological variation) and $\hat{\sigma}_\epsilon^2$ (technical variation). Ideally, biological variation

should dominate technical variation since the replicates' variance $\hat{\sigma}_\epsilon^2$ should be minimal. Hence, we are interested in the estimated proportion of random-effect variance to total variance

$$\frac{\hat{\sigma}_b^2}{\hat{\sigma}_b^2 + \hat{\sigma}_\epsilon^2},$$

and we would like to see if this ratio is close to one. For each of the 177,604 peptides, we deploy this mixed-effect model, and plot the histogram of the estimated proportions of variances.

Histogram of peptide–level proportion of random–effect variance to total variance
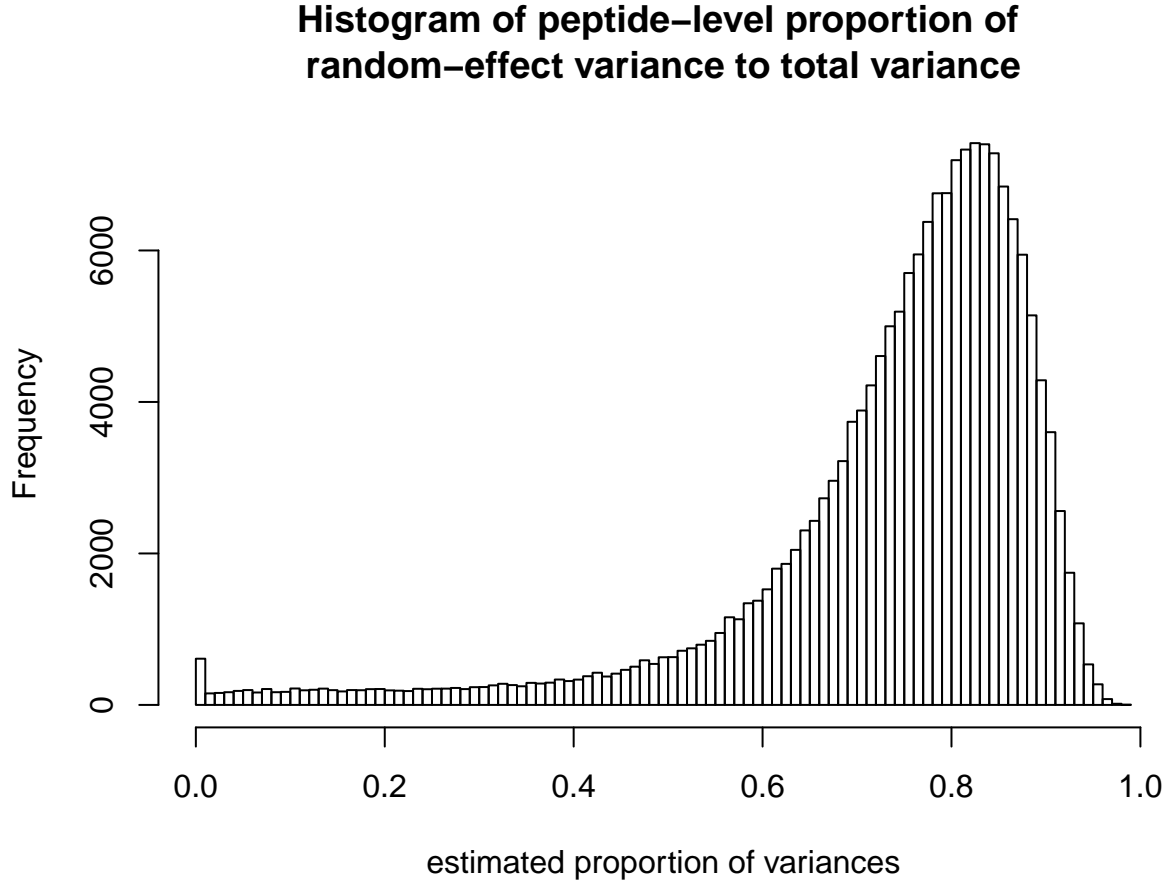
As expected, the histogram amasses at values near one, indicating that most of the variation in the ($\log_2$) fluorescence data is attributable to the biological variation of the patients and not the technical replicates themselves, which also suggests reproducibility of the replicates. The little spike at zero estimated proportions is due to the 477 singular cases where the fitted random-effect variance $\hat{\sigma}_b^2$ is close/equal to zero.

3 Tests on Binary Calls

We acknowledge the consensus that the binary calls on a peptide of a patient could be conservative – out of 177,604 peptides, only 37919 of them have at least one call among all patients. For each of these 37919 peptides, we run a logistic regression based on these binary calls of the patients in order **to determine if calls are significantly different among patients of different cancer stages**.

Specifically, for each of these 37919 peptides, we fit the following model:

$$\text{logit}\left(y_{ij}^{\text{calls}}\right) = \mu + \beta_i + \epsilon_{ij},$$

where

- y_{ij}^{calls} denotes the binary call of the peptide of a patient: 1 if the fluorescence levels meet the signal threshold in at least two replicates of the patient, and 0 otherwise,
- μ denotes the grand mean/intercept,
- β_i denotes the cancer stage,
- ϵ_{ij} denotes the random error of the model, with j indexing the patients,

and compute the deviance p-values: (null_deviance - residual_deviance) $\sim \chi^2$ with 4 degrees of freedom. We plot the histogram of the 37919 p-values.

Logistic Regression Deviance p–values
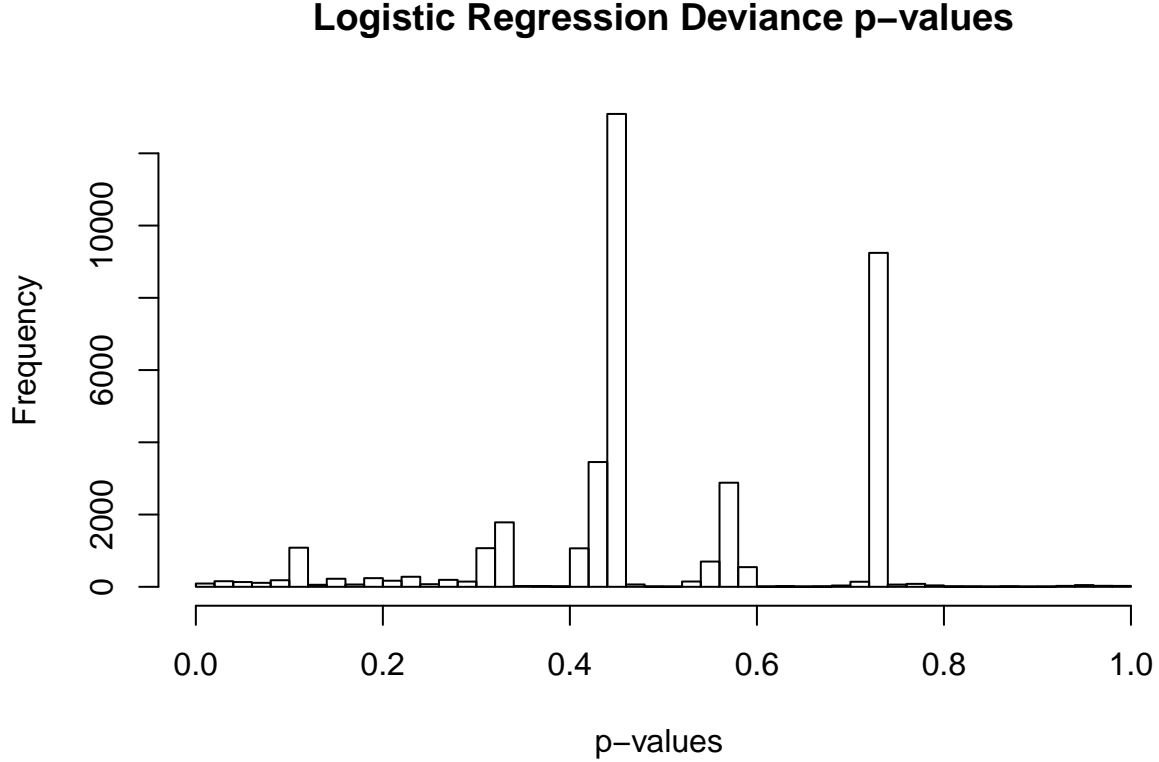
It appears that there are hardly any signals of different calls pattern among patients of different cancer stages, which corroborates with the results in Figures 2, 3 and 4 in the main manuscript. After correcting for false discovery rate, no peptides appear to be significant.

4   Tests on Fluorescence Levels with Binary-calls-filtering

We could lose a lot of information due to the conservative signal threshold of the calls data. Instead, what if we use the calls data as a filtering criterion. Specifically, we first compute the median of the $\log_2$ fluorescence levels of each patient's replicates, and then subset the 37919 peptides have at least one call among all patients. For each of these 37919 peptides, we perform one-way ANOVA (analysis of variance) to test the following:
H_0: Antibody responses are the same for patients at different stages of prostate cancer.
H_1: Antibody responses are not the same for patients at different stages of prostate cancer.

After getting p-values for the 37919 peptides, we plot the p-value histogram.

p−values distribution for 37919 peptides
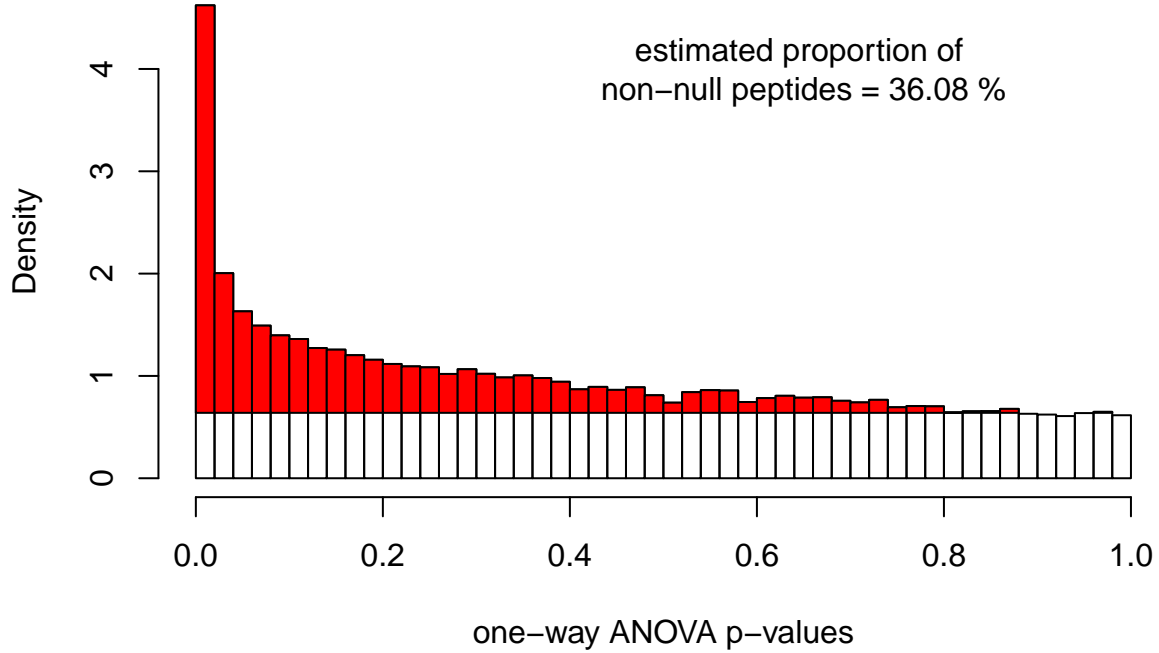
If cancer-stage effect is not present in our peptide array data, then the p-values from the ANOVA would have a uniform distribution between 0 and 1, and we expect to see a rather flat-shaped histogram of p-values.

However, the p-values histogram exhibits large counts of significant p-values (p-values close to zero), and the shape of histogram flattens off exponentially with larger p-values. Such a large count of significant p-values may not be explained by false discovery alone, and that perhaps cancer-stage effect is indeed present in some of the peptides in our profile. The red-shaded regions of the histogram represents the estimated proportion of non-null peptides in the dataset based on Storey's q-values calculation obtained via the R package `fdrtool`. The q-value is similar to the well known p-value, except that it is a measure of significance in terms of the false discovery rate rather than the false positive rate.

We apply the Benjamini-Hochberg (BH) method on the ANOVA p-values to control for false discovery rate (FDR). The peptide counts at various BH FDR thresholds are tabulated below.

FDR threshold	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
Peptide counts	2	3	13	181	793	1074	1326	1556	1678	1810

We also tabulate the peptide counts at various q-value thresholds as shown below.

FDR threshold	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
Peptide counts	3	18	675	1147	1522	1736	1982	2208	2419	2638

It appears that the BH method gives a more conservative significant peptide counts.

4.1 Visualizing Results

We have identified 793 peptides at 5% Benjamini-Hochberg(BH) false discovery rate (FDR). We would like to obtain some graphical representations to illustrate how the $\log_2$ fluorescence levels differ across different cancer stages for these 793 peptides.

For each peptide, we remove the grand mean (row mean) of the $\log_2$ fluorescence levels for all patients, before applying the following visualization techniques:

- Principal Component Analysis (PCA)
- t-distributed Stochastic Neighbor Embedding (t-SNE)
- Heatmap

Based on our one-way-ANOVA model assumption, if there is no cancer-stage effect, we expect these residual $\log_2$ fluorescence to be random noises. Any observed (clustering) patterns among these residual data points reveal the effects of various stages of prostate cancer.

For purpose of uniformity, we also use the same color scheme to distinguish the different stages of cancer patients (notice how the spectrum of colors changes with severity of the cancer stages):

- navy for healthy subjects
- cornflower_blue for `new_dx` newly diagnosed patients
- turquoise for `nmCSPC` patients
- light_pink for `mCSPC` patients – these patients have no technical replicates and are excluded from this analysis
- dark_orange for `nmCSPC` patients
- dark_red for `mCSPC` patients

First, the PCA plot is given by
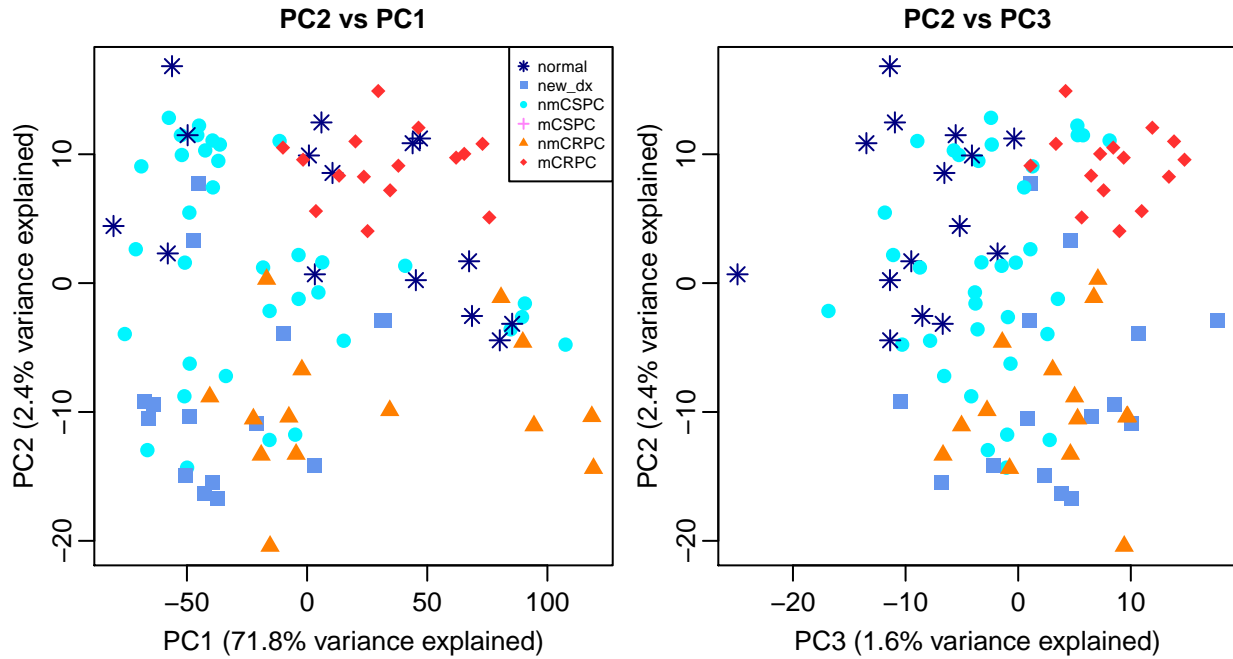
From the "*PC2 vs PC1*" plot, we observe that all mCRPC points are clustered at the topright of the panel, whereas nmCRPC observations hover at the bottom of panel. The percentage of variance explained for each principal component (PC) is shown on the axis. Note that the first principal component manages to capture most of the variation in the data.

Next, we obtain the t-SNE plot. Just like PCA, t-SNE is a dimensionality reduction technique which was first introduced by van der Maaten and Hinton in 2008.
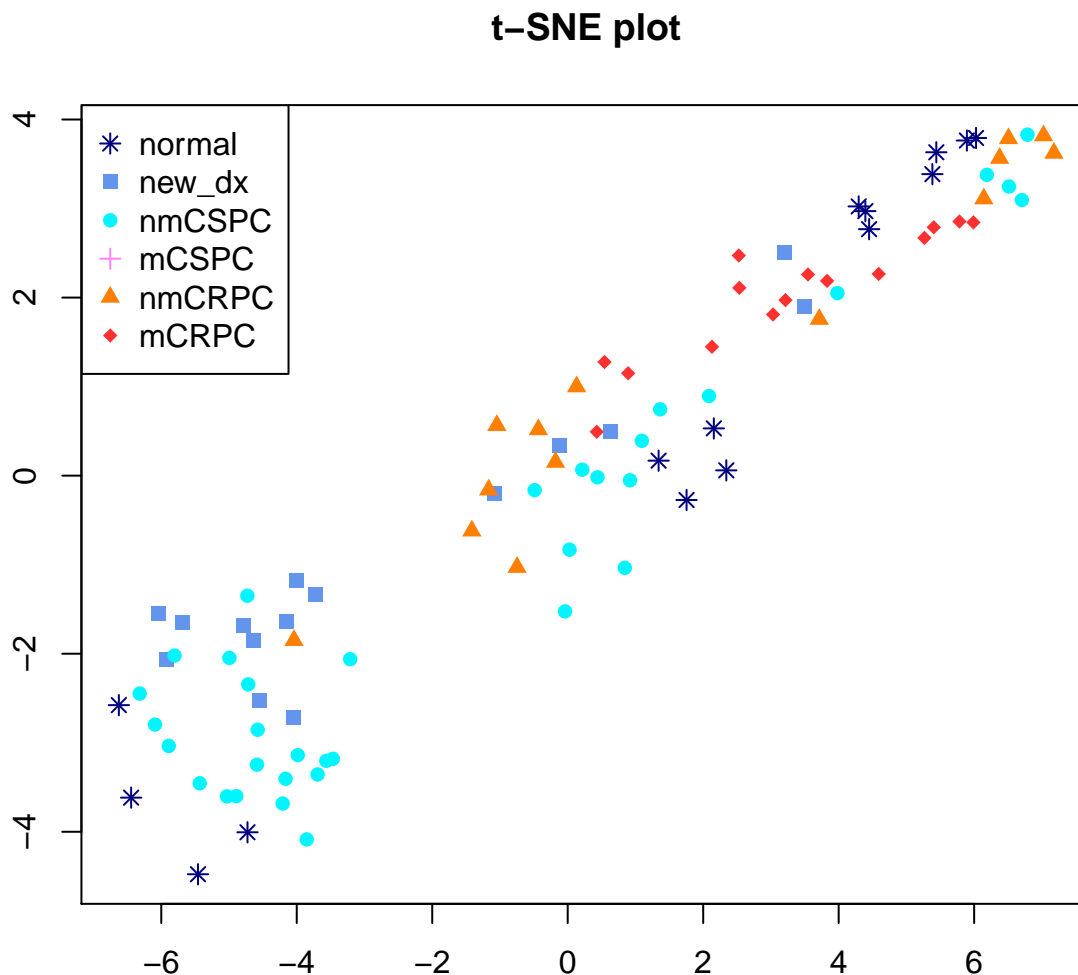
t−SNE plot

From the t-SNE plot, we observe that:

- The mCRPC points are clustered together near top-right of the plot.
- Most of the nmCRPC patients are not too far off from the mCRPC subjects in the middle of the plot.
- Most of the new_dx and nmCSPC patients are clustered at the bottom left and middle of the plot.
- Normal subjects are somewhat "all over the place".

Finally, we plot the heatmap of the $\log_2$ fluorescence (after removing row means) of the 793 peptides at 5% BH-FDR.
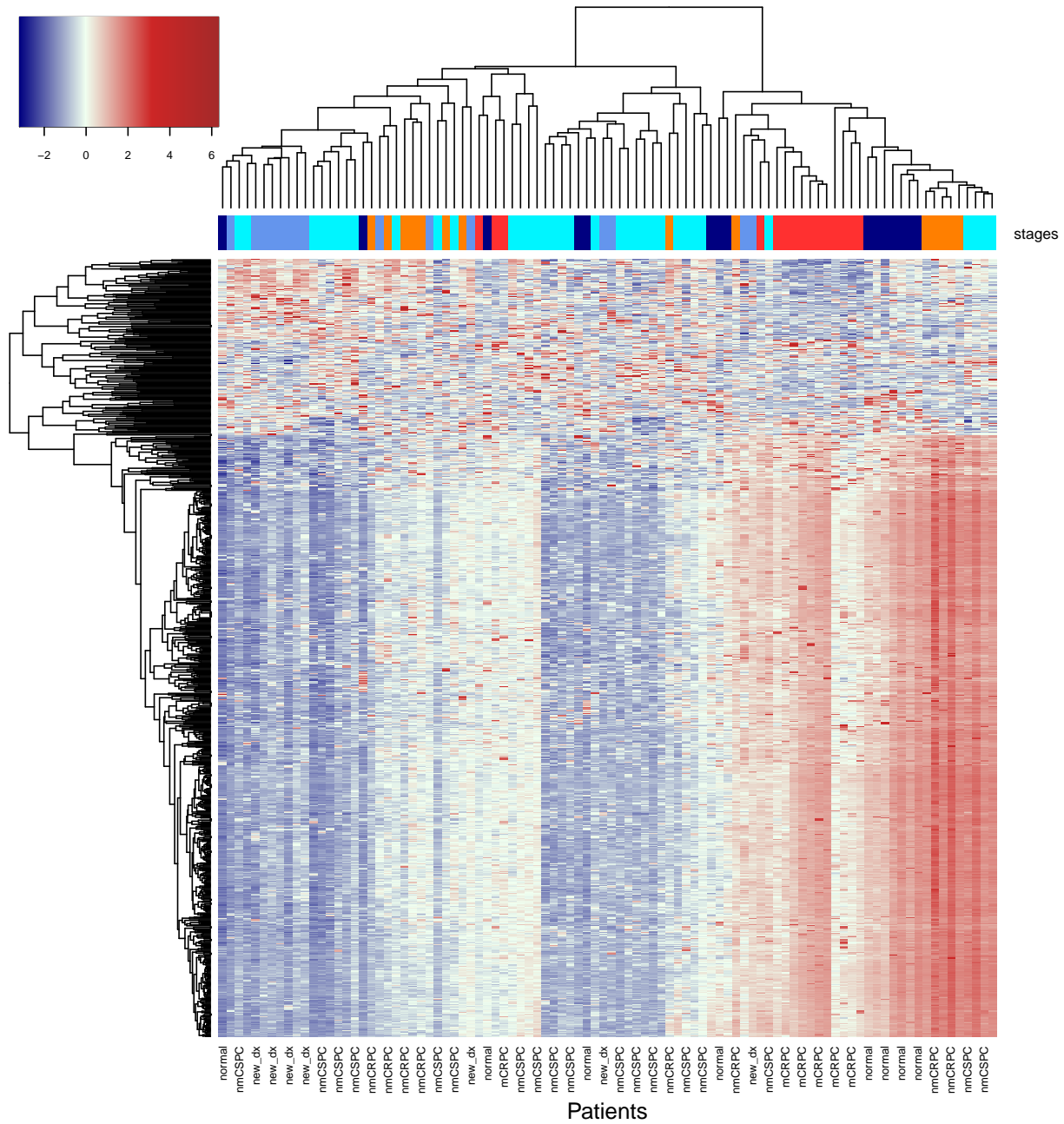
Again, we observe similar patterns that nmCSPC (colored turquoise) and new_dx (colored cornflowerblue) subjects are mostly clustered to the left part of the heatmap, whereas the mCRPC (colored red) patients and most of the nmCRPC patients (colored darkorange) are clustered to the right part of the heatmap. Again, normal subjects (colored navy) are "all over the place".

4.2 Comparing Healthy Subjects vs Cancer Patients

The one-way ANOVA is helpful in revealing peptides that exhibit significant difference in the group means of $\log_2$ fluorescence levels among patients at different stages of cancer. We may be interested to find out, for example, **if antibody response is significantly higher in cancer patients than in healthy subjects**. Using the linear-contrast method, we are able to further test this hypothesis after computing the ANOVA.

Specifically, it is a one-sided t-test comparing the averages of the $\log_2$ fluorescence levels among cancer patients and among healthy subjects, except that now, the degrees of freedom is based on the one-way ANOVA in the previous step, and the standard error of the t-test statistic is based on the root mean-squared error obtained from the one-way ANOVA and adjusted for the number of the healthy and cancer patients. This allows us to pull samples from different cancer stages and increase power – with the caveat/assumption that the patients of different cancer stages share the same level of variation in antibody response. (Note that we could avoid doing the multiplicity control like Tukey's Honest Significant Difference if we are only interested in a particular contrast/comparison and not every possible combination of pairwise comparisons)

Among the 793 peptides at 5% BH-FDR, only 28 of them are associated with higher antibody response in cancer patients than in normal subjects. We tabulate these peptides with their differences (of average $\log_2$ fluorescence levels between cancer and healthy subjects) and p-values of their t-statistics.

peptide_id	difference	tstat__pvalues
539__MAML1__9794;657	1.8619	0.0071
584__TUBA1C__84790;273	2.0708	0.0108
1098__ARGLU1__55082;197	1.7184	0.0140
1046__SRRT__51593;9	1.4562	0.0159
655__CCT2__10576;517	1.9199	0.0207
65__ZKSCAN1__7586;473	1.4761	0.0207
608__CCDC6__8030;301	1.1939	0.0248
1133__SMARCA1__6594;109	1.0842	0.0260
974__EIF3L__51386;349	0.7739	0.0284
1196__ZFP62__643836;381	1.3175	0.0290
267__ACTB__60;113	1.2223	0.0327
431__ACTG1__71;113	1.2223	0.0327
491__KDELR2__11014;145	0.1176	0.0355
591__ECHDC2__55268;1	0.1945	0.0359
478__VDAC1__7416;37	1.2779	0.0377
611__KIAA1429__25962;1717	0.7533	0.0388
PRO14;2497	1.0935	0.0394
743__AK4__205;41	1.4307	0.0394
141__CAMKK2__10645;21	1.1280	0.0411
CTA1;69	0.7611	0.0418
1004__TRIP12__9320;1145	0.6571	0.0441
150__YWHAG__7532;1	1.1831	0.0443
1406__CCDC85C__317762;109	1.2628	0.0445
633__USP54__159195;1001	1.0586	0.0457
1364__DHCR7__1717;337	0.1251	0.0464
49__U2AF1__7307;49	1.0161	0.0474
538__THRAP3__9967;649	0.8212	0.0478
985__CALD1__800;549	1.1548	0.0486

For purpose of visualization, we also produce the boxplots of $\log_2$ fluorescence levels of subjects from these two groups for some of these peptides.

539_MAML1_9794;657

log2(fluorescence)

584_TUBA1C_84790;273
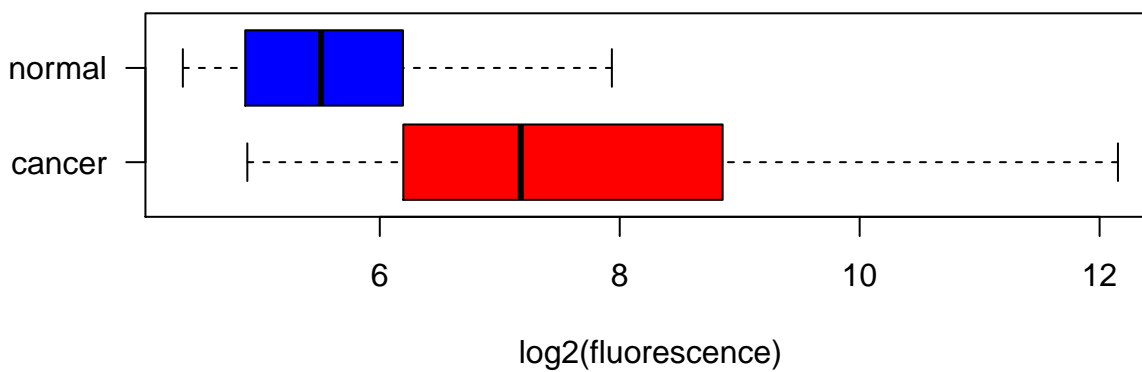
log2(fluorescence)
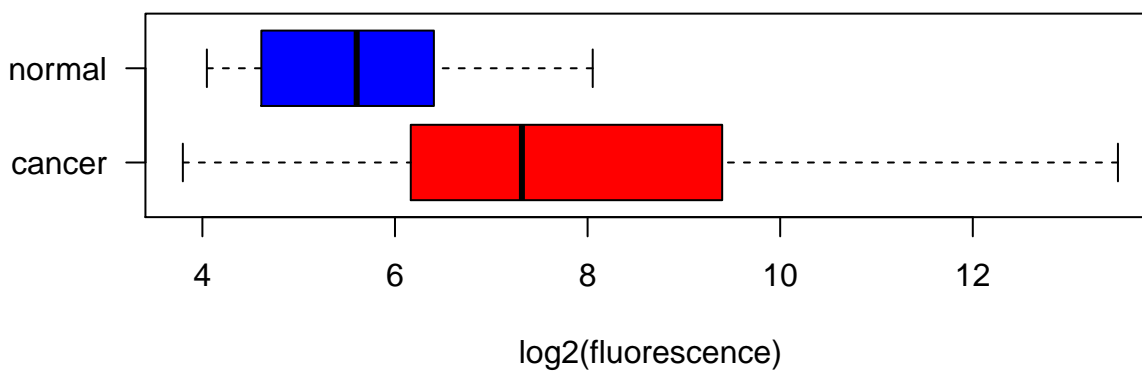
1098_ARGLU1_55082;197
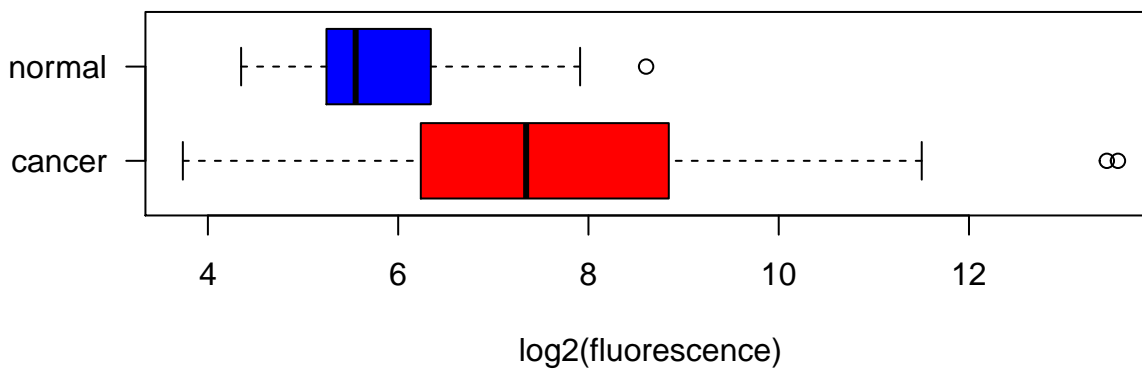
log2(fluorescence)

5 Tests on Fluorescence Levels without Filtering

As a comparison, we repeat the same calculation in the previous section, except that now, we do not apply any filtering on the peptides at all. First, we plot the histogram of p-values from the ANOVA.

p–values distribution for 177604 peptides

estimated proportion of
non–null peptides = 45.13 %
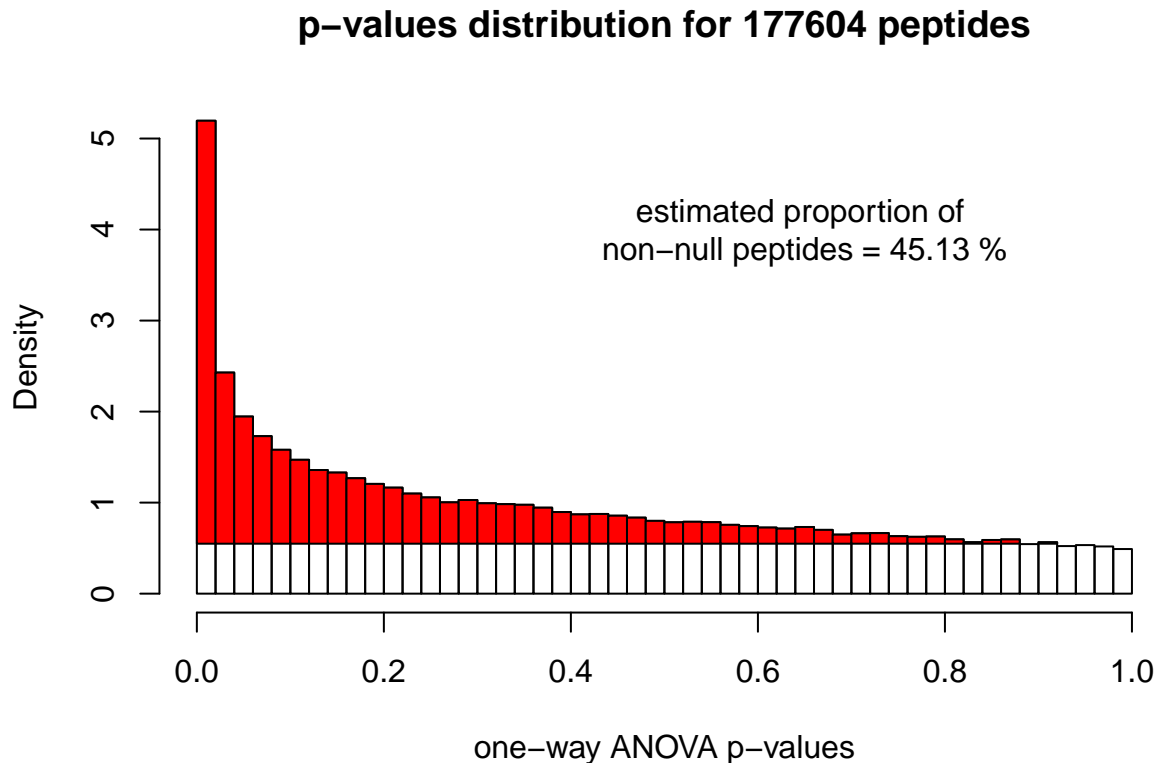
We also tabulate the peptide counts at different BH FDR thresholds.

FDR threshold	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
Peptide counts	82	204	738	1807	3128	4396	5562	6861	7858	8885

Here are also the peptide counts at different q-values thresholds.

FDR threshold	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
Peptide counts	169	1380	3732	5938	7977	9897	11708	13641	15513	17395

5.1 Visualizing Results

It appears that we have identify 3128 significant peptides at 5% BH FDR. Now, plot the resulting PCA, t-SNE and heatmap of the residual $\log_2$ fluorescence of these peptides.
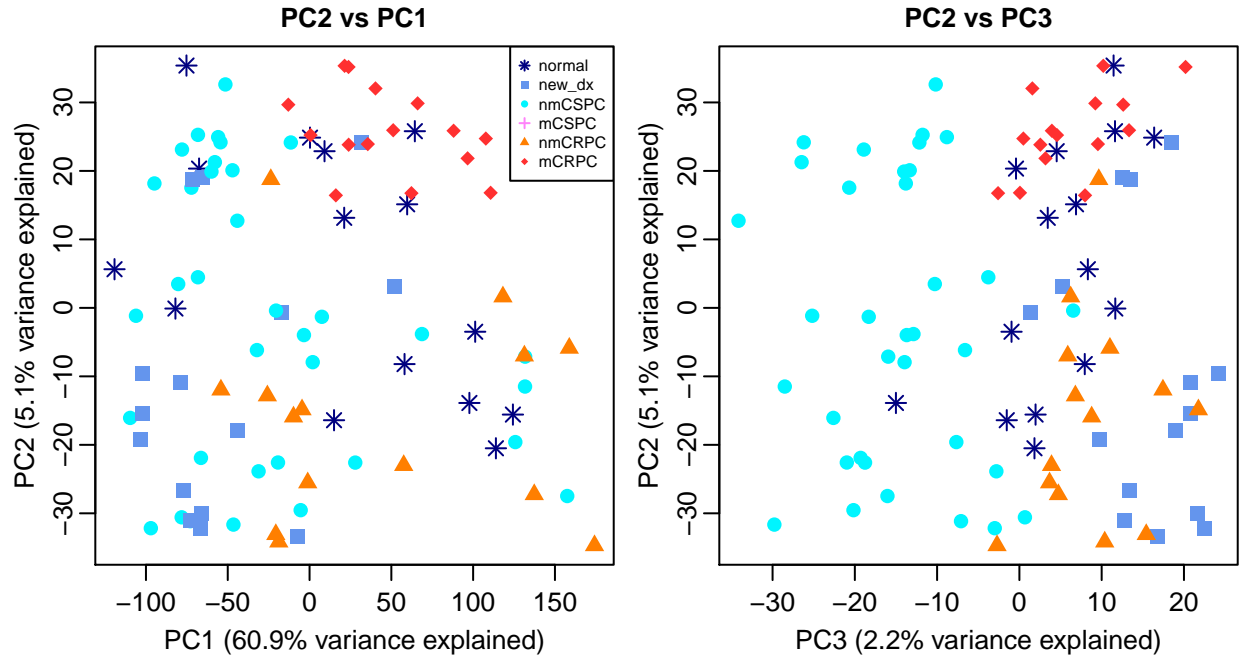
We observe similar patterns in the PCA plot as we did when we apply the calls-filtering. Obviously, now that we have working with more peptides, the percentage of variance explained in the first principal component is slightly lower than what we achieved when we apply the calls-filtering.
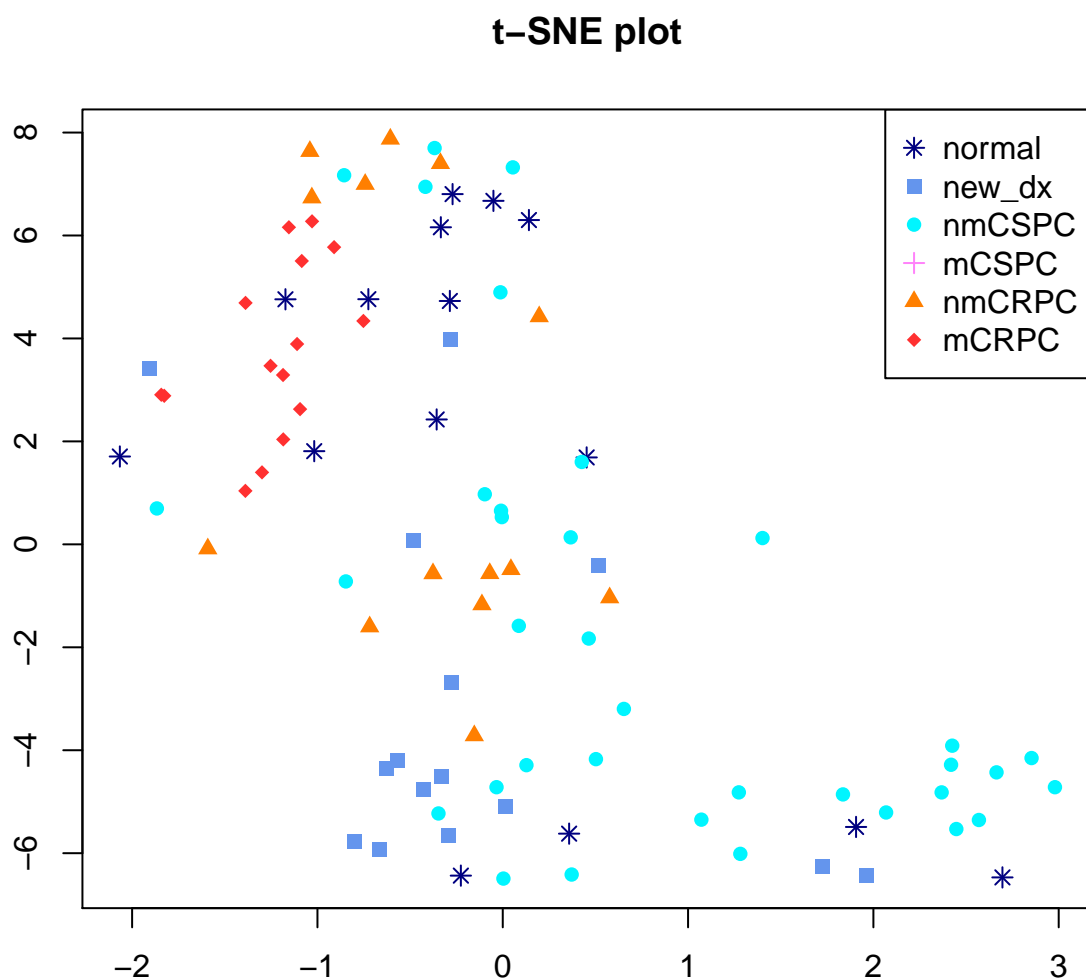
t−SNE plot

Again, we observe similar patterns in the t-SNE plot as we did when we apply the calls-filtering. There seems to be 2 clusters of nmCRPC patients.
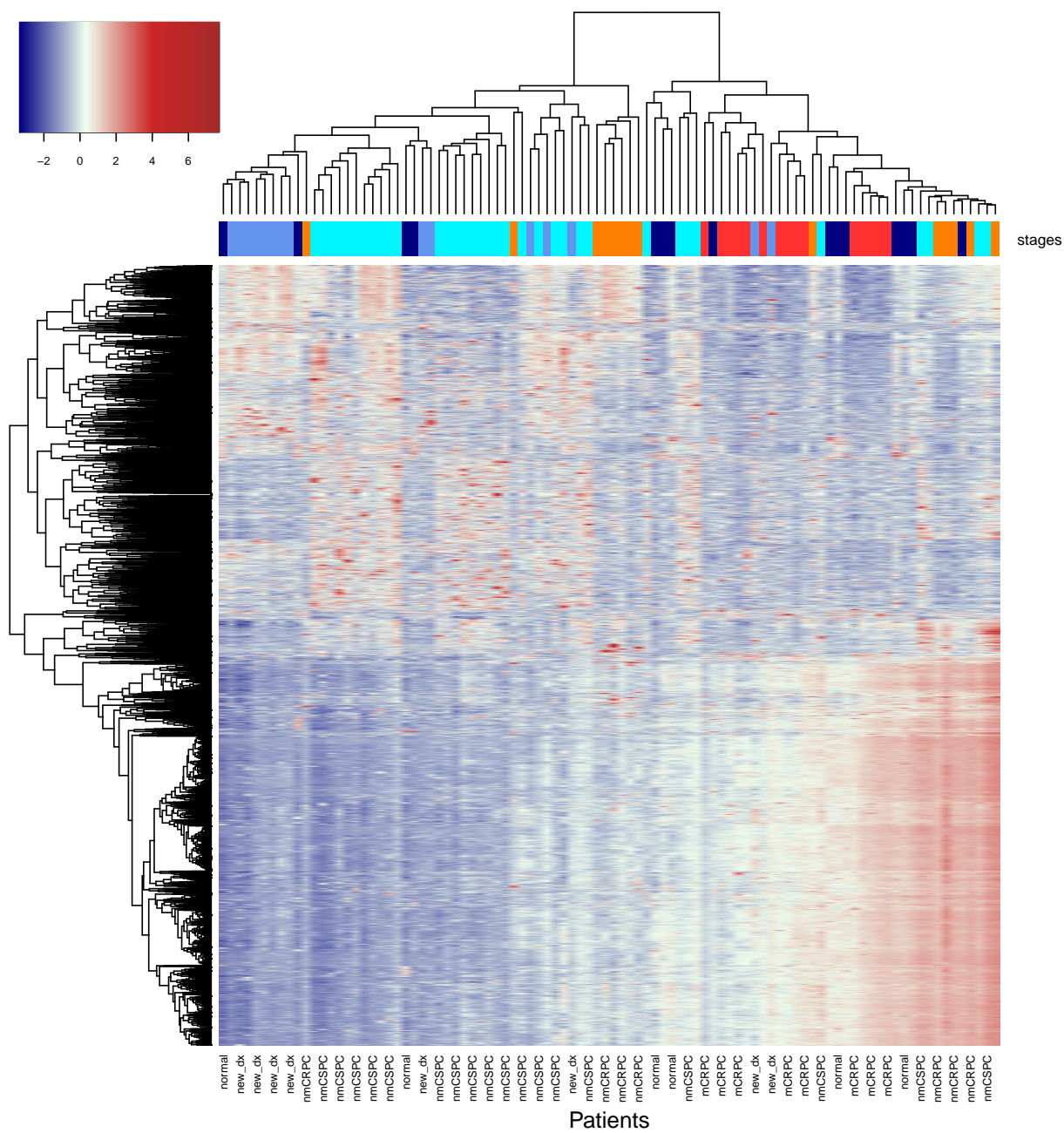
Again, we observe similar clustering patterns in the heatmap as we did when we apply the calls-filtering.

5.2 Comparing Healthy Subjects vs Cancer Patients

Finally, we repeat the linear contrast method on the 3128 peptides at 5% BH FDR to determine if fluorescence levels of these peptides are significantly higher among cancer patients than among healthy subjects.

There are 356 peptides that are associated with higher antibody response in cancer patients than in normal subjects. We tabulate these peptides with their differences (of average $\log_2$ fluorescence levels between cancer and healthy subjects) and p-values of their t-statistics.

peptide_id	difference	tstat_pvalues
539__MAML1__9794;657	1.8619	0.0071
1386__RALGAPA2__57186;997	1.4539	0.0076
788__IVNS1ABP__10625;305	0.1645	0.0084
255__RANBP2__5903;1441	1.3179	0.0085
646__FASTK__10922;153	0.2059	0.0087
385__EIF3D__8664;45	1.4657	0.0089
190__ND5__4540;249	0.2192	0.0094
912__EIF2B1__1967;125	0.1535	0.0095
1242__THOC2__57187;1189	1.7438	0.0099
CAT297__T196839__G045377__1__2510__3112__603;1	0.8343	0.0101
114__SRRM2__23524;1569	1.3350	0.0102
1081__DDX3Y__8653;301	1.6009	0.0106
1429__FRMD8__83786;213	1.0546	0.0106
191__CMTM6__54918;25	0.3225	0.0107
584__TUBA1C__84790;273	2.0708	0.0108
882__ARL8B__55207;1	0.3196	0.0115
453__NCL__4691;597	1.3914	0.0116
1026__CLIP1__6249;337	1.1904	0.0122
1204__TAF15__8148;581	1.3479	0.0126
654__RPS3__6188;109	0.2659	0.0129
385__EIF3D__8664;49	1.5236	0.0130
200__BRI3__25798;93	0.1482	0.0131
500__F11R__50848;9	0.1297	0.0132
662__AUTS2__26053;573	1.6983	0.0133
135__TRMT112__51504;5	0.2724	0.0135
489__MTMR4__9110;869	0.8476	0.0135
1098__ARGLU1__55082;197	1.7184	0.0140
471__NDUFB8__4714;81	1.3128	0.0142
556__POGZ__23126;869	1.1663	0.0143
1045__STRBP__55342;213	0.1353	0.0144
859__KIAA1522__57648;229	1.4817	0.0148
PRO30;853	0.1989	0.0150
979__RPS2__6187;153	0.3368	0.0151
CAT583__T271074__G062826__2__704__1135__432;37	0.1400	0.0153
PRO25;25	1.0527	0.0157
1046__SRRT__51593;9	1.4562	0.0159
909__RNF149__284996;281	0.1877	0.0159
1269__SLC37A3__84255;117	0.1341	0.0160
1201__UIMC1__51720;225	1.5096	0.0161
1330__PDIA4__9601;5	0.1795	0.0161
797__MARCH6__10299;465	0.1788	0.0162
1422__CADM1__23705;385	0.1290	0.0165
640__VPS35__55737;193	0.7225	0.0167
19__NPIP__9284;213	0.7591	0.0168
1063__ITGAV__3685;453	0.2723	0.0169
1173__ATP1B1__481;241	0.4046	0.0173
591__ECHDC2__55268;5	0.4165	0.0178
1308__DPM1__8813;85	1.6117	0.0179
120__NPIPL3__23117;649	1.1170	0.0183
120__NPIPL3__23117;733	1.1170	0.0183
30__LOC100132247__100132247;649	1.1170	0.0183
30__LOC100132247__100132247;733	1.1170	0.0183
306__LBR__3930;313	0.1955	0.0184
682__ARFGAP3__26286;49	0.1919	0.0185
1268__XYLT2__64132;301	0.7833	0.0187
578__WSB1__26118;81	0.9738	0.0187
101__HSPH1__10808;421	0.1672	0.0190
130__HLA-C__3107;213	0.7960	0.0190
336__STEAP1__26872;257	0.0973	0.0190
1334__HGSNAT__138050;193	0.1127	0.0192

peptide_id	difference	tstat__pvalues
1419__TBL1XR1__79718;285	1.0590	0.0192
289__NDUFA8__4702;89	0.8656	0.0194
402__C12orf62__84987;37	1.0040	0.0194
571__CMTM4__146223;133	0.2155	0.0196
639__HECTD1__25831;429	0.5327	0.0196
523__TRMT5__57570;49	1.1709	0.0197
1010__LDLR__3949;797	0.1348	0.0200
1343__DAAM1__23002;381	0.5539	0.0203
1386__RALGAPA2__57186;85	0.2588	0.0203
979__RPS2__6187;21	0.6197	0.0204
556__POGZ__23126;1101	0.9573	0.0206
655__CCT2__10576;517	1.9199	0.0207
65__ZKSCAN1__7586;473	1.4761	0.0207
559__BSDC1__55108;169	1.3160	0.0208
653__NUFIP2__57532;9	1.0511	0.0208
656__ANK3__288;797	0.1351	0.0209
872__RASD1__51655;261	1.2922	0.0209
885__TBCD__6904;969	0.1575	0.0209
198__ATP9A__10079;325	0.1770	0.0213
93__PABPC3__5042;421	0.7367	0.0213
386__LRRC8A__56262;265	0.1778	0.0216
854__SORL1__6653;1365	0.9160	0.0217
PCAT1__T351126__G082910__2__1371__1655__285;25	0.3069	0.0219
538__THRAP3__9967;73	0.3852	0.0219
1321__MLLT4__4301;765	0.3395	0.0221
474__RPL18A__6142;17	0.2767	0.0222
ADT12;129	0.6684	0.0223
2__AR__367;325	1.0911	0.0223
61__KLC3__147700;145	1.1848	0.0224
731__F5__2153;117	1.0325	0.0224
1178__C12orf51__283450;649	0.1925	0.0225
163__ATP13A3__79572;121	1.0535	0.0225
966__TARS__6897;433	1.0728	0.0225
683__MRPL3__11222;25	0.1831	0.0226
830__ZNF664__144348;97	0.8137	0.0228
266__UTRN__7402;401	0.9203	0.0230
1196__ZFP62__643836;265	1.3867	0.0231
1196__ZFP62__643836;293	1.3867	0.0231
314__TUG1__55000;25	0.2978	0.0232
677__ZNF598__90850;753	0.6518	0.0232
1046__SRRT__51593;605	0.2071	0.0233
978__CHERP__10523;413	1.1558	0.0235
448__GOLGA4__2803;229	0.9787	0.0236
646__FASTK__10922;473	0.1410	0.0239
1032__C1orf9__51430;161	0.1158	0.0241
1158__SEMA4C__54910;789	1.0786	0.0245
893__UPF1__5976;693	0.2995	0.0245
1386__RALGAPA2__57186;81	0.3405	0.0246
1062__BCLAF1__9774;73	0.8133	0.0247
843__DDB1__1642;1065	0.7300	0.0247
1146__SSBP1__6742;1	0.1938	0.0248
608__CCDC6__8030;301	1.1939	0.0248
ADT3;513	1.0341	0.0249
1174__GAS5__60674;329	0.2048	0.0249
610__USP9X__8239;1857	0.1192	0.0250
1269__SLC37A3__84255;209	0.1080	0.0254
311__EFNA1__1942;173	0.1627	0.0254
677__ZNF598__90850;285	0.6974	0.0256
872__RASD1__51655;89	0.1083	0.0256
1037__SUZ12__23512;89	0.3501	0.0257
1387__RPA1__6117;205	0.6860	0.0257
630__AHCYL1__10768;349	1.2614	0.0258
923__MRPS6__64968;57	0.1107	0.0258
497__STARD3NL__83930;101	0.1306	0.0259
1133__SMARCA1__6594;109	1.0842	0.0260
1318__GANAB__23193;773	0.9192	0.0260
529__BTBD6__90135;341	0.5582	0.0260
1258__SRSF10__10772;105	0.5432	0.0263
692__ANAPC5__51433;269	0.1629	0.0266

peptide_id	difference	tstat__pvalues
1055__KIAA0146__23514;89	1.0008	0.0268
775__KIAA2013__90231;137	0.1772	0.0268
PRCAT104.4__T230577__G053084__2__3913__4350__438;9	0.4732	0.0269
1072__CTSB__1508;197	1.0329	0.0269
632__NOS1__4842;81	0.2329	0.0272
781__ATP1A1__476;301	0.1147	0.0274
MIR205HG.2__T029617__G006597__1__387__779__393;21	0.1765	0.0275
PRCAT3.1__T183162__G042038__1__6945__7547__603;133	0.3818	0.0275
1378__CHD7__55636;2017	0.7821	0.0278
PRCAT236__T273012__G063396__2__1244__1561__318;53	0.1859	0.0280
144__SPON2__10417;233	0.1545	0.0280
CTA19;117	0.7004	0.0281
285__NPRL3__8131;397	0.1337	0.0281
21__GTF2I__2969;777	0.6205	0.0282
354__SLC12A7__10723;981	0.8015	0.0283
407__EDEM3__80267;645	0.3003	0.0284
974__EIF3L__51386;349	0.7739	0.0284
1076__TANC1__85461;33	1.4025	0.0285
1256__C19orf28__126321;325	0.5919	0.0286
1007__SEL1L3__23231;1037	0.1395	0.0287
1007__SEL1L3__23231;169	0.2695	0.0289
951__ENTPD6__955;61	0.2075	0.0289
1196__ZFP62__643836;381	1.3175	0.0290
477__SRSF5__6430;89	0.9165	0.0293
1431__ZNF462__58499;1653	0.1964	0.0297
17__CYTB__4519;89	0.1066	0.0297
692__ANAPC5__51433;709	0.1669	0.0299
188__TOMM7__54543;33	1.1299	0.0301
389__RPL6__6128;41	0.2497	0.0304
470__SCCPDH__51097;293	0.1407	0.0304
1032__C1orf9__51430;781	0.1501	0.0305
912__EIF2B1__1967;25	1.0086	0.0305
1064__TMEM184B__25829;305	0.1544	0.0306
37__RPL34__6164;73	0.5731	0.0306
77__SP100__6672;281	0.1798	0.0307
PRO3;525	0.2980	0.0308
CTA6;93	0.8390	0.0309
1062__BCLAF1__9774;61	0.5798	0.0312
1062__BCLAF1__9774;85	0.4328	0.0313
946__NKTR__4820;485	0.6716	0.0313
1433__SASH1__23328;993	0.8614	0.0315
CAT1686__T107057__G025268__2__4537__5037__501;137	0.4395	0.0316
818__ILF3__3609;237	0.9665	0.0316
1340__VSIG10L__147645;789	0.1461	0.0317
178__ZC3H11A__9877;697	1.2809	0.0317
449__GALNT11__63917;17	0.1474	0.0318
PRCAT28.1__T183157__G042036__1__3585__4718__1134;5	0.4663	0.0319
PRCAT28.4__T183159__G042036__1__1382__2515__1134;5	0.4663	0.0319
833__ZMPSTE24__10269;381	0.0905	0.0319
1306__SLC38A1__81539;361	0.1330	0.0320
1335__KTN1__3895;1101	1.3326	0.0320
379__ELP2__55250;201	0.8238	0.0320
PRCAT193__T143309__G033348__1__7398__8183__786;237	0.5505	0.0324
1041__MLL__4297;2129	0.2140	0.0326
267__ACTB__60;113	1.2223	0.0327
431__ACTG1__71;113	1.2223	0.0327
1392__C20orf108__116151;149	0.2062	0.0328
309__TMC8__147138;537	0.0883	0.0328
1248__DHTKD1__55526;21	0.3491	0.0329
749__STEAP2__261729;305	0.1247	0.0331
29__HSP90B1__7184;685	0.2225	0.0333
1072__CTSB__1508;53	0.6642	0.0335
411__PPDPF__79144;5	0.4175	0.0337
854__SORL1__6653;481	0.1991	0.0338
847__TMC5__79838;741	0.2156	0.0340
163__ATP13A3__79572;245	0.2478	0.0341
543__RHOU__58480;13	0.6965	0.0342
ADT14;669	1.1259	0.0343
231__B4GALT1__2683;21	0.1181	0.0343

(continued)

peptide_id	difference	tstat__pvalues
611__KIAA1429__25962;121	0.4611	0.0344
959__NCOR2__9612;1605	0.8209	0.0344
1080__IRX4__50805;501	1.3802	0.0345
191__CMTM6__54918;137	0.1211	0.0345
1082__MCM7__4176;513	0.9584	0.0346
1178__C12orf51__283450;1365	0.6867	0.0347
1137__KAT6A__7994;733	0.1366	0.0348
314__TUG1__55000;41	1.0880	0.0348
749__STEAP2__261729;405	0.1576	0.0349
5__GGT6__124975;185	0.2371	0.0353
491__KDELR2__11014;145	0.1176	0.0355
918__MT1G__4495;41	1.0365	0.0355
935__HIPK3__10114;1049	0.8737	0.0355
1362__ADD1__118;381	0.3540	0.0357
591__ECHDC2__55268;1	0.1945	0.0359
869__COX7B__1349;41	0.1017	0.0363
165__LSM14A__26065;397	0.6848	0.0364
CAT634.1__T277460__G064362_2__111__689__579;69	0.6698	0.0367
1137__KAT6A__7994;29	0.9300	0.0369
1076__TANC1__85461;705	0.1408	0.0371
33__ANKK1__255239;677	0.6941	0.0372
676__EIF3A__8661;481	0.8135	0.0372
77__SP100__6672;305	0.9443	0.0374
PRO30;837	1.0270	0.0375
1108__CUX1__1523;1285	0.1900	0.0375
490__ABCC5__10057;329	0.2442	0.0375
1178__C12orf51__283450;257	0.6901	0.0376
833__ZMPSTE24__10269;209	0.1445	0.0376
478__VDAC1__7416;37	1.2779	0.0377
562__COX17__10063;53	0.1181	0.0377
1361__IARS__3376;725	0.4215	0.0380
1196__ZFP62__643836;217	1.1239	0.0383
1375__ARID4B__51742;601	0.0943	0.0383
692__ANAPC5__51433;129	0.1322	0.0383
1099__WDR90__197335;585	0.3226	0.0384
881__PYCR1__5831;197	0.7078	0.0384
1144__ATP5I__521;5	0.1766	0.0385
1156__PREPL__9581;509	0.7617	0.0385
306__LBR__3930;329	0.0773	0.0388
611__KIAA1429__25962;1717	0.7533	0.0388
96__ND1__4535;1837	0.1898	0.0389
1041__MLL__4297;3925	0.1623	0.0391
1037__SUZ12__23512;109	0.2470	0.0393
PRO14;2497	1.0935	0.0394
743__AK4__205;41	1.4307	0.0394
364__TCF25__22980;605	0.0923	0.0395
1419__TBL1XR1__79718;421	1.0343	0.0397
884__SAP18__10284;141	0.5235	0.0398
1010__LDLR__3949;737	1.0057	0.0400
843__DDB1__1642;1069	0.6412	0.0401
975__AGT__183;45	0.8280	0.0401
430__NPC2__10577;1	0.1377	0.0402
CAT1454__T071522__G016513_1__1122__1892__771;137	0.0823	0.0403
519__NBPF11__200030;497	0.4605	0.0403
1138__SNRNP200__23020;1081	0.2192	0.0404
86__MCL1__4170;177	0.5519	0.0404
542__TM2D3__80213;1	0.5211	0.0405
183__CD24__100133941;65	0.1017	0.0406
456__PDCD6__10016;153	0.1276	0.0406
56__ALPK1__80216;1233	0.9445	0.0408
725__GAA__2548;265	0.1014	0.0409
CAT1686__T107057__G025268_1__5122__5934__813;113	0.8336	0.0410
1207__PAFAH1B1__5048;341	0.1112	0.0411
141__CAMKK2__10645;21	1.1280	0.0411
29__HSP90B1__7184;505	0.8768	0.0411
1199__PRKAR2A__5576;297	0.1654	0.0412
459__CLTC__1213;785	0.0893	0.0413
1242__THOC2__57187;385	0.1807	0.0414
653__NUFIP2__57532;381	1.0389	0.0416

(continued)

peptide_id	difference	tstat_pvalues
BOLA3_AS1_T191973_G044207_2_402_665_264;1	0.7923	0.0417
CTA1;69	0.7611	0.0418
1108_CUX1_1523;1289	0.1519	0.0419
1280_FZD4_8322;265	0.1153	0.0420
739_RGS10_6001;81	0.6912	0.0421
682_ARFGAP3_26286;357	0.7731	0.0423
PRO18;209	0.5309	0.0424
1039_ABAT_18;425	0.2485	0.0424
842_NEURL1B_54492;517	0.3419	0.0425
797_MARCH6_10299;373	0.1323	0.0427
1236_KLHDC10_23008;73	0.2909	0.0428
1285_MICAL2_9645;969	1.0813	0.0428
639_HECTD1_25831;985	0.2287	0.0428
1160_CYFIP1_23191;209	0.4152	0.0429
255_RANBP2_5903;301	0.6503	0.0429
1261_PIGT_51604;381	0.0794	0.0431
639_HECTD1_25831;1841	0.1438	0.0431
1108_CUX1_1523;1405	1.0791	0.0432
549_REPIN1_29803;69	0.1650	0.0433
1269_SLC37A3_84255;93	0.1795	0.0435
660_GRB10_2887;389	0.1690	0.0435
701_ECI1_1632;13	0.3184	0.0435
PRCAT116_T182782_G041985_2_2888_3328_441;85	0.4604	0.0436
1178_C12orf51_283450;2965	0.1336	0.0439
552_ADAR_103;909	0.1193	0.0439
PRCAT282_T344692_G081083_1_3008_3415_408;117	0.2234	0.0440
366_TMEFF2_23671;305	0.1232	0.0440
CTA15;193	0.0882	0.0441
1004_TRIP12_9320;1145	0.6571	0.0441
337_ATP1B3_483;89	0.7357	0.0441
820_MON1B_22879;353	0.4067	0.0441
1196_ZFP62_643836;741	1.0218	0.0443
150_YWHAG_7532;1	1.1831	0.0443
516_TACSTD2_4070;121	0.2607	0.0443
1406_CCDC85C_317762;109	1.2628	0.0445
1377_TRIM2_23321;601	0.3587	0.0447
CTA11;49	0.3507	0.0448
5_GGT6_124975;181	0.1992	0.0449
351_CCT4_10575;469	0.6923	0.0452
1221_COMMD3_23412;129	0.6045	0.0453
422_EIF3C_8663;569	0.1625	0.0454
CAT2217.1_T227650_G052508_2_19310_20443_1134;33	0.1023	0.0457
CAT2217.3_T227648_G052508_2_18684_19817_1134;33	0.1023	0.0457
CAT2217.4_T227647_G052508_2_13944_15077_1134;33	0.1023	0.0457
633_USP54_159195;1001	1.0586	0.0457
120_NPIPL3_23117;85	0.3616	0.0459
30_LOC100132247_100132247;85	0.3616	0.0459
152_FAM162A_26355;73	0.8205	0.0460
449_GALNT11_63917;405	1.0908	0.0460
578_WSB1_26118;409	0.1331	0.0462
1364_DHCR7_1717;337	0.1251	0.0464
1405_ATXN2_6311;1221	0.8494	0.0468
514_NR2F2_7026;377	0.0839	0.0468
7_HDAC10_83933;589	0.2989	0.0468
818_ILF3_3609;225	0.1098	0.0469
865_PSME1_5720;157	0.7166	0.0469
1402_A2M_2;357	0.1071	0.0470
CALML3_AS1.5_T036343_G008259_2_385_1059_675;1	0.3479	0.0472
1151_SAV1_60485;33	0.2578	0.0472
CAT625.2_T275901_G064124_1_899_2041_1143;221	0.0805	0.0473
1431_ZNF462_58499;1389	0.9080	0.0473
251_COX2_4513;33	0.5396	0.0473
549_REPIN1_29803;389	0.7453	0.0473
1108_CUX1_1523;837	1.1472	0.0474
985_CALD1_800;509	1.0753	0.0474
1189_HDAC1_3065;21	0.4489	0.0476
689_ROMO1_140823;33	0.1456	0.0476
ADT14;381	0.1201	0.0477
1129_EZH2_2146;273	0.1525	0.0478

peptide_id	difference	tstat_pvalues
1380__PSMD3__5709;97	0.1774	0.0480
16__ORAI1__84876;121	0.0922	0.0481
300__GNL3__26354;241	0.6180	0.0481
1252__SMG1__23049;1285	0.1774	0.0483
284__DSP__1832;1533	0.6084	0.0486
985__CALD1__800;549	1.1548	0.0486
813__RPS29__6235;5	0.5594	0.0487
1031__SOCS7__30837;293	1.1408	0.0488
PRO3;37	0.5806	0.0489
894__PPP1R16A__84988;321	0.3836	0.0489
CTA11;133	1.0684	0.0492
587__BTF3__689;161	0.7396	0.0492
1127__IGF1R__3480;209	0.5713	0.0494
46__ERGIC1__57222;129	0.8547	0.0494
PRCAT67__T183175__G042042__1__16548__18869__2322;317	0.6842	0.0496
1424__ITPRIPL2__162073;253	0.2101	0.0496
PCA29;305	0.5536	0.0498
1034__OCIAD2__132299;97	0.1671	0.0498
361__EIF1__10209;25	0.7168	0.0499
369__KIAA1244__57221;1709	0.1818	0.0499

6   Concluding Remarks

- Calls are conservative – most of the calls are zero anyway. No cancer-stage signals are observable from the calls data.
- There appears to be cancer stage effects present in the $\log_2$ fluorescence data with or without filtering peptides based on calls. The clustering patterns of data points are similar with or without calls-filtering.