Supplementary Figure S1

Study Design for the MS experiments

Supplementary Figure S2 Proteins and Peptides coverage Plots:

The left vertical axis indicates MS experiment. The right vertical axis indicates the number of peptides or proteins detected in that MS experiment. For example, 3259 peptides were detected in MS experiment number 1. The bottom vertical axis is the peptide or protein rank according to prevalence. The top vertical axis indicates the number of peptides or proteins detected in 1, 2, etc., up to all MS experiments. For example, 2057 peptides were detected in only 1 MS experiment, while 1587 peptides were detected in all 13 MS experiments. A vertical grey line within the plot indicates that peptide or protein was detected in that MS experiment. Thus, a peptide or protein detected in all 13 MS experiments would show as a vertical grey line over the entire plot. The red line indicates the number of peptides or proteins detected for a given number of MS experiments. For example, the first step indicates the 2057 peptides detected in only 1 MS experiment.

Supplementary Figure S3. Box and Whisker Plots

Pre- (left panel) and post-normalization (right panel) box and whisker plots showing the global peptide abundance distribution. The bottom, middle bar, and top of the boxes indicate the 25^{th}, 50^{th} and 75^{th} percentiles, respectively. The whiskers (vertical dashed lines) extend to 1.5 times the interquartile range (i.e., 1.5 * [50^{th} percentile – 25^{th} percentile]) or the maximum value, whichever is smaller. The x-axis indicates MS

experiment; within an experiment boxes are sorted by tag. The y-axis is peptide abundance on the log10 scale with labels indicating raw scale. Global shifts before normalization indicate the need for normalization, and the absence of these shifts after normalization indicates successful normalization.

Supplementary Figure S4: Co-efficient of Variation (CV) versus Mean
The vertical axis is the smoothed (via loess) within MS experiment CV. This is calculated as the standard deviation between the abundance values on the natural log scale for each MS/MS observation of a peptide.

Supplementary Figure S5. "Volcano Plots"
The horizontal axis indicates fold change on the log2 scale. The vertical axis indicates –log10(p-value). Horizontal lines indicate various levels of statistical significance. Points in the top right and left corners of the plots are the most statistically significant and have the greatest fold changes. Lines indicating 2-fold change are included for reference only.

Supplementary Figure S6
Results of IPA networks identified as significant hits in the analysis centered on NfKB complex

Supplementary Figure S7

Results of IPA networks identified as significant hits in the analysis centered on P38-MAPK complex

Supplementary Methods

Serum Sample preparation

Processing and handling of serum specimens followed HUPO/PPP standardization methods [1]. For serum processing, blood was collected in BD SST^{TM} 6.0 mL. vacutainers for serum processing on a single occasion. All specimens underwent centrifugation at 3,000 rpm for 10 minutes to generate serum and were stored at -70ºC within 30 to 45 minutes of sample collection from patients. After the initial centrifugation a protease inhibitor cocktail was added the ingredients of which included 10mL PBS (Invitrogen No. 14190300);1 tablet complete of mini, EDTA-free protease inhibitor(Roche No. 11 836 170 001); Sodium Vanadate Na3VO4 and PMSF (Sigma No. P7626-5G). No serum specimen underwent any freeze-thaw cycles prior to performing affinity depletion and preparation for isobaric mass tags for relative and absolute quantitation (iTRAQ) labeling.

Methods for Affinity depletion of high abundant proteins from human sera and protein digestion/iTRAQ labeling

HPLC buffers were obtained from the column manufacturer and the separation followed the manufacturer's recommended protocol with uv monitoring at 280nm. Each serum sample (40 µL) was diluted to 200 µL with HPLC Buffer A (Equilibrate/Load/Wash) and filtered through a 0.22 µm cellulose acetate spin filter (Agilent Technologies) prior to injection. Duplicate 95 µL injections were run for each sample, unbound (low abundance proteins) fractions and bound (high abundance

proteins) were collected as separate fractions. The unbound fractions for each sample were pooled, concentrated and buffer exchanged to 500mM triethylammonium bicarbonate pH 8.5 using BioMax 5K MWCO spin ultrafiltration cartridges (Millipore Corporation, Bedford, MA USA). The protein concentration for each of the depleted samples was determined by Bradford assay using BSA as the calibrant, and further quantified by running SDS-mini gel followed by ImageQuant Software (GE Healthcare).

Protein digestion and iTRAQ labeling

A total of 50 µg protein of each sample was denatured with 1% SDS, reduced with 5 mM tris-(2-carboxylethyl) phosphine at 60 ºC for 1 h and the cysteine residues were blocked with 8 mM methyl methanethiosulfonate (MMS) for 10 min at room temperature. Protein samples were digested with 4 µg of sequencing-grade-modified trypsin at 37 ºC for 16 h. Tryptic peptides from 4 different randomized (see statistical methods) samples were each labeled with iTRAQ reagents 114, 115, 116, and 117, respectively according to the manufacturer's protocols (document #4351918A and 4350831C downloaded from http://docs.appliedbiosystems.com/search.taf; Applied Biosystems, Foster City, CA). A total of 13 sets of iTRAQ 4-plex tagged samples were prepared and designed as shown in Figure 1 covering all 50 individually depleted samples; paired samples 100059.early and 100037.late were duplicated in mass spectrometry (MS) experiment 4 and MS experiment 5. After labeling, the four samples were pooled in each of the 13 MS experiments and subjected to cation exchange chromatography using an Applied Biosystems cation-exchange cartridge system. The pooled iTRAQ labeled tryptic peptides (~400 µL) samples were

evaporated completely in a SpeedVac concentrator and reconstituted with 1 mL of loading buffer (10 mM potassium phosphate pH 3.0, 25% ACN). The pH of the sample was adjusted to 3.0 with formic acid prior to cartridge separation. After conditioning of the SCX cartridge with loading buffer, the sample (~200 µg) was loaded and washed with an additional 2 mL of loading buffer. The peptides were eluted in one steps by 1 mL of loading buffer containing 500 mM KCl. Desalting of SCX fractions was carried out using solid phase extraction (SPE) on Sep-Pak® Cartridges (Waters, Milford, MA) and ready for the first dimensional LC fractionation via a high pH reverse phase chromatography as described below.

High pH reverse phase fractionation

High pH reverse phase fractionation was completed using a Dionex UltiMate3000 HPLC system with built-in micro fraction collection option in its autosampler and UV detection (Sunnyvale, CA). The iTRAQ tagged tryptic peptides were reconstituted in buffer A (20 mM ammonium formate pH 9.5 in water), and loaded onto a hybrid silica column XTerra MS C18 (2.1 mm × 150 mm) from Waters (Milford, MA). Buffer B consists of 20 mM ammonium formate pH 9.5 with 90% ACN in water. The high pH RP chromatography separation was performed at a flow rate of 200 µL/min using a gradient of 5% B for 3 min, 5-45% B in 30 min, and ramp up to 90% B in 5 min. During gradient elution, forty-eight fractions were collected at 1 min per fraction. On the basis of the UV trace at 214 nm, several fractions were pooled to yield a total of 19 high pH RP fractions which were dried and reconstituted in 2% ACN-0.5% formic acid for subsequent nanoLC-MS/MS analysis.

Nano-scale reverse phase chromatography and tandem mass spectrometry (nanoLC-MS/MS).

The Orbitrap was interfaced with an UltiMate3000 MDLC system (Dionex, Sunnyvale, CA). An aliquot of high pH RP peptide fractions (2.0-10.0 µL) was injected onto a PepMap C18 trap column (5 µm, 300 µm × 5 mm, Dionex) at 20 µL/min flow rate for rapid sample loading and then separated on a PepMap C-18 RP nano column (3 µm, 75µm x 15cm), and eluted using a 60 min gradient of 5% to 35% acetonitrile (ACN) in 0.1% formic acid at 300 nL/min., followed by a 3-min ramping to 95% ACN-0.1%FA and a 5-min holding at 95% ACN-0.1%FA. The column was re-equilibrated with 2% ACN-0.1%FA for 20 min prior to the next run. The eluted peptides were detected by Orbitrap through a nano ion source containing a 10-µm analyte emitter (NewObjective, Woburn, MA). The Orbitrap Velos was operated in positive ion mode with nano spray voltage set at 1.5 kV and source temperature at 225 ℃. Either internal calibration using the background ion signal at m/z (mass to charge ratio) 445.120025 as a lock mass or external calibration for FT mass analyzer was performed. The MS and MS/MS data were acquired at data-dependent acquisition (DDA) mode using FT mass analyzer for one survey MS scan for precursor ions followed by ten data-dependent HCD-MS/MS scans for precursor peptides with multiple charged ions above a threshold ion count of 5000 with normalized collision energy of 45%. MS survey scans at a resolution of 30,000 (fwhm at *m/z* 400), for the mass range of m/z 400-1400, and MS/MS scans at 7,500 resolution for the mass range m/z 100-2000. All data are acquired under Xcalibur 2.1 operation software (Thermo-Fisher Scientific).

Mass spectrometry data processing, protein identification and analysis

All MS and MS/MS raw spectra from iTRAQ experiments were processed using Proteome Discoverer 1.1 (PD1.1, Thermo) and the spectra from each DDA file were output as an MGF file for subsequent database search using in-house license Mascot (Matrix Science, London, UK; version 2.2), Sequest (Thermo Fisher Scientific, San Jose, CA, USA; version 27, rev. 12), X!Tandem (The GPM, thegpm.org; version 2009.09.15.3) and Scaffold (version 3.0, Proteome software, Inc. Portland, OR). The *SwissProt Human* 2010_05 database containing 20,277 sequence entries downloaded from Uniprot (http://uniprot.org) was used for database query. The default search settings used for 4-plex iTRAQ quantitative processing and protein identification in Scaffold were: one mis-cleavage for full trypsin with fixed MMTS modification of cysteine, fixed 4-plex iTRAQ modifications on lysine and N-terminal amines and variable modifications of methionine oxidation and 4-plex iTRAQ on Tyrosine. The peptide mass tolerance and fragment mass tolerance values were 10 ppm and 50 mDa, respectively. To estimate the false discovery rate (FDR) for a measure of identification certainty in each replicate set, a decoy database search was performed by augmenting the database with a set of reverse sequences. A filter was applied using the Scaffold software, requiring peptide and protein probability at >95% and at least 2 peptides per protein. The estimated FDR for these thresholds was <1%. Intensities of the reporter ions from iTRAQ tags upon fragmentation were used for quantification.

Statistical Methods for Proteomic analyses

In order to ensure that sample groups of interest were well balanced over potential experimental effects such as time, iTRAQ labeling tag and MS experiment, randomized block design principles (23) were used to create two randomization schemes to determine: 1) the order of sample processing, and 2) allocation to the iTRAQ 4-plex experiments. The paired pre/post ADT samples were forced to be in the same MS experiment to minimize variability for the paired comparison. One matched pair from the early/late group was replicated for quality control purposes. Thus, a total of 52 samples from 35 subjects were assayed in 13 MS experiments with the 4-plex iTRAQ system (Supplementary Figure **S1**). In addition to the laboratory quality-control assessments, global quality and bias were assessed graphically [ref: PMID: 19712457; ref: PMID: 23176383]. Box-and-whisker plots were used to assess global shifts in the distribution of abundance across samples and experiments. Plots of Coefficients of Variation (CV) versus the mean were used to assess the mean-variance relationship, which was not found to be a function of the mean in these data (Supplementary Figure **S4:** "CV Versus Mean"). Finally, coverage plots were used to assess the number of identified proteins and peptides across iTRAQ experiments. Based upon these quality control plots, two MS experiments were repeated due to incorrect machine settings. Subsequently, all experiments and samples were deemed of good quality and all 52 specimens were retained for further analysis.

The natural log peptide abundance values of the 52 samples were normalized using a linear model to remove global experimental effects (24). The residual values from the model fit were used as normalized values and kept for further analysis. Box-and-whisker plots were used to confirm that normalization was successful (Supplementary Figure

S3: "Box and Whisker" Plot). Per-protein linear mixed effects models with contrast statements were used to assess differential abundance using the normalized data. Three comparisons were conducted: pre-ADT vs. post-ADT (15 paired samples), early vs. late failure (ten samples each), and post-ADT vs. any failure (early plus late failure cohorts). A random effect for subject was used in order to account for correlation between paired samples for the pre-ADT vs. post-ADT analysis. Only the first replicates from the two duplicated samples were included in differential abundance analyses. Statistical significance was assessed using a False Discovery Rate (FDR), with a threshold of 0.20 for each comparison (25). Volcano plots for each comparison were created, with fold changes being reported on the log2 scale for ease of interpretation. Differentially abundant proteins/peptides in each of these comparisons are presented in Venn diagrams illustrating the overlap of candidates in these comparisons.

Pathway enrichment methods

Experimentally observed candidates with differential abundance in these comparisons were then grouped using pathway network analyses in order to identify key genes and their translational serum products for prospective validation. Using the Ingenuity Pathway Analysis tool, genes that correspond to candidate proteins were mapped to relevant pathways and networks based on their functional annotation and known molecular interactions in Ingenuity Knowledge Base (IKB). IKB is a manually curated repository containing thousands of biological interactions and functional annotations extracted from published literature [2]. A molecular network of direct or indirect physical, transcriptional, and enzymatic interactions between mammalian orthologs was

computed from IKB. The corresponding gene names were uploaded into IPA along with the gene identifiers and corresponding protein fold change values. The genes were then overlaid onto a global molecular network in IKB. In the network analysis, networks of these genes are then algorithmically generated based on their connectivity. Two genes are considered to be connected if there is a path in the network between them. Highly-interconnected networks likely represent significant biological function. IPA constructs networks that optimize for both interconnectivity and number of focus genes (defined based on triangular connectivity) under the constraint of maximal network size (default is 35 genes).

References:

1. Rai, A.J., et al., *HUPO Plasma Proteome Project specimen collection and handling: towards the standardization of parameters for plasma proteome samples.* Proteomics, 2005. **5**(13): p. 3262-77.
2. Calvano, S.E., et al., *A network-based analysis of systemic inflammation in humans.* Nature, 2005. **437**(7061): p. 1032-7.

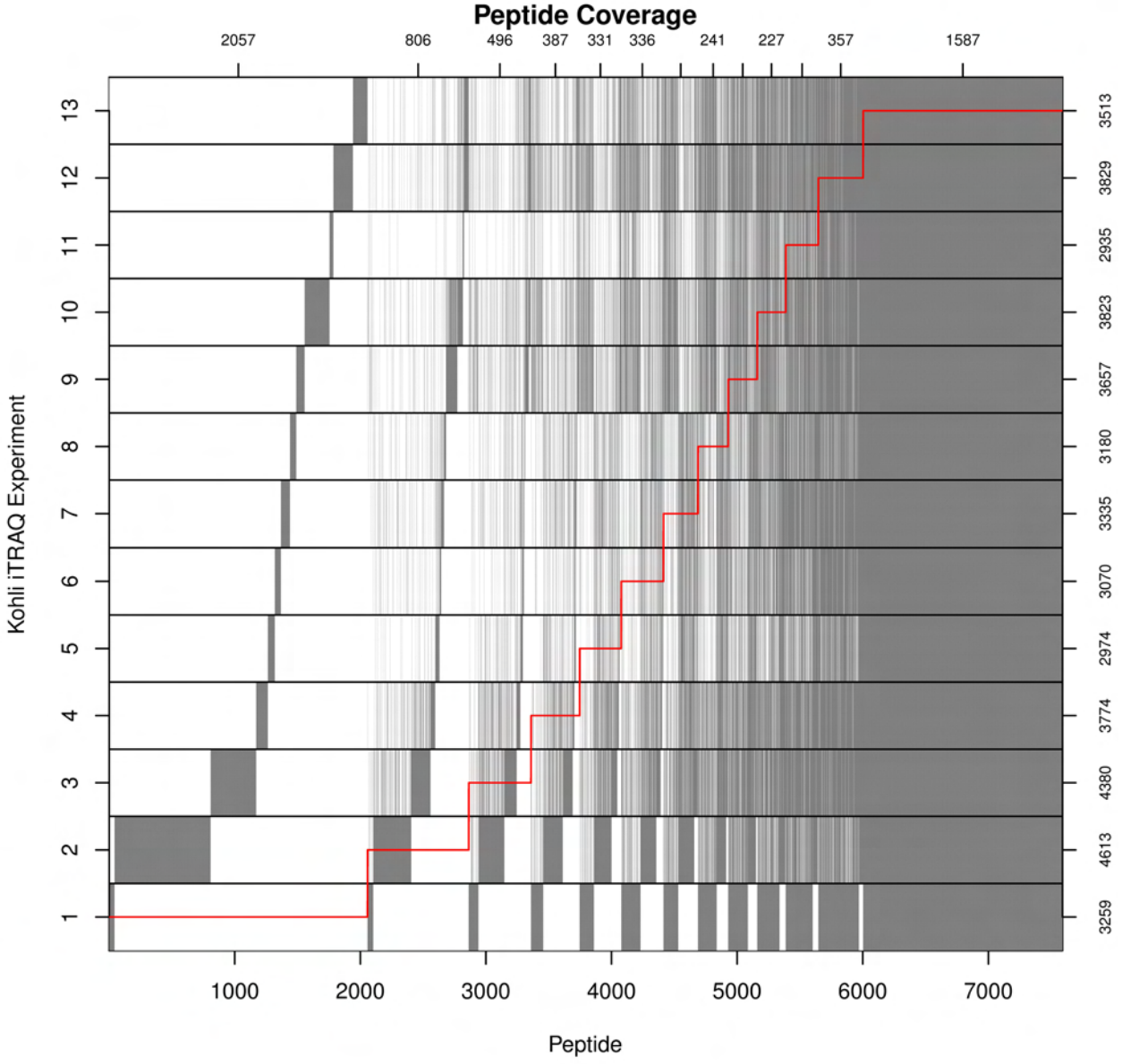
S1: iTRAQ Experimental Design

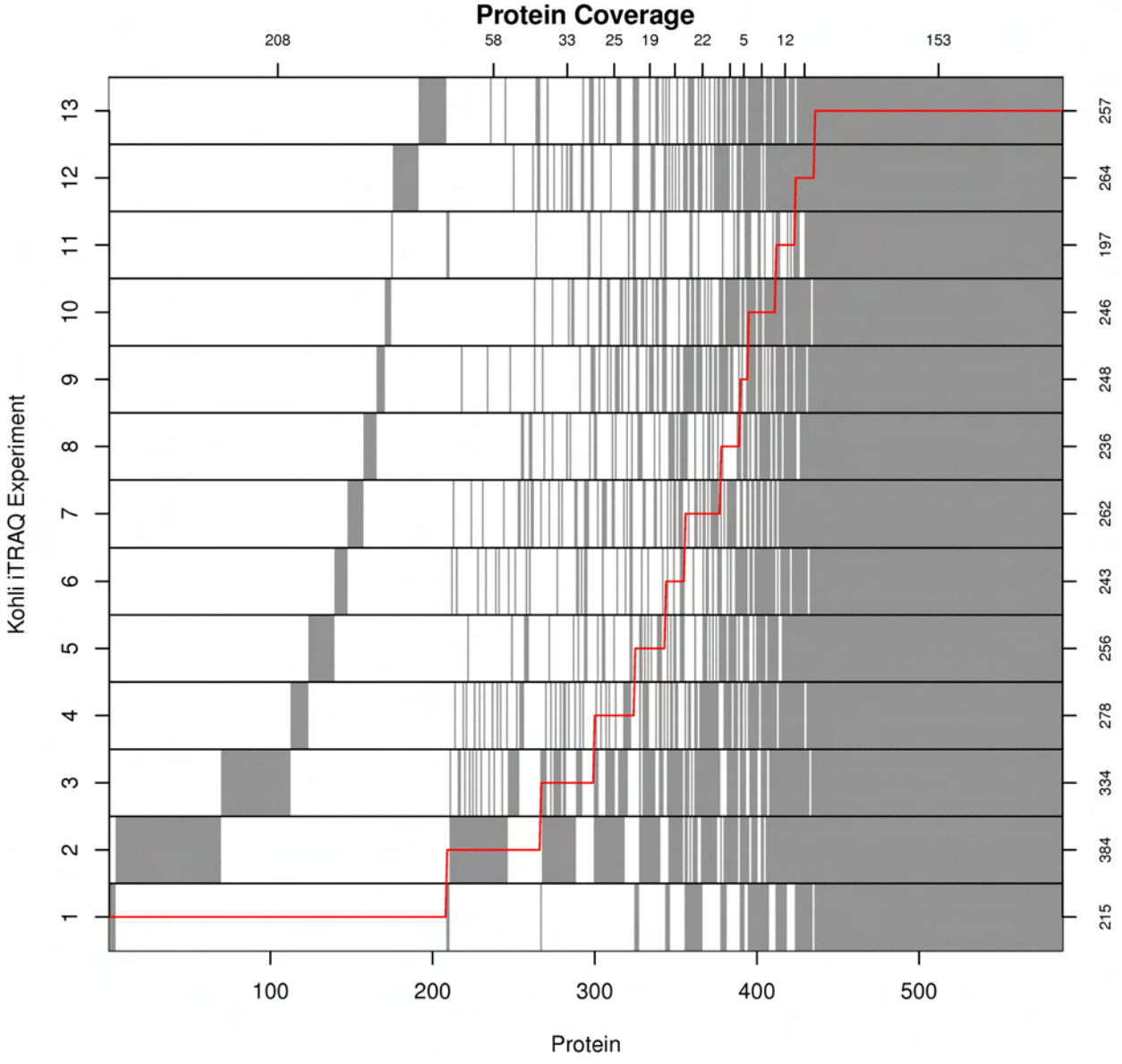
iTRAQ Experiment Number	Channel 114	Channel 115	Channel 116	Channel 117
1	100270*.pre	100011.late	100263.early	100270.post
2	100083.post	100285.early	100176.late	100083.pre
3	100033.early	100245.post	100245.pre	100271.late
4	100278.post	100059.early	100037.late	100278.pre
5	100037.late	100043.pre	100043.post	100059.early
6	100006.late	100149.early	100154.post	100154.pre
7	100169.pre	100169.post	100140.early	100175.late
8	100282.pre	100163.pre	100282.post	100163.post
9	100155.early	100005.late	100240.pre	100240.post
10	100096.post	100118.post	100118.pre	100096.pre
11	100063.post	100063.pre	100021.late	100016.early
12	100025.early	100283.pre	100152.late	100283.post
13	100204.post	100116.late	100204.pre	100090.early

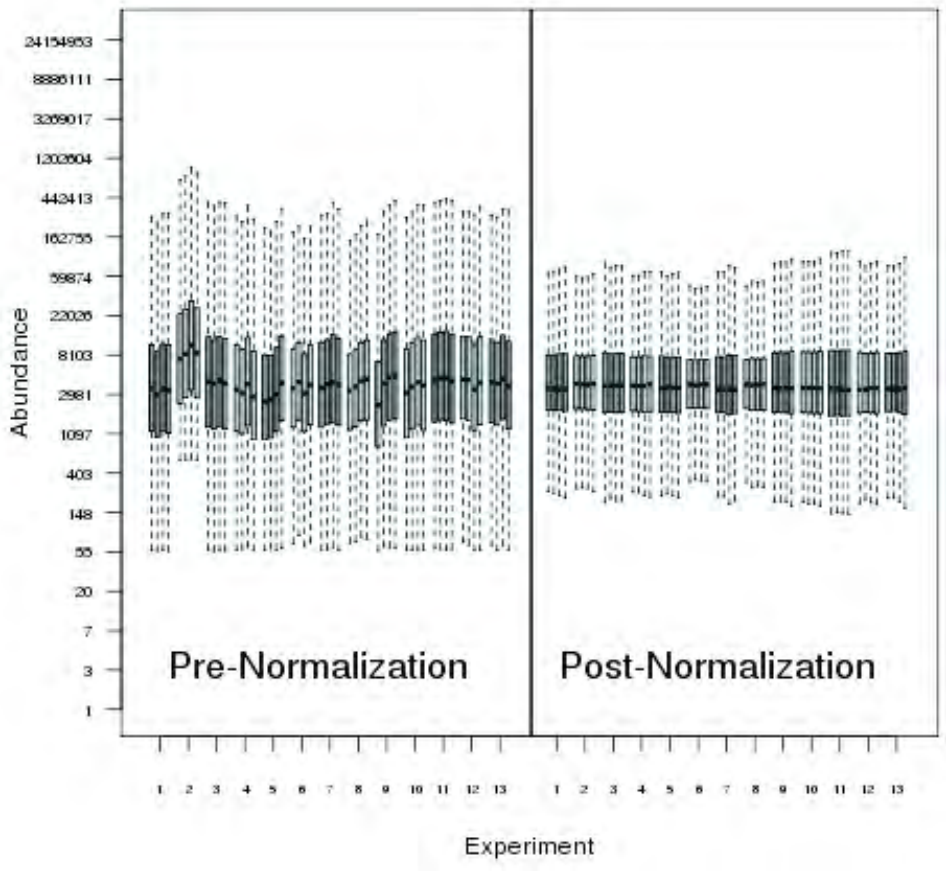
* Specimen IDs annonymized from 100005 to 100278

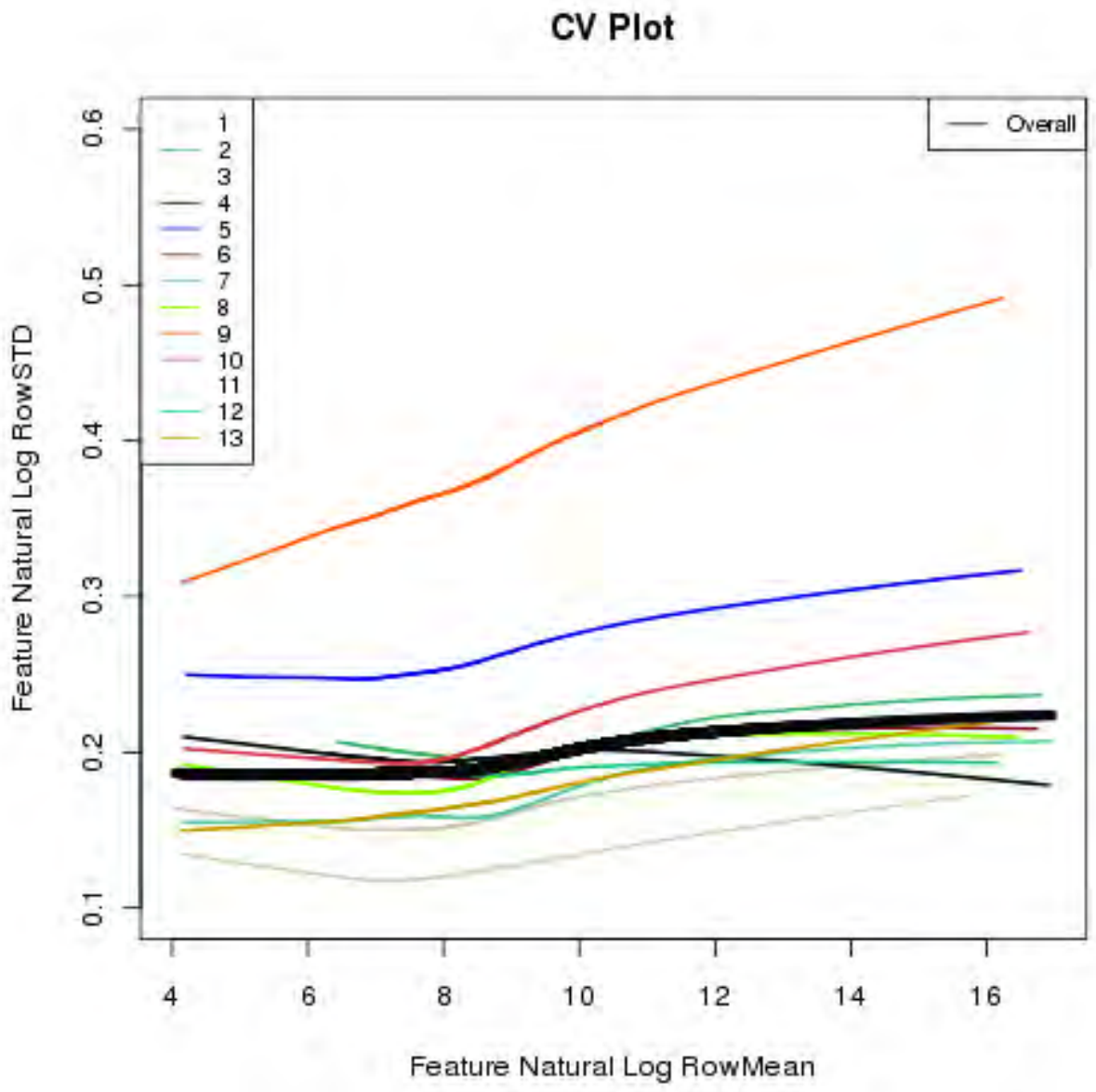
Color Key

- Pre-ADT
- Post-ADT
- Early ADT Failure
- Late ADT Failure

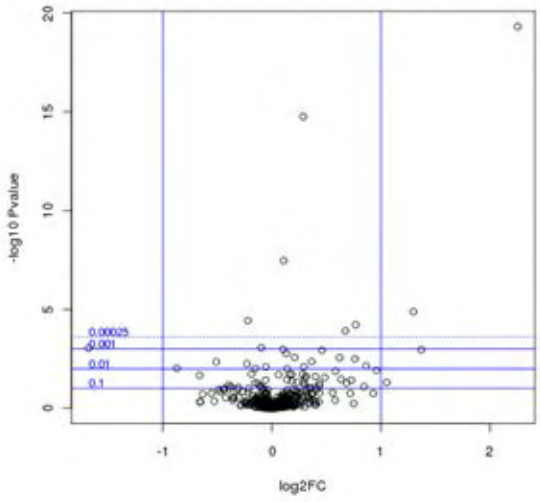

Peptide Coverage
2057
806
496
387
331
336
241
227
357
1587
Kohli iTRAQ Experiment
13
12
11
10
9
8
7
6
5
4
3
2
1
3513
3829
2935
3823
3657
3180
3335
3070
2974
3774
4380
4613
3259
1000
2000
3000
4000
5000
6000
7000
Peptide


Protein Coverage
208
58
33
25
19
22
5
12
153
Kohli iTRAQ Experiment
13
12
11
10
9
8
7
6
5
4
3
2
1
257
264
197
246
248
236
262
243
256
278
334
384
215
100
200
300
400
500
Protein

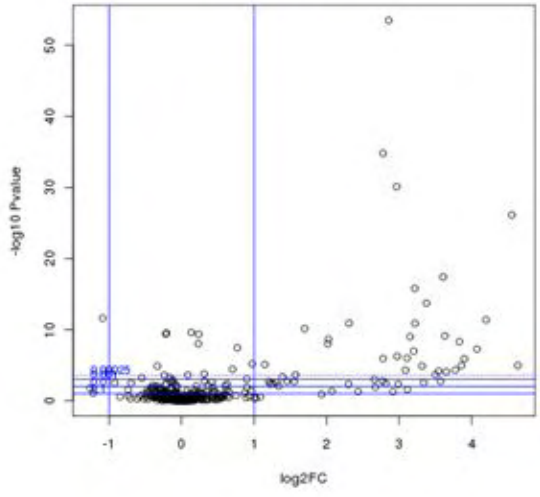

Abundance
24154953
8886111
3269017
1202604
442413
162755
59874
22026
8103
2981
1097
403
148
55
20
7
3
1
Pre-Normalization
Post-Normalization
1 2 3 4 5 6 7 8 9 10 11 12 13
1 2 3 4 5 6 7 8 9 10 11 12 13
Experiment


CV Plot
Overall
1
2
3
4
5
6
7
8
9
10
11
12
13
Feature Natural Log RowSTD
0.1
0.2
0.3
0.4
0.5
0.6
4
6
8
10
12
14
16
Feature Natural Log RowMean

Analysis 1 - Pre vs Post Paired



Analysis 2 - Early vs Late



Analysis 3 - Post vs Early/Late