Model Report

Project - XXX-YYY

Date - 3/2/2021

Project leader Customer: ..

Project leader PamGene: …

Customer

Name

Address

Country

PamGene International B.V.

Wolvenhoek 10, ’s-Hertogenbosch

The Netherlands

PamGene Model Report

3/2/2021

Contents

1. Summary

This report represents a model study report of PamGene data. It contains data from an actual experiment; however, the information has been anonymized and replicate numbers were scrambled for the purpose of this report. The arrays of the treatments and controls are as in the original data.

We investigated the kinases implicated in treatments T1, T2, T3 compared to untreated controls by measuring the kinase activity of the CELL LINE A samples using PamGene's PTK (Protein Tyrosine Kinase) and STK (Serine/Threonine Kinase) assays. In the context of this report, T1-T3 refer to 3 treatment timepoints.

Based on PamGene's QC criteria of signal strength, number of peptides and the variation, we consider the data quality of this study to be "good" (i.e., green flag, ●).

For a definition of these flags and how they are applied to the data please refer to the separate Supplementary document (see supplement.docx).

In summary, for both PTK and STK assays, there is a significant stimulation effect seen in the 3 treatments (Treatments 1, 2, 3) compared to controls. Treatment 2 shows the most significant effect.

The main result is a list of kinases (PTK and STK) for each comparison, ranked according to their putative involvement in the differential effect between treatment and control (Kinase interpretation (Differential)).

This information can be used for further biological interpretation, hypothesis generation and validation.

2. Study Aims

Aim

This study aims to determine kinases implicated in treatments T1, T2, T3 compared to untreated controls.

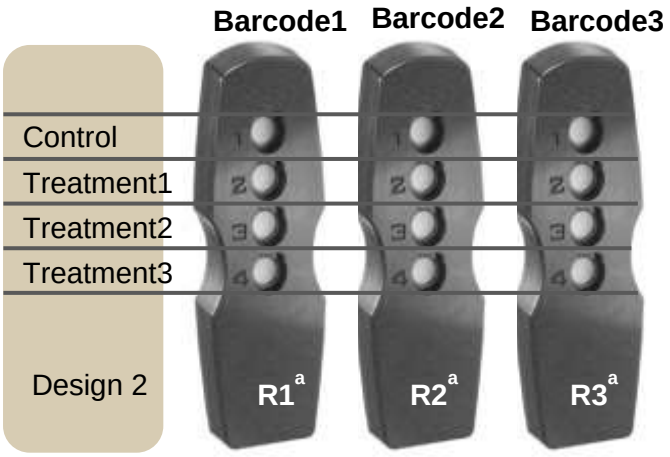
Background

CELL LINE A was treated in vitro with 3 treatments conditions, with 3 biological replicates per condition.

Typically, the experimental design/set-up for this study type follows the scheme shown below (Design 2 of our Experimental Design).

Design 2

Comparing multiple treatment conditions in the same modelb (E.g.: in vitro treatment of cell cultures off chip or ex vivo spike-in of cell lysates on chip).
Comparing different modelsb with different genotypes or phenotypes (E.g: wild-type and mutants, or diseased and normal).
Time or concentration course of the same treatment in the same modelb (E.g.: in vitro treatment of cell cultures off chip).



[a] R1, R2, R3 - Refer to Replicates. These can be Technical replicates (1 lysate on 3 arrays) or Biological replicates (3 lysates on 3 arrays)

[b] Models - Refer to the model system used, including but not limited to, cell lines, tumour tissue, PDX, mouse models etc.

Control and Treatments - Refer to the conditions in the study, see Applications Box for examples

Note - Sample annotations should include the "treatment" details and the generic annotation as shown

PamGene kinase activity profiles are measured for the following samples:

Treatment: (3 Time points; Biological Replicates 1, 2, 3 on PamChip Barcodes 1, 2, 3)

- Control
- Treatment Time point 1 (T1)
- Treatment Time point 2 (T2)
- Treatment Time point 3 (T3)

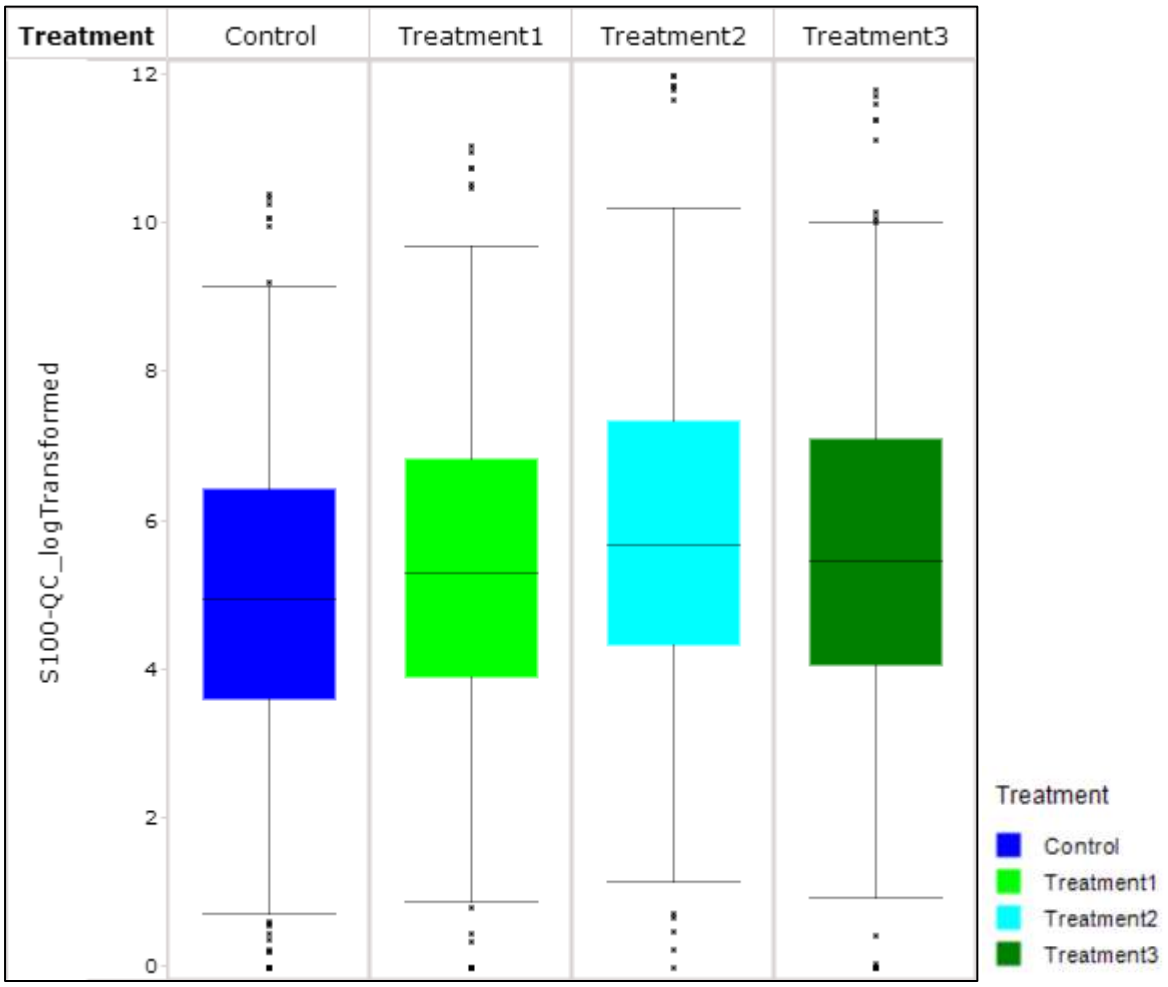
3. Results

3.1. Overview of measurements for all samples

3.1.1. Box Plot overview of signals

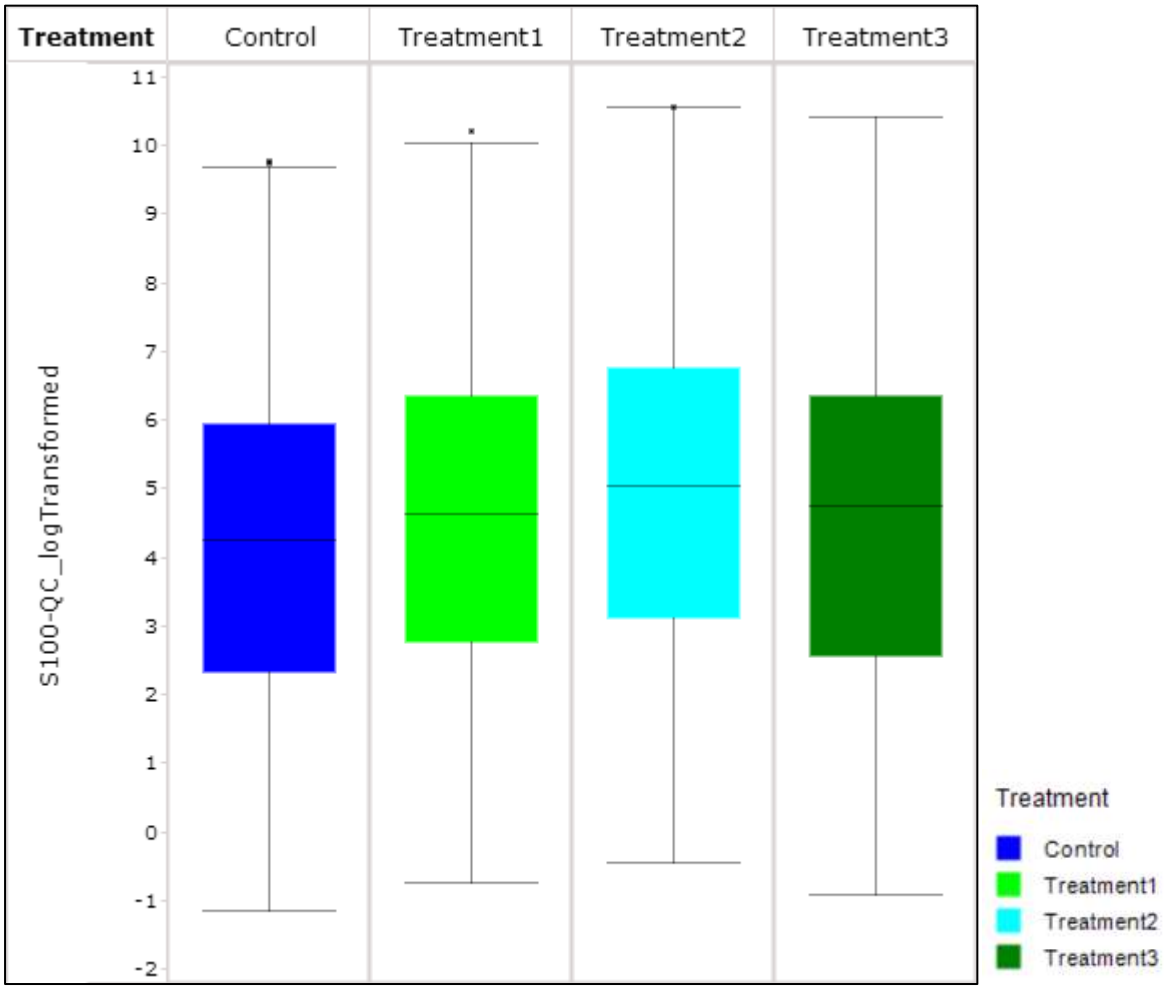
Log2 Signals of phosphorylated peptides which have passed the QC (See Supplement) are depicted in Box plots, to visualize overall sample variance and group differences for the 4 conditions (C, T1, T2, T3) with 3 biological replicates per condition.

Box Plot (PTK) Log2 signals



PTK - Log2 signals (Treatments and Control) (Columns) for each treatment condition (Box plot).

Box Plot (STK) Log2 signals



STK - Log2 signals (Treatments and Control) (Columns) for each treatment condition (Box plot).

3.2. Differential Analysis Results: “Treatment” versus “Control”

For Phosphosite statistical analysis (differential): The following statistical test is used to generate a list of differentially significant phosphorylated peptides:

- A type of ANOVA-PostHoc Test for multiple treatments versus Control (MTvC)

For Kinase interpretation (differential): In order to generate a ranked list of putative kinases responsible for differences in the peptide phosphorylation, PamGene’s in-house method called Upstream Kinase Analysis (UKA) is used (for details refer to the section “Kinase interpretation (Differential)”).

For Pathway interpretation (differential): In order to generate a ranked list of possible canonical pathways (and networks) responsible for differences in the peptide phosphorylation, there are many open-source tools to perform this. On request, this report may be supplied.

3.2.1. Phosphosite statistical analysis (differential), MTvC

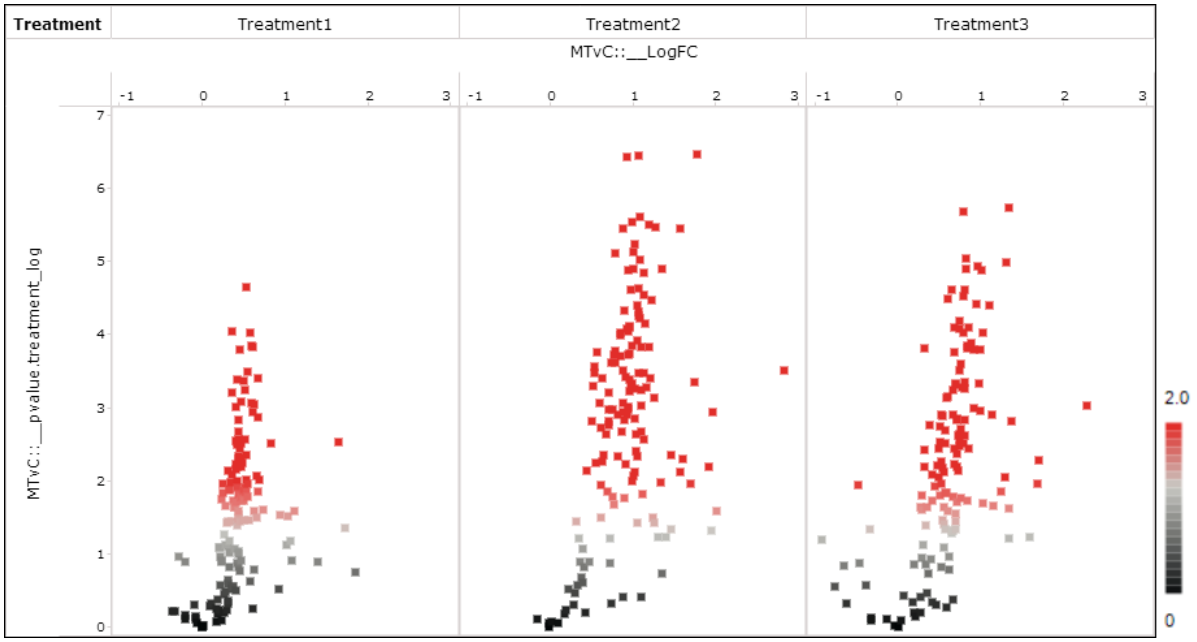
A Multiple Treatment versus Control (MTvC) tool is used to generate a list of differentially phosphorylated peptides.

This tool performs an ANOVA-like test for significant effects anywhere between the treatments (including control) and for significant effect between all separate Treatment vs. Control pairs.

The following plots are generated:

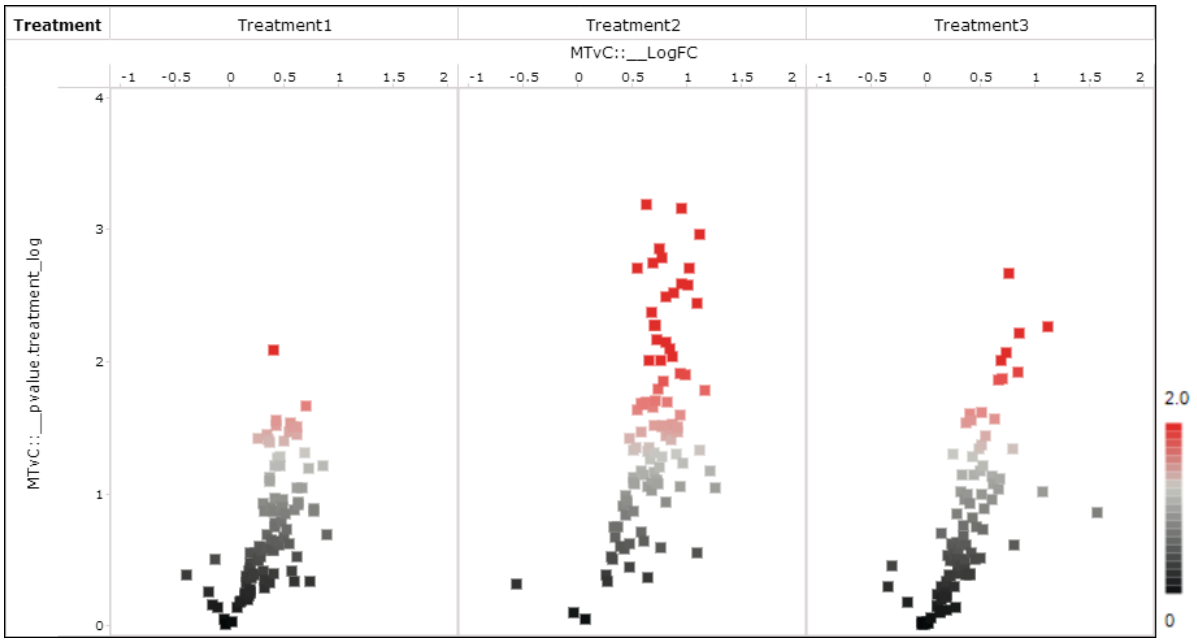
- Volcano plot, useful for a treatment effect overview. A volcano plot is used to give an overview of size, direction and significance.
- Peptide cluster view, useful for group peptide profiles and identification of trends in the data series.
- Significant peptides LFC heatmap, useful for the identification of the important peptides

Volcano (PTK) Treatment Time points v Control



The volcano plot visualizes the result of the tests by plotting - for each test - the effect size (x-axis, LFC or delta; >0 indicates activation and <0 indicates inhibition relative to Control) versus significance (y-axis, -log10(pvalue)) of the test. Red spots are peptides that show significant difference compared to control (p<0.05).

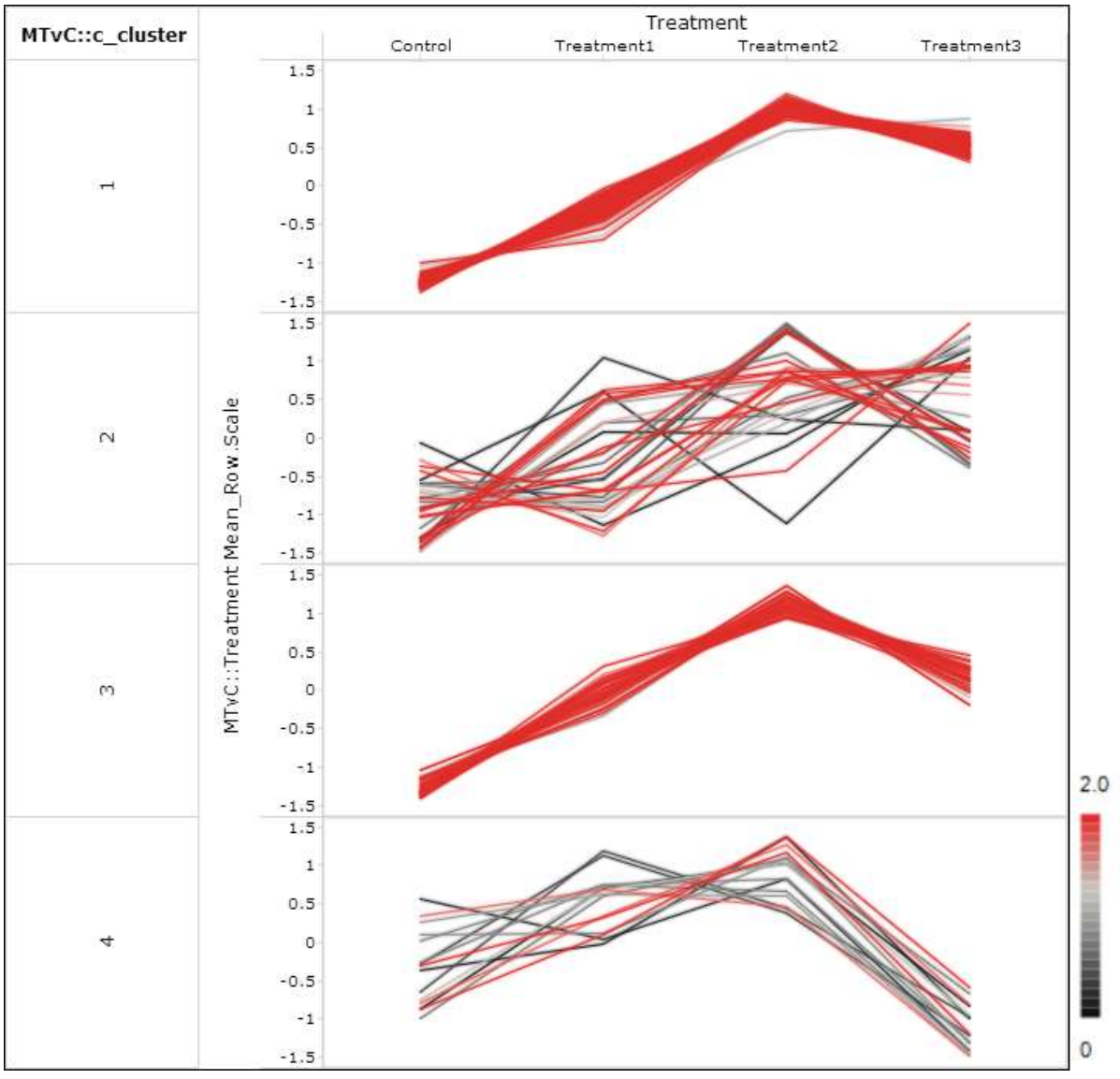
Volcano (STK) Treatment Time points v Control



The volcano plot visualizes the result of the tests by plotting - for each test - the effect size (x-axis, LFC or delta; >0 indicates activation and <0 indicates inhibition relative to Control) versus significance (y-axis, -log10(pvalue)) of the test. Red spots are peptides that show significant difference compared to control (p<0.05).

Peptide cluster (PTK) Treatment Time points v Control

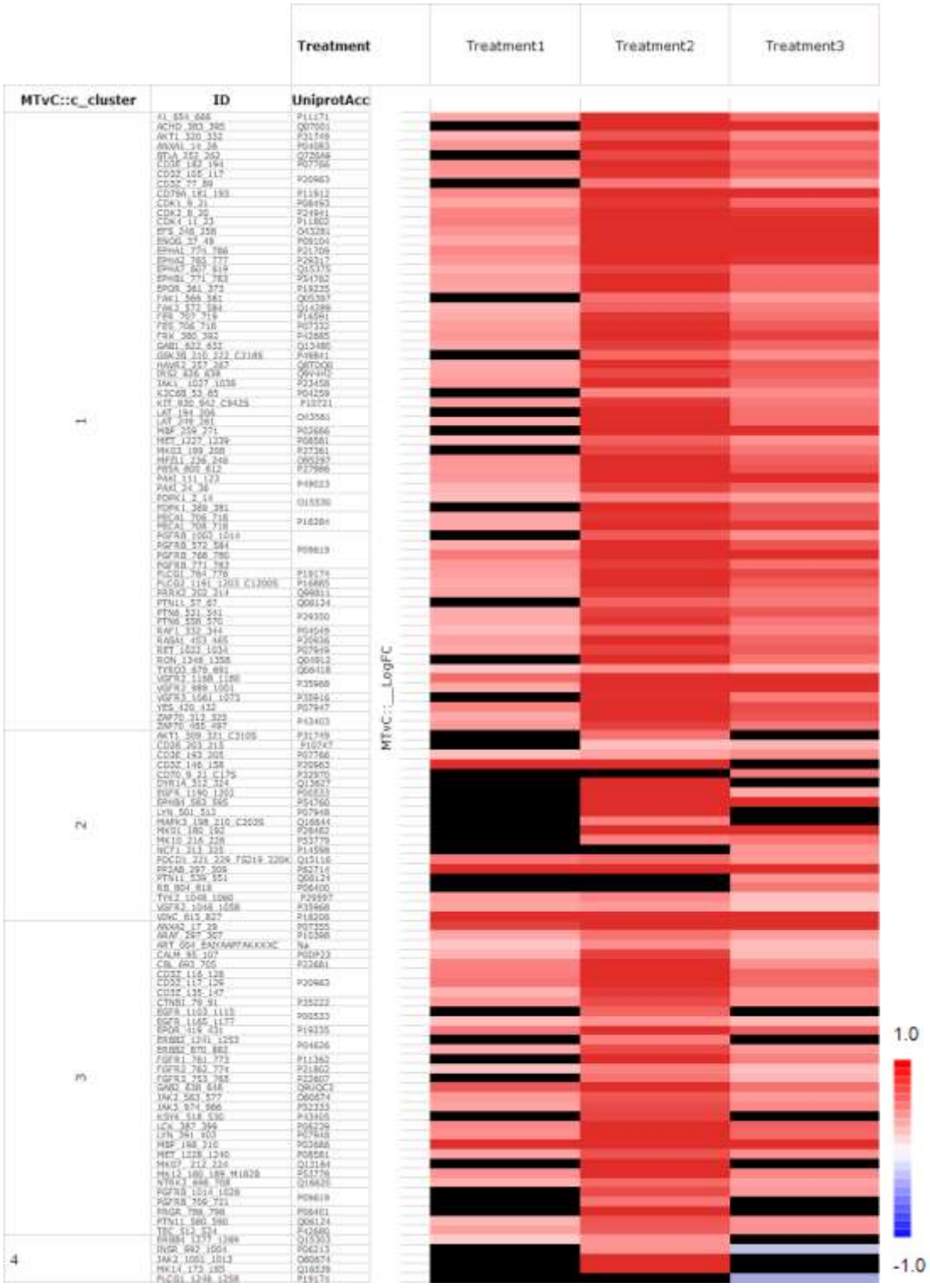
The peptide cluster plot is used to identify the different patterns of signal as a function of treatment. Each cluster contains peptides showing a similar response to treatment and may indicate similar biological function. The most relevant clusters are likely those that contain many red lines (see legend).



Line plots representing the response of peptides to treatment, clustered using a model-based algorithm. The signals are scaled per row. The y-axis is the z-score of each peptide across all treatments. The lines are colored according to the significance of any one of treatment effects (using ANOVA): a red line indicates the corresponding peptide with a significant treatment effect, black denotes not significant.

Significant peptides LFC (PTK) Treatment Time points v Control

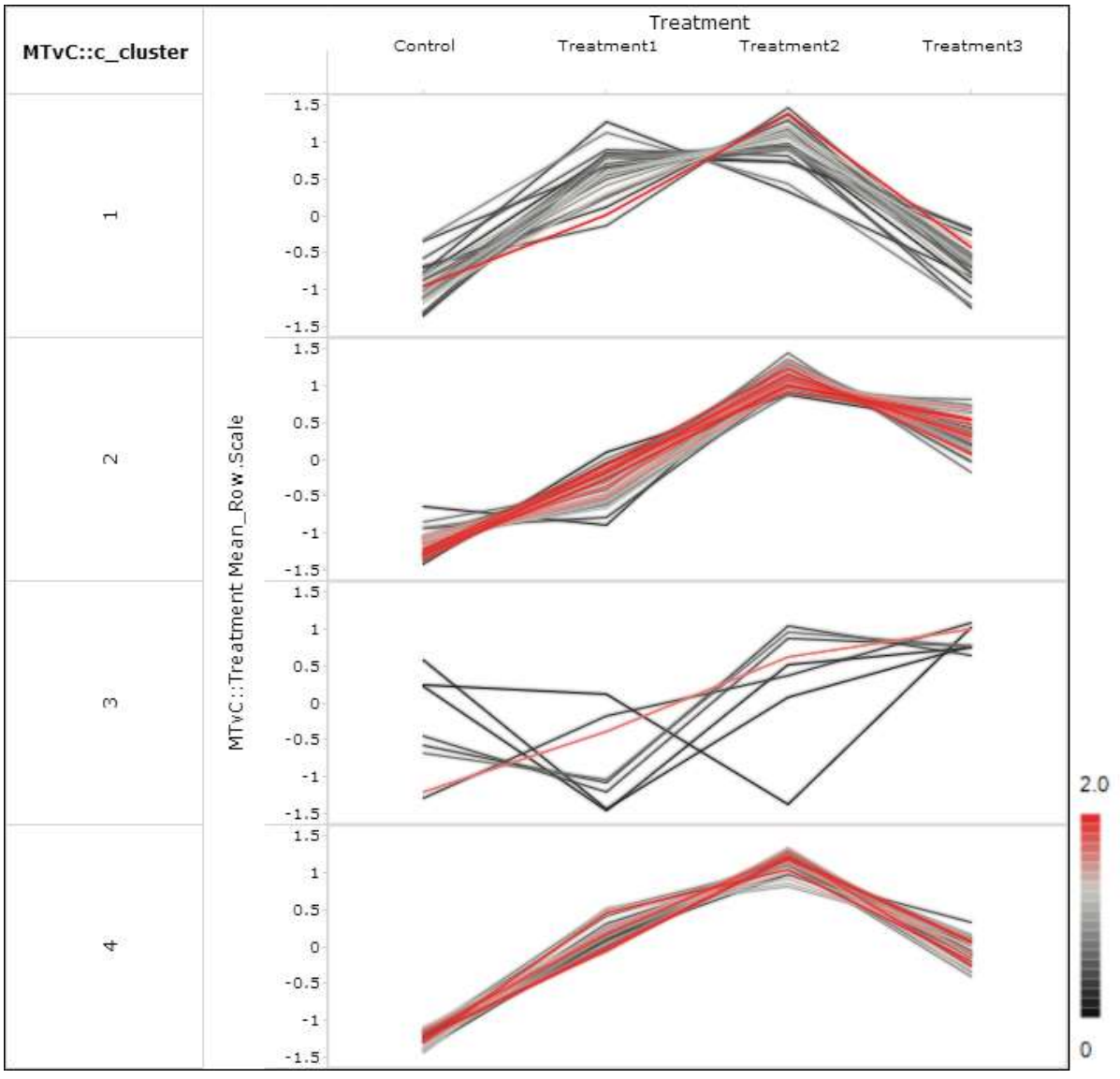
A list of peptides which are significantly differentially phosphorylated is visualized using a log fold change (LFC) heatmap.



The LFC heatmap for significant peptides shows the values of significantly differentially phosphorylated peptides ($p<0.05$, treatment vs. control). The treatment effects are log-ratios (log2(treatment)-log2(control)). Black color are peptides which did not pass the significance threshold.

Peptide cluster (STK) Treatment Time points v Control

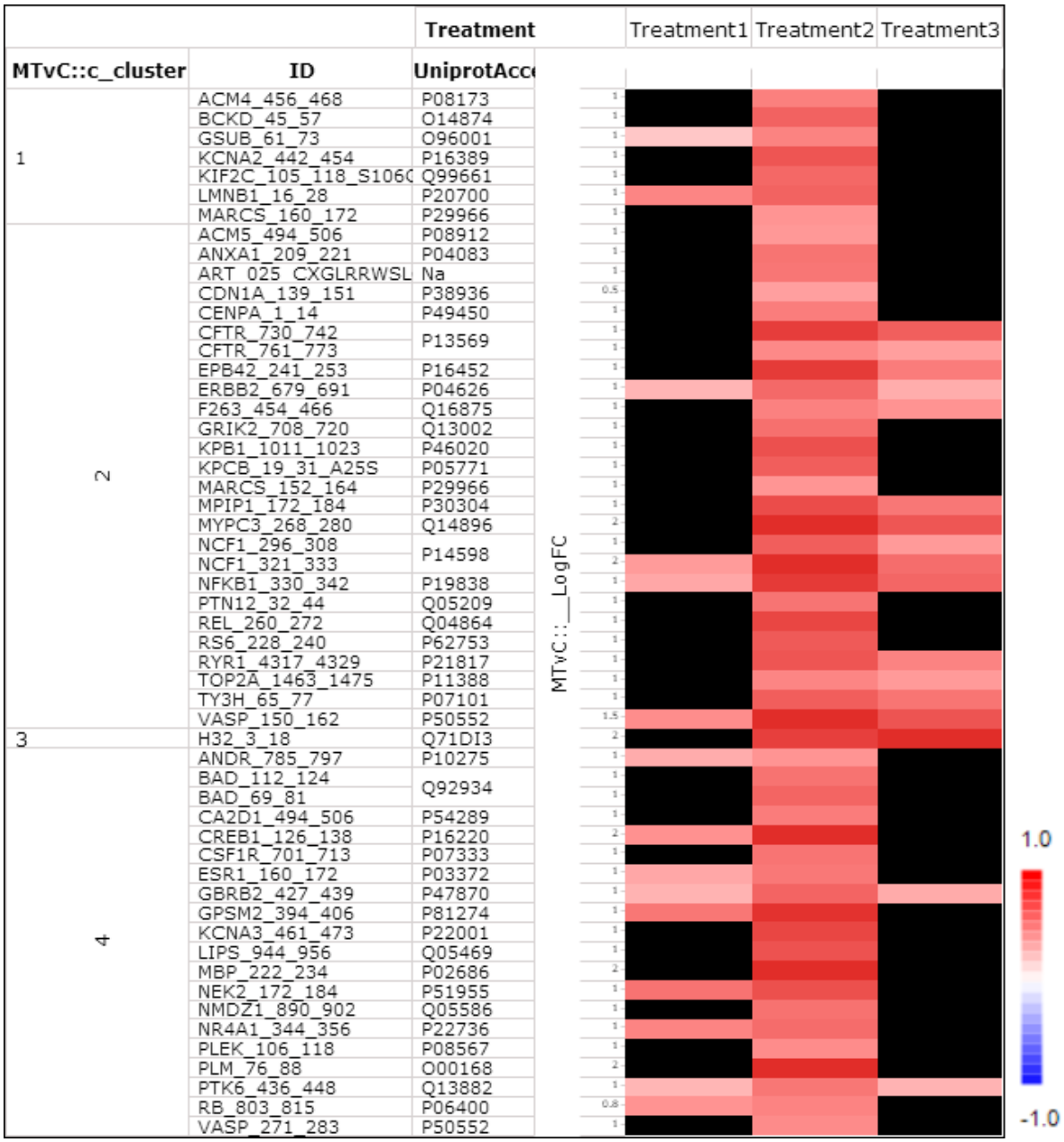
The peptide cluster plot is used to identify the different patterns of signal as a function of treatment. Each cluster contains peptides showing a similar response to treatment and may indicate similar biological function. The most relevant clusters are likely those that contain many red lines (see legend).



Line plots representing the response of peptides to treatment, clustered using a model-based algorithm. The signals are scaled per row. The y-axis is the z-score of each peptide across all treatments. The lines are colored according to the significance of any one of treatment effects (using ANOVA): a red line indicates the corresponding peptide with a significant treatment effect, black denotes not significant.

Significant peptides LFC (STK) Treatment Time points v Control

A list of peptides which are significantly differentially phosphorylated is visualized using a log fold change (LFC) heatmap.



		Treatment	Treatment1	Treatment2	Treatment3
MTvC::c_cluster	ID	UniprotAcc			
1	ACM4_456_468	P08173			
	BCKD_45_57	O14874			
	GSUB_61_73	O96001			
	KCNA2_442_454	P16389			
	KIF2C_105_118_S106(	Q99661			
	LMNB1_16_28	P20700			
	MARCS_160_172	P29966			
2	ACM5_494_506	P08912			
	ANXA1_209_221	P04083			
	ART_025_CXGLRRWSL	Na			
	CDN1A_139_151	P38936			
	CENPA_1_14	P49450			
	CFTR_730_742	P13569			
	CFTR_761_773				
	EPB42_241_253	P16452			
	ERBB2_679_691	P04626			
	F263_454_466	Q16875			
	GRIK2_708_720	Q13002			
	KPB1_1011_1023	P46020			
	KPCB_19_31_A25S	P05771			
	MARCS_152_164	P29966			
	MPIP1_172_184	P30304			
	MYPC3_268_280	Q14896			
	NCF1_296_308	P14598			
	NCF1_321_333				
	NFKB1_330_342	P19838			
	PTN12_32_44	Q05209			
	REL_260_272	Q04864			
	RS6_228_240	P62753			
	RYR1_4317_4329	P21817			
	TOP2A_1463_1475	P11388			
	TY3H_65_77	P07101			
	VASP_150_162	P50552			
3	H32_3_18	Q71DI3			
4	ANDR_785_797	P10275			
	BAD_112_124	Q92934			
	BAD_69_81				
	CA2D1_494_506	P54289			
	CREB1_126_138	P16220			
	CSF1R_701_713	P07333			
	ESR1_160_172	P03372			
	GBRB2_427_439	P47870			
	GPSM2_394_406	P81274			
	KCNA3_461_473	P22001			
	LIPS_944_956	Q05469			
	MBP_222_234	P02686			
	NEK2_172_184	P51955			
	NMDZ1_890_902	Q05586			
	NR4A1_344_356	P22736			
	PLEK_106_118	P08567			
	PLM_76_88	O00168			
	PTK6_436_448	Q13882			
	RB_803_815	P06400			
	VASP_271_283	P50552			

The LFC heatmap for significant peptides shows the values of significantly differentially phosphorylated peptides ($p<0.05$, treatment vs. control). The treatment effects are log-ratios (log2(treatment)-log2(control)). Black color are peptides which did not pass the significance threshold.

3.2.2. Kinase interpretation (Differential)

An Upstream Kinase Analysis (UKA) tool is used to generate a putative list of kinases responsible for phosphorylating the phosphosites on the PamChip. The tool is integrated into PamGene's proprietary data analysis software, BioNavigator v6.3. It is an interpretation and is highly dependent on the contents of the underlying phosphorylation databases.

The following plots are generated:

- Kinome Tree - useful to group the kinases into sequence families (plotted using a free Tool CORAL: http://phanstiel-lab.med.unc.edu/CORAL/. Additional information and tutorial are available on request.
- Kinase Score Table - useful to identify the top kinases
- Kinase Score Plot - useful to identify the specificity of kinases (can be additionally provided on request).

The tree and tables and plots show results from the Upstream Kinase Analysis (UKA) tool, a functional scoring tool which rank-orders the top kinases differential between the two groups. The ranking factor is the Final (Median) Kinase Score. This score is based on a combined sensitivity score (difference between "Treatment" and "Control" groups) and specificity score (for a set of peptides to kinase relationship derived from current databases). The Kinase statistic indicates the group differences, with effect size (values) and direction (+ or >0 = activation; - or 0 = inhibition).

This is an experimental tool and the results should be regarded as speculative for generating hypothesis, for target identification, pathway elucidation, biomarker discovery and other applications. The selected kinases need to be further validated using different approaches. Whether they make sense in the biological context should be further considered.

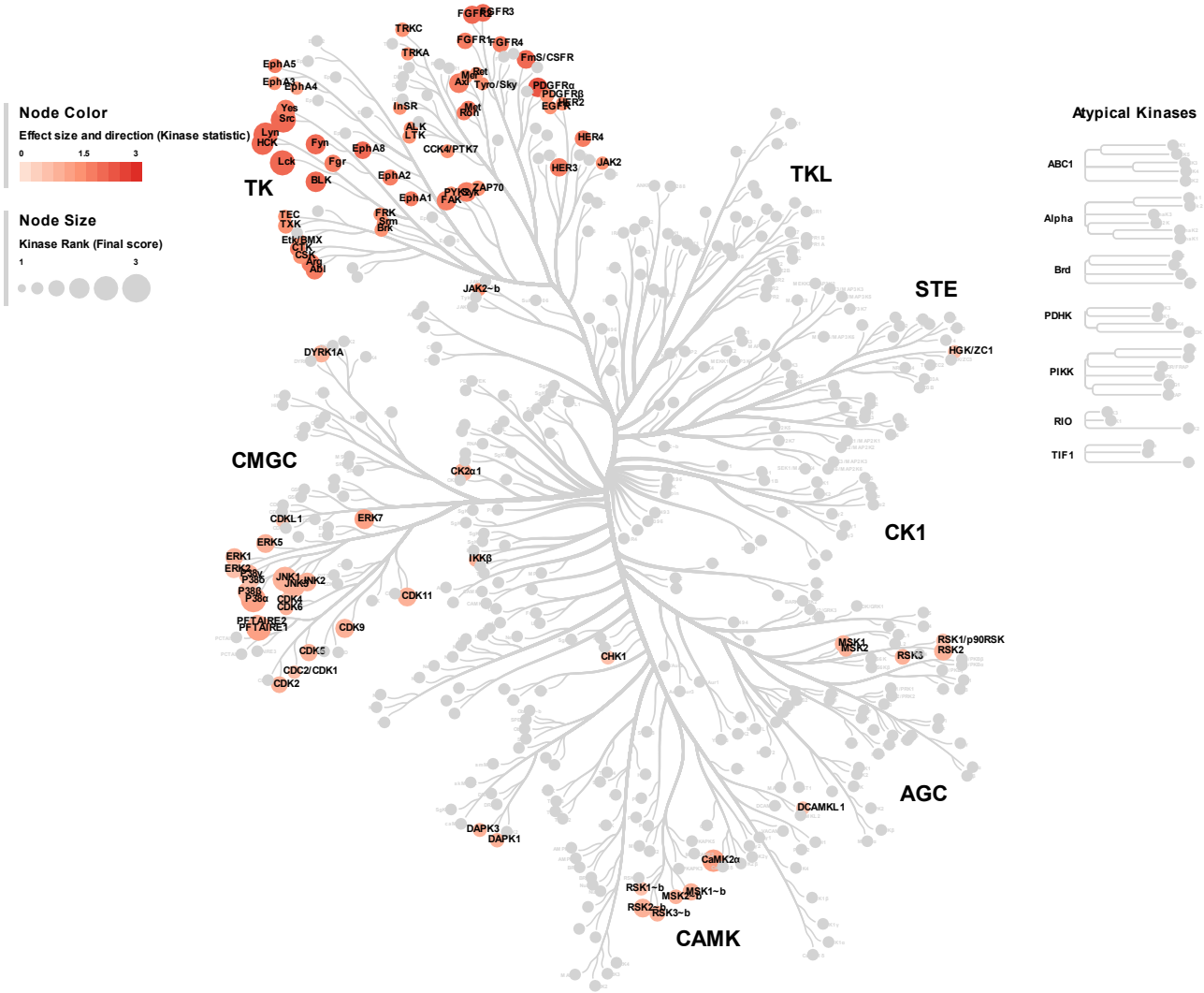
A general guideline below can be used to infer biological interpretation from the UKA results. PamGene can support and discuss these approaches.

1. Use the UKA tool to give a ranked predictive list of kinases
2. Select top kinase(s) based on:
    a. Ranking order
    b. Threshold cut-off (subjective; guideline is to use (Median) Final score > 1.2)
    c. Known Biological information
3. Further validate the selected kinase based on:
    a. Biological context inferred from current knowledge base and literature
    b. Other platforms and technologies

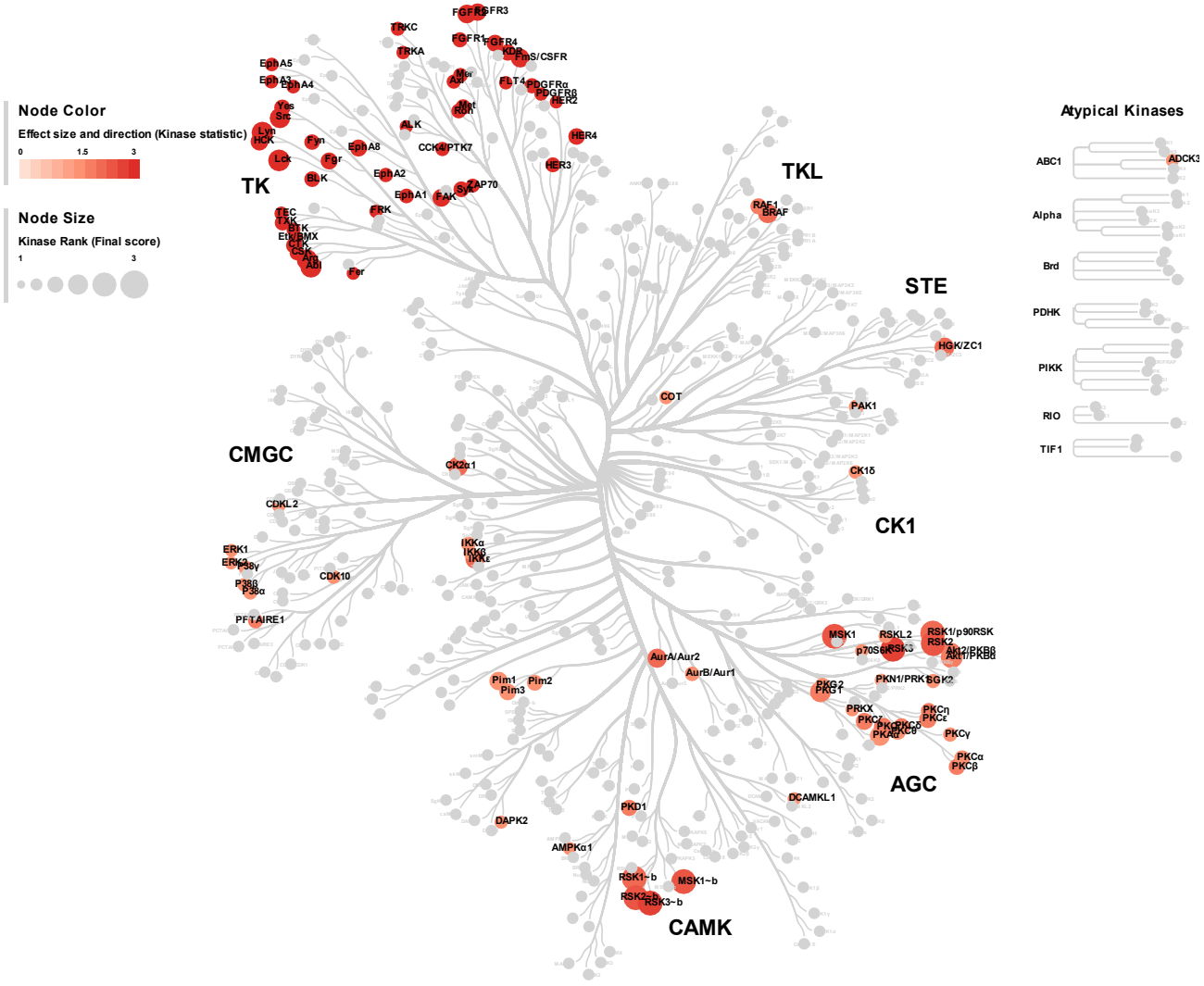
Further details of this algorithm are available on request.

3.2.2.1. Kinome Tree views (Combined; CORAL)

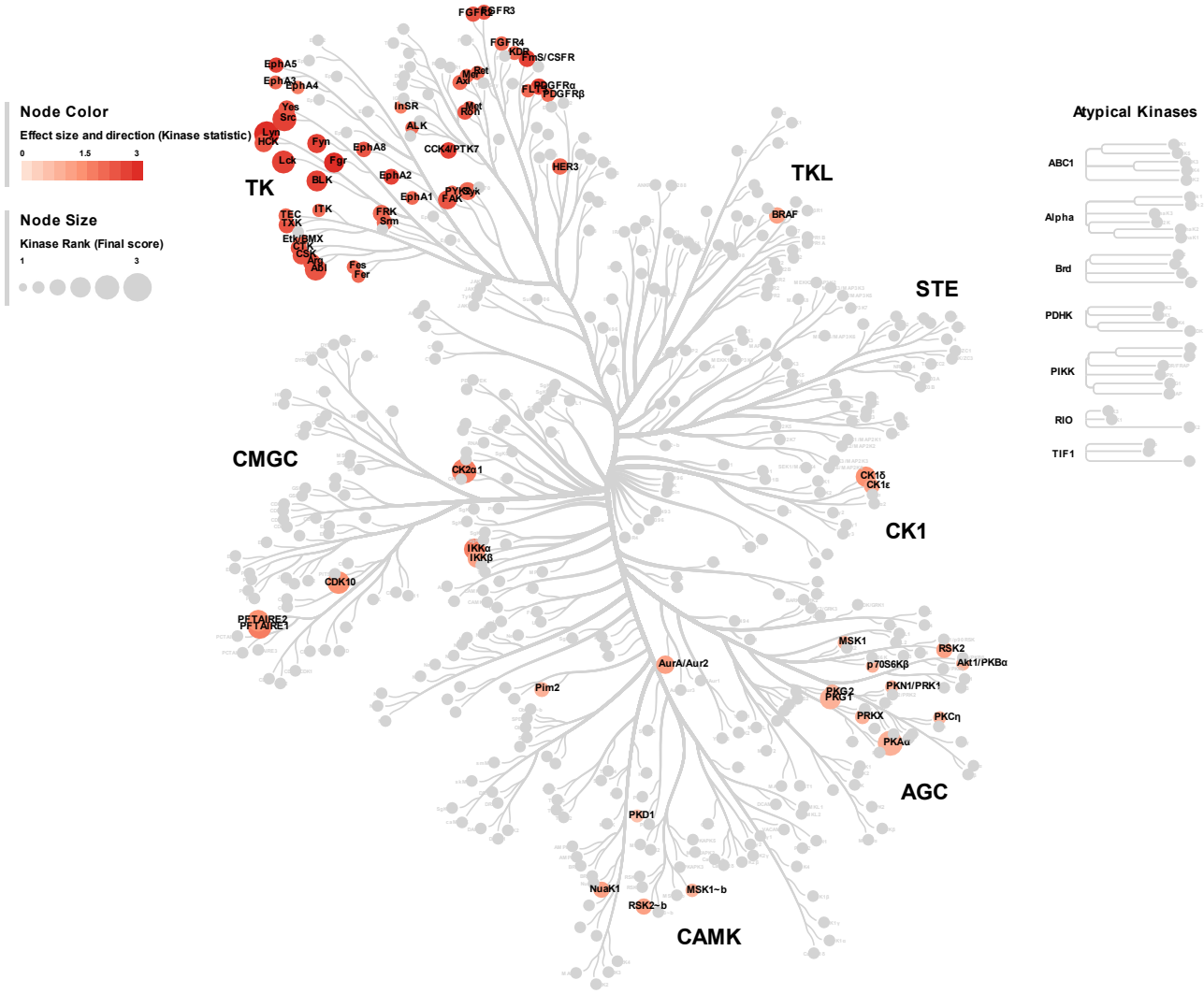
Time point T1 v C



Time point T2 v C



Time point T3 v C



3.2.2.2. Kinase Score Tables

Kinase Score Table (PTK) Time point T1 v C

Rank	Kinase.Name	Median.Final.score	Mean.Kinase.Statistic
1	Lck	3.43	2.02
2	Src	3.43	1.95
3	Lyn	2.89	2.08
4	HCK	2.56	1.99
5	BLK	2.32	1.95
6	Fyn	2.26	1.96
7	Axl	2.19	1.77
8	Syk	2.18	1.62
9	PDGFR[alpha]	2.17	2.20
10	FAK1	2.17	1.80
11	Yes	2.14	1.87
12	FmS/CSFR	2.04	1.95
13	Arg	2.03	1.73
14	HER3	1.99	1.72
15	FGFR2	1.96	2.09
16	Abl	1.91	1.69
17	FGFR3	1.89	2.07
18	EphA8	1.87	2.17
19	Fgr	1.84	1.97
20	HER4	1.72	1.71

Kinase Score Table (STK) Time point T1 v C

Rank	Kinase.Name	Median.Final.score	Mean.Kinase.Statistic
1	MAPK14	3.43	1.10
2	PFTAIRE1	3.26	1.18
3	JNK1	2.96	0.88
4	JNK3	2.96	0.88
5	CaMK2[alpha]	2.59	1.12
6	ERK7	2.23	1.11
7	p38[gamma]	2.20	0.94
8	CDK9	2.11	0.93
9	p38[beta]	2.11	1.00
10	CDK11	2.07	1.02
11	JNK2	2.05	0.86
12	ERK5	2.05	0.90
13	RSK2	2.04	0.92
14	CK2[alpha]1	1.96	1.01
15	ERK1	1.90	0.83
16	MSK1	1.84	0.89
17	CDK5	1.83	0.88
18	DYRK1A	1.83	0.94
19	ERK2	1.80	0.79
20	p38[delta]	1.76	0.82

Kinase Score Table (PTK) Time point T2 v C

Rank	Kinase.Name	Median.Final.score	Mean.Kinase.Statistic
1	Lck	2.52	5.16
2	Lyn	2.50	5.49
3	Abl	2.42	4.83
4	Arg	2.36	4.77
5	Src	2.31	4.87
6	FmS/CSFR	2.06	5.51
7	FGFR2	2.06	5.55
8	FAK1	1.94	4.66
9	HCK	1.93	4.99
10	FGFR3	1.87	5.43
11	EphA8	1.83	5.91
12	Fgr	1.81	5.77
13	FGFR4	1.81	4.85
14	HER4	1.75	4.60
15	KDR	1.72	4.70
16	Yes	1.68	4.60
17	CTK	1.68	4.65
18	Syk	1.62	4.25
19	BLK	1.61	4.57
20	Met	1.59	4.45

Kinase Score Table (STK) Time point T2 v C

Rank	Kinase.Name	Median.Final.score	Mean.Kinase.Statistic
1	RSK2	3.73	2.33
2	RSK1/p90RSK	3.73	2.52
3	MSK1	3.03	2.20
4	RSK3	3.03	2.11
5	Akt1/PKB[alpha]	2.58	1.74
6	PKG1	2.38	1.67
7	PKA[alpha]	2.25	1.64
8	HGK/ZC1	2.23	2.01
9	CK2[alpha]1	2.21	2.13
10	BRAF	2.19	2.08
11	AurA/Aur2	2.13	2.03
12	IKK[epsilon]	2.12	1.93
13	Pim1	2.04	1.58
14	PKC[iota]	2.02	1.78
15	PKC[epsilon]	1.87	1.70
16	PKC[delta]	1.85	1.64
17	PKC[zeta]	1.83	1.78
18	Akt2/PKB[beta]	1.79	1.69
19	Pim3	1.79	1.58
20	PKC[theta]	1.78	1.68

Kinase Score Table (PTK) Time point T3 v C

Rank	Kinase.Name	Median.Final.score	Mean.Kinase.Statistic
1	Lyn	3.42	2.97
2	Src	2.82	2.53
3	Lck	2.68	2.64
4	Abl	2.52	2.41
5	BLK	2.38	2.58
6	Fgr	2.26	3.03
7	Fyn	2.13	2.62
8	FAK1	2.08	2.36
9	HCK	2.00	2.47
10	Syk	1.92	2.17
11	CSK	1.91	2.29
12	Arg	1.87	2.26
13	FRK	1.83	2.24
14	FmS/CSFR	1.77	2.56
15	Yes	1.75	2.33
16	CCK4/PTK7	1.73	2.72
17	TXK	1.73	2.31
18	HER3	1.71	2.15
19	PDGFR[alpha]	1.68	2.58
20	CTK	1.65	2.22

Kinase Score Table (STK) Time point T3 v C

Rank	Kinase.Name	Median.Final.score	Mean.Kinase.Statistic
1	PKA[alpha]	3.41	1.09
2	CK2[alpha]1	2.95	1.89
3	PFTAIRE1	2.80	1.87
4	CDK10	2.61	1.61
5	PFTAIRE2	2.50	1.38
6	IKK[alpha]	2.45	1.64
7	PKG1	2.41	1.07
8	CK1[delta]	2.27	1.54
9	AurA/Aur2	2.00	1.35
10	IKK[beta]	1.97	1.37
11	BRAF	1.69	1.26
12	RSK2	1.67	1.23
13	NuaK1	1.67	1.17
14	PKG2	1.58	1.00
15	CK1[epsilon]	1.58	1.17
16	Pim2	1.49	0.98
17	PRKX	1.47	0.99
18	Akt1/PKB[alpha]	1.29	0.97
19	MSK1	1.26	1.03
20	PKN1/PRK1	1.22	1.05

4. Discussion and Conclusion

For both PTK and STK assays, there is a significant stimulation effect seen in the 3 treatments (Treatments 1, 2, 3) compared to controls. Treatment 2 shows the most significant effect.

The main result is a list of kinases (PTK and STK) for each comparison, ranked according to their putative involvement in the differential effect between treatment and control.

**Example discussion, actual discussion would include further analysis.*

PTK results (Tyrosine kinases):

The treatment time points T1, T2, T3 stimulate the Src family tyrosine kinases.

STK results (Ser/Thr Kinases):

Treatment timepoint 1 stimulates the AGC family and Treatment time point 2 stimulates the CMCG family, and also a few other STK kinases, as shown in the Kinase plots and Kinome trees. Treatment 3 stimulates to lesser extent, the AGC family.

It is recommended to perform pathway analysis using the significant peptides from the ANOVA (MTvC) results, to elucidate significant signalling pathways.

This information can be used for further biological interpretation, hypothesis generation and validation.

5. Appendix I. Data Analysis

5.1. Workflow

An overview of the workflow is given and includes some possible next data analysis steps.

1. Check and modify sample annotation (PamChip Annotator ®)
   - Annotation correctness is essential for the statistics.
2. Perform image analysis for signal quantification (BioNavigator®)
   - Converts the images to numbers for each spot (peptide)
3. Integrate exposure times (BioNavigator® and R)
   - Maximizes the dynamic range of the measurements
4. Perform quality control assessment of experiment, peptides and samples (BioNavigator® and R)
   - Assess the quality of the study, removal of low signal spots is important for the statistics
5. Visually check of measurements and annotations for all PamChip arrays (BioNavigator® and R)
   - Indicates possible outliers and annotation issues
6. Normalize the measurements, if applicable (BioNavigator® and R)
   - Rectification of systematic biases (e.g., total protein amount, runs, batch)
7. Perform a differential analysis for each control vs treatment pair (BioNavigator® and R)
   - Generate list of differential phosphosites
8. Perform differential kinase interpretation for each control vs treatment pair
   - Generate list of putative kinases
9. Generate report (BioNavigator®, R and RStudio)

All details and settings of the standard analysis workflow, as applied to the study, are contained in the BioNavigator protocol file (*.bn6) and requires the BioNavigator® software

5.2. Additional analysis

The following analysis is available (please inquire):

10. Pathway/network interpretation (Clarivate Analytics®, MetaCore)
    - Using an input list of differential phosphosites
    - find a set of relevant canonical pathways
    - find interaction networks
11. Pathway/network interpretation (Enrichr)
    - Using an input list of differential phosphosites
    - find a set of relevant canonical pathways
12. Protein analysis using Proteome Maps
13. Integration of PamChip measurements with RNAseq measurements
    - Venn diagrams showing common up and down regulation
    - Pathway analysis of the common entities

Please refer the separate supplement.docx for more background and explanation.