molecular systems biology

Mapping the human phosphatome on growth pathways

Francesca Sacco, Pier Federico Gherardini, Serena Paoluzi, Julio Saez-Rodriguez, Manuela Helmer-Citterich, Antonella Ragnini-Wilson, Luisa Castagnoli, Gianni Cesareni

Corresponding author: Gianni Cesareni, University of Rome Tor Vergata

Review timeline:

Submission date:	05 March 2012
Editorial Decision:	19 April 2012
Revision received:	24 May 2012
Editorial Decision:	20 June 2012
Revision received:	02 July 2012
Acceptance letter:	10 July 2012
Accepted:	10 July 2012

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 19 April 2012

Thank you again for submitting your work to Molecular Systems Biology. We have now heard back from two of the three referees who agreed to evaluate your manuscript, and we have decided to render a decision now to avoid further delay. As you will see from the reports below, the referees find the topic of your study of potential interest. They raise, however, substantial concerns on your work, which, I am afraid to say, preclude its publication in its present form.

While both reviewers appreciated the goals of this work and the potential value of the underlying dataset, they had substantial concerns regarding the conclusiveness of key aspects of this work, and both felt that the additional work would be needed to help extend these findings. Reviewer #2 provides a series of detailed and constructive comments related to issues of statistical analysis and manuscript clarity, which will need to be thoroughly addressed. This reviewer also felt that a search for phosphatases with multiple targets, while undoubtedly challenging, would add to the broader value of this work.

In addition, Reviewer #1 raises some important conceptual issues. First, this reviewer notes that the phosphatases studied here include two mechanistically distinct families that may need to be considered separately, especially since assumptions (like the expectation of an inhibitory function) may not be equally valid for both families. The second important issue is a general concern that activation/inhibition in this targeted signaling model may not be a generalizable indicator of oncogene vs tumor suppressor activity. Additional experimental evidence supporting some of the most novel oncogene vs tumor suppressor predictions may be helpful in this regard.

If you feel you can satisfactorily address these points and those listed by the referees, you may wish to submit a revised version of your manuscript. Please attach a covering letter giving details of the way in which you have handled each of the points raised by the referees. A revised manuscript will

be once again subject to review and you probably understand that we can give you no guarantee at this stage that the eventual outcome will be favorable.

PLEASE NOTE As part of the EMBO Publications transparent editorial process initiative (see http://www.nature.com/msb/journal/v6/n1/full/msb201072.html), Molecular Systems Biology now publishes online a Review Process File with each accepted manuscript. Please be aware that in the event of acceptance, your cover letter/point-by-point document will be included as part of this file, which will be available to the scientific community. Authors may opt out of the transparent process at any stage prior to publication (contact us at msb@embo.org). More information about this initiative is available in our Instructions to Authors.

Sincerely,

Editor - Molecular Systems Biology
msb@embo.org

Referee reports

Reviewer #1 (Remarks to the Author):

The manuscript by Sacco et al. describes the development of a bioinformatics and system biology approach to analyze a large dataset of siRNA screen with its associated various signaling phenotype. They choose on one hand to address this new approach to what they identify to be the "phosphatome" from HeLa cells. On a second hand their assessment of changes in signaling pathways is made using high throughput morphological assessment as well as evaluation of five specific signaling pathway reporter proteins. By using large-scale siRNAs screens to down regulate various human phosphatases and in examining the associated consequences on various signaling network they then build a "logic modeling "
of their finding to predict the role of the targeted phosphatases.

The overall basis of the manuscript is justified and would bring interesting and valuable data to this sets of genes (or any other through the same approach). Hence in priori it is not only interesting for the system biologists and phosphatase community" but to scientific community at large. I intrinsically like the effort of the authors to organize the "phosphatome signaling" however I have serious doubt on the validity of several of the data as well as on the choice of the authors in presenting their findings.

A first major comment is the repeated statements by the author that the phosphatase represents a super gene family. As they should know the PTPs and the ser/thr phosphatases (PPPs) have absolutely no homology whatsoever in their phosphatase domains. Hence the idea that they work similarly in signaling is inappropriate. The PP1 for example requires a trimeric complex to form the physiological active enzyme. The consequences of inhibiting a subunit versus a single member of the PTP gene family will causes quite a variety of changes. Even inside these different families the different cascades and substrates (lipids vs proteins vs sugars) also brings different difficulties in assessing the biological responses. Assessing these two large groups of proteins in the same paper does not at the limit make makes more scientific sense than assessing any two other gene family of the genome.

On this issue I consider that the authors have not sufficiently justify their target genes. Choosing one of these gene families would at the end raise the value of the data and of their manuscript. It seems to this reviewer that if they would have limited their efforts on one specific gene family of phosphatases (and their associated unit) that enzymatically would work through the same mechanism (PTP, DSPTP, PP1, PP2A etc) would have been a better first step for the validation of their analysis. In contrast it is surprising that they decided to eliminate the lipid phosphatase sub family, as they are key example of phosphatase modulating several of the pathways examined. For example the negative data obtained for PTEN is certainly a difficult hurdle for accepting the dataset. This is a major modulator of the AKT/PKB pathway and should be an outstanding positive control.

A second general comment would be that the authors should have used a specific set of siRNAs

against positive control targets. They have nicely employed the INCENP gene knock-down but other genes that are well known to affect the downstream signaling of mapk, akt/PKB pathways etc should have also be used to examine the extend of the readouts signals in their system.

A third comment relates to the way that the different phosphatases action are organized in fig.2. Those that are working in parallel pathways may have partial effect only because of their redundancy of the gene family activity. Several of these are known to work on a hierarchical fashion. For example it is likely that the PTP1B is the first PTP affecting receptor kinase function, then other one like PTPN13, or PTPN22 then PTEN, etc.. Would the network figure 2a and 2b presenting the summary of all of the effect for each phosphatase as a direct effect on the selected reporter target be misleading?
The authors appears to bias there finding by stating that "This is especially relevant because the majority of the phosphatases are expected to have an inhibitory effect and therefore their silencing should result in the activation of their target protein" . This is perhaps true for the PP family but as many PTP act as oncogene than tumor suppressors this statement and the reality are likely divergent. The model presented in the manuscript as state by the authors should ideally not prejudge the function of the gene been knockdown to make the analysis.
I do consider however that the just attempt by the authors to standardize the readout to the midlevel effect is a fair effort to analyze gene functions independently.

However, their identification of oncogene versus tumor suppressor based on these nodes is an overstatement. A good example is PTP1B assume to be an oncogenic enzyme by several studies. This is getting more complicated as it is also reported to be a negative regulator of AKT/PKB pathway but an activator for the erk pathway using knockouts.

In conclusion I like the bioinformatic aspects of the manuscript but I would not lend much credibility on the identification of any gene in the dataset as oncogenic versus tumor suppressor.

Minor comments
The use of a single oligo for any results should be discarded as potentially non-specific and with off targeting effect, as the authors themselves state. Yet they decide to include them in the dataset. Is there a reason?

Reviewer #3 (Remarks to the Author):

In their manuscript, Sacco et al. present the results of a screen of the effects of knocking down phosphatases on the activity of five signaling proteins in either untreated or TNF-treated cells. A logic-based model, built from literature insights and trained against quantitative measurements of an overlapping set of signaling proteins, was then use to interpret the findings of the screen. This yields a mapping of the "point of entry" of each phosphatase in the signaling network, if a match can be found between the phosphatase knockdown phenotype and the results of simulating model perturbation of a single node. Whether the phosphatase knockdown matches simulation of node up- or down-regulation is then indicative of the directionality of its impact on this "point of entry". These results offer important insights on possible mechanisms of regulation of the modeled pathways by these phosphatases, although as the authors do point out, the effects of phosphatases on these "point of entry" signaling proteins could be direct or indirect.

The most important issues with the manuscript relate to the transformation of the data for the quantification of the effect of phosphatase knockdown in the screen and these issues need to be addressed. The computation of the final score for each readout in the phosphatase screen is complex and several aspects need to be clarified and additional data should be provided:

1. Each readout was evaluated on a cell by cell basis yet the authors only report values that are based on a measure of central tendency (is it a mean? a median?, that information should be reported). Is the distribution of the readout value across a population of cells unimodal? or is it bimodal (which could either be evidence of the fraction of cell that were effectively transfected or of an effect that is variable from cell to cell)?

2. The authors then state that to allow comparison across plates, they calculated z-scores (based on median and average absolute deviation). Section 1 of the methods states that the cells were seeded in

LabTek chambers, not in plates. How many wells per chamber? Is this number of wells sufficient to allow for valid z-score calculations?

3. The median z-score of three biological replicates was then used as the final z-score for each oligo. With only three samples, this effectively results in choosing the middle value and throwing out the other two. The mean should be used instead as a measure of central tendency.

4. Finally to combine data for all three oligos, the authors performed a chi-square test, summing the squares of the three oligos and used a seemingly arbitrary p-value threshold of 0.04. This threshold could be instead determined by false discovery rate analysis.

Other important issues:

Transfection efficiency was estimated based on the independent transfection of an siRNA targeting INCENP - the assumption that all siRNAs transfect equally well should be stated. Furthermore, as the control transfection was repeated in many well, a distribution of the transfection efficiency should be presented as it would provide a better idea of the possible range of efficiency across different wells.

Regarding siRNA off-target effects, Snijder, B. et al. (Nature, 2009) show how some effects of siRNA knockdown on cellular phenotypes can be indirect consequences of changes in cellular context due to primary effect on cell growth, cell death, etc. Are there similar variable effects of phosphatase knockdowns on the cell density, cell growth that could affect the interpretation of their results? Could these effects help explain the known functional relationships that were not recapitulated in the screen?

In the development of the model, the authors should clarify the relationship between the model compression and the model calibration steps (Figure 4B and 4D). Which data are used in the model compression step?

The normalization of the model training set data using a Hill function should be illustrated visually either in a main Figure or in a supplemental Figure. Also, do the colored scales used in Figure 4E and S3 represent the data in the raw or normalized form? Could this form of normalization overly emphasize the ON and OFF states for each readout?

The authors should briefly discuss how the "average" behavior of 1000 models may be interpreted. If average values from the models are in the middle of the range, could it mean that most of the models disagree with the data by showing either activation or inactivation of the readout while the data shows a mid-range value for the readout?

The authors have limited their analysis and mapping strategy to perturbations of single nodes in the model. With this strategy, 30% of the phosphatases could not be matched to any profiles. Even a limited analysis with perturbations of combinations of nodes to attempt to match some of these phosphatases would enhance the impact of the study. As the authors point out, the method is particularly useful for phosphatases whose perturbations influence multiple readouts. Similarly, a computable model is particularly useful to evaluate the effect of combinations of perturbations.

In describing Figure 6A, the authors state: "The experimental results largely confirm the predictions derived from the model. Notable exceptions are the predicted modulation of NFkB and p38 by PTP4A1 and PTP4A2..." From another point of view however, this could be interpreted as being wrong for 40% of the readouts for PTP4A1! Experiments and simulation actually show completely opposing results for these two readouts, and those are the only readouts that are significantly affected by the perturbations.

Other issues:

The presentation of the experimental strategy in the beginning of the results section could be clearer. Before reading the rest of the results section, the provenance of the training set data and the process of mapping is rather nebulous. The legend of figure 1 is helpful.

Throughout the paper, descriptions of siRNA and shRNA-based experiments should be clearly described as co-transfections of a pool of constructs (shRNA experiment) or independent transfection of three different siRNAs - the derived data can then be interpreted in the appropriate context.

Many phosphatases are referred to throughout the text by their abbreviation without any additional information. To improve readability for a broad spectrum audience, adding a very brief context to each would help (for example, "CDC25C, a tyrosine phosphatase that regulates cell division").

There are several typographical errors throughout the manuscript that should be corrected and some of the figure citations seem incorrect in the last sections of the manuscript.

p.12. The authors state: "The significant enrichment (p-value < 0.01 ...) of "cancer phosphatases" in the hit list suggest that our screening preferentially selects proteins that have the potential to interfere with the control of cell growth.", then go on to test this hypothesis (Figure 6C). However this result likely simply falls out from the particular setup of the screen, especially with regards to the test conditions (untreated) and the choice of readouts (many proteins involved in growth) and from the use of what they describe as a "logic growth model" (Figure 1 legend)!

p. 14. The authors state that "... the Carma1-Bcl10-Malt1 (CBM) complex. This complex binds trans-membrane receptors, such as TCR or TNFR". While there are published reports showing interaction of the CBM complex proteins with adaptor proteins that bind to these receptors, to my knowledge, none show direct interaction with the receptors themselves.

p.14. The authors state that "In conclusion this study offers a genome-wide perspective on the involvement of protein phosphatases in the regulation of cell growth...". This should be rephrased. While the study offers data for a large number of phosphatases, their involvement is evaluated only in a very limited context after no treatment or TNF stimulation and mapped on a 15-protein network!

The supplementary methods should be integrated with main methods sections as some critical details of the experiments are only described in the supplements. In section 6 of the methods, the authors should briefly describe the fuzzy logic approach that was used in addition to citing Ref. 30.

Throughout the figures the authors use a red/white/green color scale that should be replaced with red/white/blue scale or other color scheme that is equally accessible to colorblind readers.

Figure 2. The x-axes should be labeled I npart E. In part F, are only the screen hits plotted (rather than all the data points)? If so, this should be clearly stated in the legend. The authors should also more clearly indicate which experimental conditions were chosen for the screen.

Figure 4. Part C should also include a schematic representation of the hypothetical training set data (to allow comparison of how it matches the "inferred profiles").

Figure 5. Why were the experimental conditions used in Part C different from those used in the screen? Were the simulations modified accordingly?

Figure 6. In Figure 6A and 6B, the labels on the data quantification tables need to be aligned with the columns. The legend repeatedly states that "For each measurement, the average value and standard deviation of three independent biological replicates was calculated", however only the average values are ever reported. The standard deviations (and average values) should be reported in the supplements to allow the readers to evaluate the significance of the data.

1st Revision - authors' response 24 May 2012

Thank you for your and the referees' comments, which we address here. In the following paragraphs we have marked the questions as "Q" and our answers as "A". Unfortunately we were not able to

up-load the supplementary information reporting single cell values, since these files exceed the required maximal size. However, it is down-loadable on https://www.dropbox.com/s/r9ruqxqrj8rrx1p/Single_cell_values.zip [Editor's Note: This dataset has now also been uploaded to the MSB website, and can be found among the supplementary materials]. We are confident that the manuscript is now acceptable for publication.

Reviewer 1

Q 1.1

A first major comment is the repeated statements by the author that the phosphatase represents a super gene family. As they should know the PTPs and the ser/thr phosphatases (PPPs) have absolutely no homology whatsoever in their phosphatase domains. Hence the idea that they work similarly in signaling is inappropriate. The PP1 for example requires a trimeric complex to form the physiological active enzyme. The consequences of inhibiting a subunit versus a single member of the PTP gene family will causes quite a variety of changes. Even inside these different families the different cascades and substrates (lipids vs proteins vs sugars) also brings different difficulties in assessing the biological responses. Assessing these two large groups of proteins in the same paper does not at the limit make makes more scientific sense than assessing any two other gene family of the genome.
On this issue I consider that the authors have not sufficiently justify their target genes. Choosing one of these gene families would at the end raise the value of the data and of their manuscript. It seems to this reviewer that if they would have limited their efforts on one specific gene family of phosphatases (and their associated unit) that enzymatically would work through the same mechanism (PTP, DSPTP, PP1, PP2A etc) would have been a better first step for the validation of their analysis. In contrast it is surprising that they decided to eliminate the lipid phosphatase sub family, as they are key example of phosphatase modulating several of the pathways examined. For example the negative data obtained for PTEN is certainly a difficult hurdle for accepting the dataset. This is a major modulator of the AKT/PKB pathway and should be an outstanding positive control.

A 1.1

We apologize for the confusion created by our incorrect use of the term superfamily. We have revised the manuscript making sure that there is as little confusion as possible about the evolutionary relationships and the functional differences between the proteins studied in this manuscript. Although we understand the remarks made by the referee, we feel that it is not completely unjustified to combine in the same report the analysis of two enzyme families such as PTPs and PPPs sharing the function of removing phosphate groups from proteins, albeit with different mechanisms. This is a functional screening aimed at the identification of the biological processes in which phosphatases are involved. As such it has little to do with the molecular mechanisms used by the enzymes to carry out their function. Therefore we feel that addressing the function of all the enzymes that remove phosphates from peptides, instead of restricting the analysis to specific phosphatase subfamilies, is an added value of our work. Although our primary screening included lipid and sugar phosphatases, in the validation and functional analysis we did not consider these two classes because our validation library did non include siRNA plasmids against these enzymes.
Concerning false negatives (i.e. PTEN and others) we were equally concerned as the referee. This is discussed in the *"false negatives"* section. However, it is well known that large scale analysis suffer from false positives and false negatives problems. In large-scale siRNA experiments there is no way to control the efficiency of gene expression (and protein level) knock down. Inefficient knock down is probably the case for the PTEN gene/protein since an independent study by Ellenberg and colleagues, which used the same technological platform to study mitosis, failed to identify PTEN as a regulator of cell proliferation (Neumann et al., 2010:Nature 464, 721-7).

Q 1.2

A second general comment would be that the authors should have used a specific set of siRNAs against positive control targets. They have nicely employed the INCENP gene knock-down but other genes that are well known to affect the downstream signaling of mapk, akt/PKB pathways etc should have also be used to examine the extend of the readouts signals in their system.

A 1.2

There are three things that we would like to control in a "perfect" experiment:

- the siRNA transfection efficiency in our experimental conditions. To this end we have used the INCENP siRNA, a well characterized siRNA that is known to cause a clearly detectable morphological phenotype due to interference with the cell cycle. We assume that if the INCENP siRNA efficiently enter a cell, all the remaining siRNA will do the same.
- the knock down of mRNA levels and eventually protein levels. These processes are siRNA/gene/protein specific and cannot be controlled in a large scale experiment.
- the consistency of our cell specific signaling model. That is, as hinted by the referee, we would like to know if, by down regulating the levels of an upstream node/activity, we affect in a predictable manner our readouts. This could be done by knocking down the level of well characterized genes by siRNA or, as we have done in our validation of the cell model, by perturbing the cell by incubating in different experimental conditions and with different inhibitors

In addition to these controls, to assess the reliability of our assays, before starting the screening, we performed preliminary experiments by measuring the five readouts, after inhibition of specific upstream kinases by small molecules. We have added a new figure (supplementary Figure S1) that reports the results of these experiments and supports the robustness of our experimental methodology.

Q 1.3

A third comment relates to the way that the different phosphatases action are organized in fig.2. Those that are working in parallel pathways may have partial effect only because of their redundancy of the gene family activity. Several of these are known to work on a hierarchical fashion. For example it is likely that the PTP1B is the first PTP affecting receptor kinase function, then other one like PTPN13, or PTPN22 then PTEN, etc… Would the network figure 2a and 2b presenting the summary of all of the effect for each phosphatase as a direct effect on the selected reporter target be misleading?

A 1.3

We are not sure we understand the referee's comment. Figure 2A and B are graphical representation of the results of the siRNA experiments. This is a bipartite graph where two sets of nodes are represented: 1) phosphatases (circles) and 2) phenotypes/readouts (rectangles). If a phosphatase knockdown was observed to affect a phenotype, then the corresponding rectangle and circles are linked by an edge. We thought this was a clear way to graphically represent the screening results instead of presenting in the main article the long table included in the supplementary materials. The graph is not misleading because it only represents the experimental results without any further assumption. In particular it does not contain any information about the relation between the phosphatases. The complications mentioned by the referee for sure will affect the outcome of our "simple" screening and will be detected by screenings in which siRNA knockdowns will be examined in pair. However, they have little to do with the graph representation in Figure 2.
We have revised the figure legends and text to make sure that the reader understands what the graph is showing.

Q 1.4

The authors appears to bias there finding by stating that "This is especially relevant because the majority of the phosphatases are expected to have an inhibitory effect and therefore their silencing should result in the activation of their target protein". This is perhaps true for the PP family but as many PTP act as oncogene than tumor suppressors this statement and the reality are likely divergent.

A 1.4

We completely agree with the referee. What we meant is that phosphorylation is often (erroneously) associated to activation and thus phosphatases are considered negative regulators. We have revised the text and changed the sentence to (pag.9, paragraph "Mapping of phosphatase hits")**:** "The experimental data used to calibrate the HeLa cell model did not include the results of inhibiting all the nodes in the network. Moreover the effect of up-regulating the nodes is completely absent from the experimental data. This is especially relevant for the phosphatases that inhibit their target node, as their silencing should result in an up-regulation of the target."

Q 1.5

The model presented in the manuscript as state by the authors should ideally not prejudge the function of the gene been knockdown to make the analysis. I do consider however that the just attempt by the authors to standardize the readout to the midlevel effect is a fair effort to analyze gene functions independently. However, their identification of oncogene versus tumor suppressor based on these nodes is an overstatement. A good example is PTP1B assume to be an oncogenic enzyme by several studies. This is getting more complicated as it is also reported to be a negative regulator of AKT/PKB pathway but an activator for the erk pathway using knockouts. In conclusion I like the bioinformatic aspects of the manuscript but I would not lend much credibility on the identification of any gene in the dataset as oncogenic versus tumor suppressor.

A 1.5

We understand the referee's concerns. However, from his/her statement it is not clear what he/she is specifically referring to. The classification of oncogenes and oncosuppressors with the yellow/blue dots in Figure 5 is based on an analysis of the literature and not on our results. In addition we state that the hits in our screening are enriched for genes that have been reported in the literature as oncogenes or oncosuppressors. Thus the additional hits are likely to contain new oncogenes or oncosuppressors. We think this is fair. Finally we show that some of our hits perturb the timing of cell cycle similarly to characterized oncogenes and oncosuppressors.
As far as we understand, a fair definition of an oncogene is a gene that, when mutated, can cause tumors or that is frequently up-regulated in tumors. Such an analysis is beyond the scope of this work. Thus we mitigated the conclusions of our work by changing, when referring to genes that have not been thoroughly characterized, the terms "oncogenes and oncosuppressors" with "up and down regulators of pathways that modulate cell growth".

Minor comments

Q 1.6

The use of a single oligo for any results should be discarded as potentially non-specific and with off targeting effect, as the authors themselves state. Yet they decide to include them in the dataset. Is there a reason?

A 1.6

In our experience relatively few siRNA in the 888 phosphatase oligos collection hit the target. Although we risk considering some false positives we preferred to keep phosphatase hits validated by a single oligo to avoid discarding potentially interesting information. Of all the hits considered in this work 35% (31) were identified by a single oligo. Among them we found that 4 were false positives, while 14 were validated in the secondary screening. Unfortunately, for the remaining 13, no siRNA oligos were available in the library that was used for the secondary screening. Since a high percentage (78%) of hits identified by a single oligo were validated we decided to include also the remaining 13 in the analysis.
In the new figure 5B the phosphatases that have been identified as hits on the basis of the results of a single oligo and could not be validated in the secondary screening are highlighted. These include ACPT, DUPD1, DUSP5, PPP2R2D, PPPRH, PTPRK and SSH3.

Reviewer 2

Q 2.1

The full quantitative data set and code should be made available

A 2.1

We have now added to the supplementary materials the files reporting the single cell feature values. The code that we have used for modeling signaling networks is in the CellNetOpt software that can be freely downloaded from (http://www.ebi.ac.uk/saezrodriguez/cno/). We wrote several scripts to analyze the data but, since they are specific to this project, we do not think that there is great value in releasing them. We are obviously prepared to do so if the referee wants us to.

Q 2.2

The wording is in parts to hyperbolic and should be toned down. E.g. whether measuring several signaling pathway outcomes should be called "high-content" could be a matter for debate. It would help - in particular in the abstract - to more clearly describe what has been done.

A 2.2

We have tried to address the referee's remark by toning down the wording. The term "high-content" is commonly used to describe phenotyping based on automatic fluorescence microscopy. However, in the abstract and text we have substituted "high-content" with " mutiparametric". In addition, as suggested, we have extended the abstract to provide more information (pag.1, abstract section).

Q 2.3

Not all experimental steps are well described, e.g. how many cells were analyzed? How was the variability between measurements assessed?

A 2.3

We have expanded the "Materials and Methods" section by adding more experimental information (pages 22-23, paragraphs "siRNA screen" and "Image analysis").

Q 2.4

"The interference of 35 of the 58 phosphatase hits (60%) results in a profile that matches (distance=0) one of those inferred by inactivating or activating in silico one of the nodes of the

model. These phosphatases were therefore mapped accordingly (Fig. 5A)." - unclear what is meant.

A 2.4

In this section, we compare the activation profile of the five sentinel proteins in two different data sets: the experimental one, (results of the siRNA screening), and the one, obtained *in silico* by simulating the up or down regulation of each node in the model.
Our mapping strategy is based on the assumption that, if the effect of inhibiting a phosphatase is the same as up-regulating a node, then the phosphatase is an inhibitor of that node. The matching between the experimental and in-silico datasets is done by comparing the activities of the five sentinel proteins. In 35 cases this matching is perfect, i.e. we can identify an in-silico perturbation that has exactly the same effect on the five sentinel proteins as the down-regulation of a phosphatase in the screening. Moreover, as described in the subsequent paragraphs, we also included in the final mapping a number of cases where there was a single mismatch.
We have now rephrased this statement to make it clearer to the reader, changing the text to the following state (pag. 10, paragraph "Mapping of phosphatase hits"): "For 35 out of 58 phosphatase hits (60%) we can identify an *in-silico* perturbation that has exactly the same effect (i.e. distance = 0) on the five sentinel proteins as the down-regulation of the phosphatase in the siRNA screening."

Q 2.5

How do the author justify using the described distance measure? How do other distance measures perform?

A 2.5

We had first used the euclidean distance between the vectors containing the quantitative values of the sentinel proteins in the simulation and in the siRNA screening. However, we found this method to be extremely cumbersome, especially because we needed a meaningful way of normalizing to a common scale the absolute values from the two data sources. In the end we tried several different strategies and we observed several inconsistencies with what little is known about the targets of some phosphatases. We therefore decided to choose the most unbiased way of discretizing the simulation values and used this approach to compare the simulations with the results of the screening, which were already in a qualitative form (i.e. up/down regulator of a given readout). We think that this approach is completely unbiased and easier to interpret. In particular it suggested that some experiments could be repeated in different conditions to help explain a number of imperfect mappings (see Fig. 5). The results with the euclidean distance were less clear and subjected to differences according to the normalization strategy.

Q 2.6

How often do profiles of individual siRNA diverge?

A 2.6

In our screening approximately 30% of the three oligos targeting the same phosphatase have the same activation profile, while most of the remaining 70% have a profile that diverges in at least one of the five readouts. This is likely due to the presence of false positive and false negative results, that are a common problem of large scale siRNA screenings. For this reason, the activation profile resulting form phosphatase down-regulation has been validated in the secondary screening.

Q 2.7

It is not fully clear to me how the authors come to the conclusion that "...The results of our screening support the notion that protein phosphatases can both up- (40%) or down-regulate (60%) signal transduction events."

A 2.7

In our siRNA screening we found that the activation level of the five readouts can be either increased or decreased by phosphatase down-regulation, suggesting that these enzymes can act both as positive and negative regulators. The figures represent the percentage of phosphatase-readout pairs that result in either an up- or down-regulation upon phosphatase knock down.

Reviewer 3

Q 3.1

Each readout was evaluated on a cell by cell basis yet the authors only report values that are based on a measure of central tendency (is it a mean? a median?, that information should be reported).
Is the distribution of the readout value across a population of cells unimodal? or is it bimodal (which could either be evidence of the fraction of cell that were effectively transfected or of an effect that is variable from cell to cell)?

A 3.1

For each image the mean of signal intensity of each cell was measured. This information has been reported in the Material and Method section. We agree with the referee that in principle using the mean mixes the signal of the cells that have been efficiently transfected with that of the cells in which the siRNA has had no effect. To address this issue we used Hartigan's diptest for bi-modality. In total we are measuring 16416 variables, corresponding to all the combinations of oligo, replicate, readout and inhibited phosphatase. We performed Hartigan's test on each one of the 16416 distributions of single cell values. After adjusting the p-value with Bonferroni's correction, only 95 distributions present evidence of bi-modality. Given this low number we think that using the mean is a reasonable approximation. The alternative route of parametrizing each distribution as a mixture of two normals (i.e. the signals of transfected and non-transfected cells) would be much more cumbersome. We feel that the low number of distributions that have evidence of bi-modality does not warrant this additional complication. Moreover the effect of using the mean is to lower the signal of transfected cells, by polluting it with those derived from non-transfected ones. This could eventually lead to the loss of some hit phosphatases, but does not in any way invalidate the hits we have already identified. Moreover, once a phosphatase has been defined as hit, we do not use the quantitative value of the readout anymore. Therefore even if it could be possible to arrive at a better estimate of the readout value by separating the two components of the signal, this would have no effect for the purpose of our analysis.

Q 3.2

The authors then state that to allow comparison across plates, they calculated z-scores (based on median and average absolute deviation). Section 1 of the methods states that the cells were seeded in LabTek chambers, not in plates. How many wells per chamber? Is this number of wells sufficient to allow for valid z-score calculations?

A 3.2

Sorry "plate" is our mistake. We meant "chambers" and we have revised the text (pag. 25, Material and Method section, paragraph "Statistical analysis of siRNA screening results"). Each LabTek chamber contains 384 spots, corresponding to the various siRNAs and controls. The z-score is calculated with respect to the distribution of these 384 values. We feel that this number allows for a valid z-score calculation.

Q 3.3

The median z-score of three biological replicates was then used as the final z-score for each oligo. With only three samples, this effectively results in choosing the middle value and throwing out the other two. The mean should be used instead as a measure of central tendency.

A 3.3

We spent a lot of time considering the problem of integrating the three replicates. We chose the median as in some cases there were extreme outliers due to artifacts in the fluorescent staining or in the image acquisition. We therefore used the median in order to have a measure as insensitive to outliers as possible. In the end each approach has its own pros and cons and we do not think that using the mean would be a clearly superior alternative.

Q 3.4

Finally to combine data for all three oligos, the authors performed a chi-square test, summing the squares of the three oligos and used a seemingly arbitrary p-value threshold of 0.04. This threshold could be instead determined by false discovery rate analysis.

A 3.4

We cannot perform a false discovery rate analysis as we do not have a suitably sized "golden standard" of phosphatases for which we know the relationship between their inhibition and the phenotypes we are measuring. We chose hit phosphatase by taking the ones with p-value less than a threshold. Whatever the value of the threshold, the ones with lower p-values are clearly the most significant and therefore this is the most appropriate value of selecting the hits for the follow-up screening. The specific value of the threshold was chosen in order to have a number of hits amenable to a secondary screening performed with "standard" transfection/culture technology.

Other important issues:

Q 3.5

Transfection efficiency was estimated based on the independent transfection of an siRNA targeting INCENP - the assumption that all siRNAs transfect equally well should be stated. Furthermore, as the control transfection was repeated in many well, a distribution of the transfection efficiency should be presented as it would provide a better idea of the possible range of efficiency across different wells.

A 3.5

As the referee suggested, we added a graph in supplementary material (Fig. S2), showing the phenotypic effect of INCENP down-regulation in the different spots of the different labTeks.

Q 3.6

Regarding siRNA off-target effects, Snijder, B. et al. (Nature, 2009) show how some effects of siRNA knockdown on cellular phenotypes can be indirect consequences of changes in cellular context due to primary effect on cell growth, cell death, etc. Are there similar variable effects of phosphatase knockdowns on the cell density, cell growth that could affect the interpretation of their results? Could these effects help explain the known functional relationships that were not recapitulated in the screen?

A 3.6

We are aware of these problems in large siRNA screenings. We cannot exclude that some of our hits are "very indirect". In order to mitigate problems related to cell density, cell death or cell growth, we first manually inspected images where the number of cells was larger than 170 to identify possible artifacts. Moreover we removed the images containing a number of cells in the top and bottom 2.5 percentile of the distribution of each experiment. This had already been done in the original version of the paper but the information was missing from the Material and Methods section. We have now added this information in the paragraph "Statistical analysis of siRNA screening results", pag. 25. We feel that the "known functional relationships that are not recapitulated in our screening" are better explained by inefficient siRNA knockdown or by the different cellular context.

Q 3.7

In the development of the model, the authors should clarify the relationship between the model compression and the model calibration steps (Figure 4B and 4D). Which data are used in the model compression step?

A 3.7

The model compression step (described in more detail in Saez-Rodriguez et al. 2009) removes from the literature-derived network non-identifiable elements, i.e. nodes whose state cannot be meaningfully derived from the analysis of the experimental dataset. These include:

1. nodes whose states are not affected by any of the inputs or perturbations
2. linear cascades of undesignated nodes (i.e. not perturbed or measured) that impinge on a designated node

Therefore the compression step depends both on the topology of the network and the design of the perturbation experiments. However it should be noted that, as stated, the nodes that are removed are not identifiable given the experimental design and therefore nothing is lost by removing them.
In the Results section, paragraph "Logic-based modeling of cancer associated pathways" (pag. 7), we have added a paragraph to make this clearer.

Q 3.8

The normalization of the model training set data using a Hill function should be illustrated visually either in a main Figure or in a supplemental Figure. Also, do the colored scales used in Figure 4E and S3 represent the data in the raw or normalized form?
Could this form of normalization overly emphasize the ON and OFF states for each readout?

A 3.8

According to the referee suggestion, we have added a supplemental figure (Fig. S4), showing the comparison between the raw and the normalized data. As shown in Fig. S4, this form of normalization does not emphasize the ON and OFF state. The usage of a Hill function maximizes the sensitivity in the intermediate range of values, which is probably where most of the interesting variations occur. A linear normalization would excessively compress the range of values in cases where the minimum and maximum values are very distant from the rest. A boolean model is only concerned with two values 0 and 1. Once we have estabilished that a value of e.g. 100 represents the ON (1) state we do not need to represent all the range of values up to a maximum of e.g. 1000.
The data represented in Figure 4E and S3 are in the normalized form.

Q 3.9

The authors should briefly discuss how the "average" behavior of 1000 models may be interpreted. If average values from the models are in the middle of the range, could it mean that most of the models disagree with the data by showing either activation or inactivation of the readout while the

data shows a mid-range value for the readout?

A 3.9

Each model is optimized by minimizing the difference between the predictions and the experimental data. If, for instance, half of the models show activation of a readout and the other half inactivation, it means that the experimental value is close to 0.5, i.e. essentially equidistant from 0 and 1. Both models would thus be equi-probable in this optimization scheme. In this case taking the average across all the models will correctly show a midrange value, while any single model is distant from the true experimental value which does not show a strong activation (1) or inactivation (0). In a certain sense no single model is correct, but the average model correctly reproduces the experimental situation. In the Results section, paragraph "Logic-based modeling of cancer associated pathways", we have added a paragraph to make this clearer (pag. 8).

Q 3.10

The authors have limited their analysis and mapping strategy to perturbations of single nodes in the model. With this strategy, 30% of the phosphatases could not be matched to any profiles. Even a limited analysis with perturbations of combinations of nodes to attempt to match some of these phosphatases would enhance the impact of the study. As the authors point out, the method is particularly useful for phosphatases whose perturbations influence multiple readouts. Similarly, a computable model is particularly useful to evaluate the effect of combinations of perturbations.

A 3.10

As suggested by the referee we performed new simulations where we computed the effect of up/down regulating all the possible pairs of nodes in the model. This turned out not be very informative since we were able to map only 3 additional phosphatases (whose profile differed from an inferred profile by only one readout in the simulation with single perturbations) while the remaining 14 phosphatase profiles were either: i) not compatible with the model, i.e they could not be reached in any of the simulated experimental conditions involving perturbation of up to two nodes; ii) not specific to the perturbation of any node, or pair of nodes. We have added a description of the results of these simulations in the Results section (pag. 12). We have also added half a page in the Discussion section (pag. 15-16) detailing the possible reasons that explain the failure in the mapping of these phosphatases

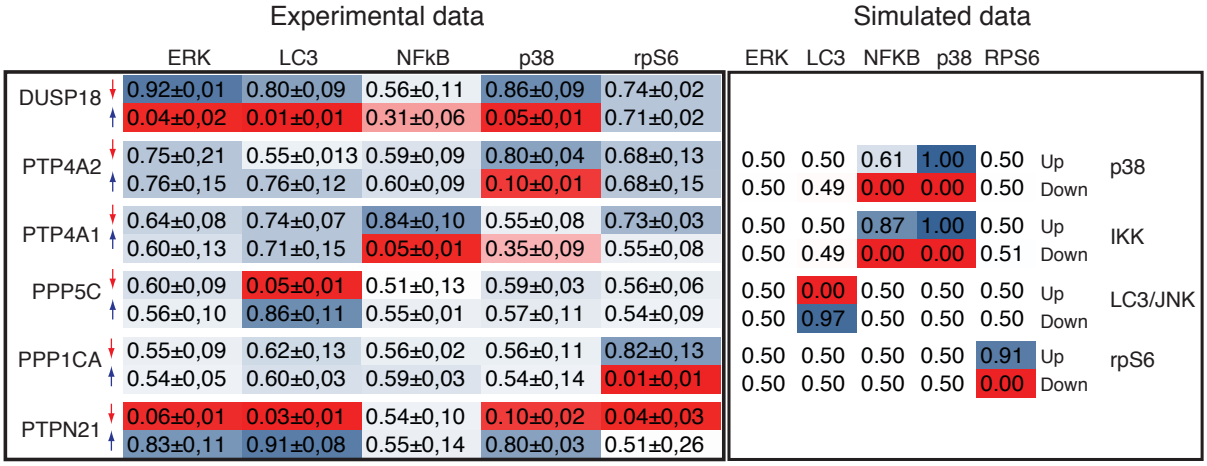
Q 3.11

In describing Figure 6A, the authors state: "The experimental results largely confirm the predictions derived from the model. Notable exceptions are the predicted modulation of NFkB and p38 by PTP4A1 and PTP4A2..." From another point of view however, this could be interpreted as being wrong for 40% of the readouts for PTP4A1! Experiments and simulation actually show completely opposing results for these two readouts, and those are the only readouts that are significantly affected by the perturbations.

A 3.11

The observation of the referee suggest us that Figure 6A is not clear. We have now revised that figure, demonstrating that our results largely confirm the predictions derived from the model, with the exception of PTP4A1 and PTP4A2. Indeed our model predicts that PTP4A1 up-regulation or down-regulation should respectively increase or decrease the NFkB and p38 activity. In our experimental results we observed that upon PTP4A1 down-regulation, only NFkB is affected, while PTP4A1 over-expression induces a dramatic reduction of NFkB activity and to a lesser extent of p38, confirming model prediction. On the other hand, our strategy infers that PTP4A2 down-regulation should inhibit only p38, while phosphatase up-regulation should affect both p38 and

NFkB activity. As shown in Fig. 6A, PTP4A2 down-regulation only activates p38, confirming the model prediction. However, upon PTP4A2 over-expression we only observed a drastic reduction of p38 activity, while NFkB is not affected.
Thus, among the 12 inferred affected readouts, 10 were confirmed by our experimental results, while 2 were not, indicating that our prediction can be successfully validated. It should be noticed that in this calculation we have not considered the readouts that are predicted, and observed experimentally, to be unaffected. If we include also these predictions, the total number of correct predictions becomes 40.



Other issues:

Q 3.12

The presentation of the experimental strategy in the beginning of the results section could be clearer. Before reading the rest of the results section, the provenance of the training set data and the process of mapping is rather nebulous. The legend of figure 1 is helpful.

A 3.12

We have made an effort to identify the sentences that would benefit from additional information and we have changed the text accordingly.

Q 3.13

Throughout the paper, descriptions of siRNA and shRNA-based experiments should be clearly described as co-transfections of a pool of constructs (shRNA experiment) or independent transfection of three different siRNAs - the derived data can then be interpreted in the appropriate context.

A 3.13

We have revised the paper to take into account the referee's remarks.

Q 3.14

Many phosphatases are referred to throughout the text by their abbreviation without any additional information. To improve readability for a broad spectrum audience, adding a very brief context to each would help (for example, "CDC25C, a tyrosine phosphatase that regulates cell division").

A 3.14

As the referee suggested we have added a very short description of the analyzed genes whenever we felt that this could help the non specialized reader. However, the referee understands that we are dealing with 300 phosphatases some of which are referred to in the text. The manuscript would become unreadable if we had to do this every time we mention a phosphatase.

Q 3.15

There are several typographical errors throughout the manuscript that should be corrected and some of the figure citations seem incorrect in the last sections of the manuscript.

A 3.15

We have corrected all the errors we identified.

Q 3.16

p.12. The authors state: "The significant enrichment (p-value < 0.01 ...) of "cancer phosphatases" in the hit list suggest that our screening preferentially selects proteins that have the potential to interfere with the control of cell growth.", then go on to test this hypothesis (Figure 6C). However, this result likely simply falls out from the particular setup of the screen, especially with regards to the test conditions (untreated) and the choice of readouts (many proteins involved in growth) and from the use of what they describe as a "logic growth model" (Figure 1 legend)!

A 3.16

The purpose of the experiments shown in Figure 6C is to show that a subset of the proteins affecting our molecular readouts also have an effect on cell cycle timing, which is a phenotypic readout of a much higher level that was not explicitly considered in the screening. Essentially, we are trying to link a low-level phenotype (the values of the measured readouts) with a high-level one (cell cycle timing). Although this may be considered far away from the main theme of the paper, it adds additional evidence that we have identified nodes whose perturbation affects cell growth. Implicitly, although we do not show it directly, we assume that this is a consequence of the perturbation of the nodes that we have identified in our experiments. Although it is true that the cells in the experiment in Fig. 6 are grown in serum and are not treated with TNFa, serum is an input node in our model, it is one of the experimental condition we used in the primary screening and its modulation in combination with the perturbation of other nodes is predicted to affect the readouts.

Q 3.17

p. 14. The authors state that "... the Carma1-Bcl10-Malt1 (CBM) complex. This complex indirectly binds trans-membrane receptors, such as TCR or TNFR". While there are published reports showing interaction of the CBM complex proteins with adaptor proteins that bind to these receptors, to my knowledge, none show direct interaction with the receptors themselves.

A 3.17

The referee is correct the interaction is indirect. The CBM complex binds PKC, TRAF6 and TRAF1 which are known interactors of TCR and TNFR. We changed the sentence to “This complex *indirectly* binds trans-membrane receptors, such as TCR or TNFR" (Discussion section, pag. 16).

Q 3.18

p.14. The authors state that "In conclusion this study offers a genome-wide perspective on the involvement of protein phosphatases in the regulation of cell growth...". This should be rephrased. While the study offers data for a large number of phosphatases, their involvement is evaluated only in a very limited context after no treatment or TNF stimulation and mapped on a 15-protein network!

A 3.18

We rephrased the sentence as follows “In conclusion this study offers a genome wide perspective on the involvement of protein phosphatases in the modulation of cell growth under TNFa stimulation” (Discussion section, pag. 17).

Q 3.19

The supplementary methods should be integrated with main methods sections as some critical details of the experiments are only described in the supplements. In section 6 of the methods, the authors should briefly describe the fuzzy logic approach that was used in addition to citing Ref. 30.

A 3.19

We have included the supplementary methods in the Material and Method section, providing a short description of the Fuzzy logic approach (paragraph “Simulating up- and down-regulation of each node”, pag. 26).
.

Q 3.20

Throughout the figures the authors use a red/white/green color scale that should be replaced with red/white/blue scale or other color scheme that is equally accessible to colorblind readers.

A 3.20

As the referee suggested, we have now changed the color scale of the figures.

Q 3.21

Figure 2. The x-axes should be labeled In part E. In part F, are only the screen hits plotted (rather than all the data points)? If so, this should be clearly stated in the legend. The authors should also more clearly indicate which experimental conditions were chosen for the screen.

A 3.21

We have now revised the text by stating in the legend that in Figure 2F only phosphatase hits are plotted and we indicated the experimental condition chosen for the screen (pag. 19).

Q 3.22

Figure 4. Part C should also include a schematic representation of the hypothetical training set data (to allow comparison of how it matches the "inferred profiles").

A 3.22

As the referee suggested, we have now added in Figure 4C a schematic example of the hypothetical training dataset.

Q 3.23

Figure 5. Why were the experimental conditions used in Part C different from those used in the screen? Were the simulations modified accordingly?

A 3.23

The experimental profile after inhibition of the phosphatases PPP2R5C, CDC25C and PPP2R1A

matched 4 out of 5 readouts of an inferred profile. The only exception being the LC3 readout. LC3 (autophagy) is predicted to be down regulated in conditions where S6 is activated while is found unchanged in the screening.
We reasoned that this imperfect match could represent a false negative caused by the low sensitivity of the LC3 western blot assay in the condition of the screening: growth in serum. Autophagy is activated at very low levels in the presence of serum. To test such a hypothesis, the effect of the over-expression or down-regulation of these three phosphatases was monitored in starved cells (no nutrients, no amino-acids). In these conditions autophagy is activated and its down-regulation can be easily monitored. Thus, the simulated values shown in figure 5C are indeed the ones obtained in these conditions (no Serum and no TNFa). As illustrated in Fig. 5C, these three phosphatases affect the LC3 level in such experimental condition, as predicted by the model.

Q 3.24

Figure 6. In Figure 6A and 6B, the labels on the data quantification tables need to be aligned with the columns. The legend repeatedly states that "For each measurement, the average value and standard deviation of three independent biological replicates was calculated", however only the average values are ever reported. The standard deviations (and average values) should be reported in the supplements to allow the readers to evaluate the significance of the data.

A 3.24

Label in Fig.6 are now aligned. For each average measurement, we have now added the standard deviations (Fig. 6A-B).

2nd Editorial Decision 20 June 2012

Thank you again for submitting your work to Molecular Systems Biology. We have now heard back from the two referees who agreed to evaluate your revised study. As you will see, the referees felt that the changes made to this work had addressed their main concerns, and they are now largely supportive. They make some important suggestions for modifications, though, which we would ask you to carefully address in a final revision of the present work.

In addition to the points raised by Reviewer #3, we ask you to address the following minor format and content issues when preparing your revised work:

1. Regarding Fig. 6C, we generally discourage the use of error bars or standard deviations with only two independent replicates. I would suggest plotting the two data points directly. I also encourage to provide these measurements, and other numeric data in the figures, as figure source data (e.g. <http://tinyurl.com/365zpej>). Please see our Instructions of Authors for more details on preparation and formatting of figure source data (<http://www.nature.com/msb/authors/index.html#a3.4.3>).

2. The text in the figures is rather small, and I am concerned that it will be hard to read when the figures if formatted within a final publication. I would recommend the following: 1) use a vertical layout for Figures 1 and 3; 2) split Fig. 6 into 2 figures (also appropriate given the very long legend); 3) increase font sizes where possible -- the text should not be smaller than about 9pt when printed to a typical single page size.

3. The Supplementary Information pdf should begin with a Table of Contents that lists all supplementary materials (figures, tables, including separate files).

Thank you for submitting this paper to Molecular Systems Biology.

Sincerely,

Editor - Molecular Systems Biology
msb@embo.org

Referee reports

Reviewer #1 (Remarks to the Author):

The manuscript has improved somehow by the additional details that were added.
I am satisfied with the argumentation of the rebuttal letter as for my questions. One must understand that this is a novel direction in assessing signaling pathways controlled by phosphatases. I remain somehow skeptical of the signaling data that are obtained in conclusion.

Reviewer #3 (Remarks to the Author):

In their revised manuscript, Sacco et al. have addressed most of my original comments. The presentation of their analyses and data is better and clearer. A few minor points should still be clarified or and corrected:

Figure S2 (and p.6): "Transfection efficiency was estimated from the appearance of polylobed nuclei (>85%) in cells silenced for the INCENP gene..." The percentages of polylobed cells (instead of the mean nuclear area) be more useful to report in Figure S2, if those data are available.

p.6 (middle paragraph): Figure referencing seems to refer to a prior version of Figure 2: 2D should be cited instead of 2B; 2E instead of 2C, 2F instead of 2D.

p.19 (Figure 2 legend): "INCEMP" instead of "INCENP" (also misspelled on the figure itself, part 2D).

p.22 (Figure 6 legend): still refers to the green/red color scheme, instead of blue/red.

p.25, (Statistical analysis of siRNA screening results): Instead of "... with respect to all the points in the same chamber."; stating "... with response to all *384* points in the same chamber." would remove any of the ambiguities that were the source of my criticisms in the first version of the manuscript.

p.25, (Statistical analysis of siRNA screening results): Adding the explanation (stated in A3.3) of why the median was chosen instead the mean within this section would allow the authors to justify their choice of metric for the readers of the paper.

p.27 (Simulating up- and down-regulation of each node): The new brief description of the logic-based modeling strategy is very helpful. However, what the authors describe is a three-state (or multistate) logic-based model, not, strictly speaking, a fuzzy logic approach. A fuzzy approach would add the possibility of a partial membership in each state. Please clarify further.

Figure 2D: Use "siINCENP" instead of "siINCEMP".

Figure 2E: labeling the x-axes as "z-score" would be helpful.

Figure 6A and 6B: In the simulated data chart (right side), the column labels still appear misaligned.

Figure S4: The figure is very helpful, although bar graphs seem more appropriate as there are categorical variables.

2nd Revision - authors' response 02 July 2012

We have now revised text and Figures according to the editor’s and referee’s comments. In the following paragraphs we have listed all the changes in the revised version of the paper. Unfortunately we were not able to up-load the supplementary information reporting single cell values, since these files exceed the required maximal size. However, it is down-loadable on

https://www.dropbox.com/s/r9ruqxqrj8rrx1p/Single_cell_values.zip. We are confident that the manuscript is now acceptable for publication.

1.1 Regarding Fig. 6C, we generally discourage the use of error bars or standard deviations with only two independent replicates. I would suggest plotting the two data points directly. I also encourage to provide these measurements, and other numeric data in the figures, as figure source data (e.g. <http://tinyurl.com/365zpej>). Please see our Instructions of Authors for more details on preparation and formatting of figure source data (<http://www.nature.com/msb/authors/index.html#a3.4.3>.
As suggested, we have revised Figure 6C, adding the source data for both the experiments reported in Figure 6 and Figure 7.
1.2. The text in the figures is rather small, and I am concerned that it will be hard to read when the figures if formatted within a final publication. I would recommend the following: 1) use a vertical layout for Figures 1 and 3; 2) split Fig. 6 into 2 figures (also appropriate given the very long legend); 3) increase font sizes where possible -- the text should not be smaller than about 9pt when printed to a typical single page size.
We have modified the figures, according to the suggestion.

1.3. The Supplementary Information pdf should begin with a Table of Contents that lists all supplementary materials (figures, tables, including separate files).
We have now added the table contents in the supplementary material section
Reviewer #3 (Remarks to the Author):

3.1 Figure S2 (and p.6): "Transfection efficiency was estimated from the appearance of polylobed nuclei (>85%) in cells silenced for the INCENP gene..." The percentages of polylobed cells (instead of the mean nuclear area) be more useful to report in Figure S2, if those data are available.
Since we do not have any pipeline to automatically identify polylobed nuclei, or to automatically calculate the percentage of polylobed cells, we have estimated the percentage of polylobed nuclei by counting the nuclei, whose area is larger than twice the mean nuclear area in scrambled controls. For a few images we have checked that this estimates match the frequency of polylobed cells as estimated visually. The results of our analysis have been added to Fig. S2 and confirms the high transfection efficiency obtained in our siRNA screening.

3.2 p.6 (middle paragraph): Figure referencing seems to refer to a prior version of Figure 2: 2D should be cited instead of 2B; 2E instead of 2C, 2F instead of 2D.
Thanks for spotting this problem. We have revised the text.

3.3 p.19 (Figure 2 legend): "INCEMP" instead of "INCENP" (also misspelled on the figure itself, part 2D).
We have revised the text.

3.4 p.22 (Figure 6 legend): still refers to the green/red color scheme, instead of blue/red.
We have revised the legend of Figure 6.

3.5 p.25, (Statistical analysis of siRNA screening results): Instead of "... with respect to all the points in the same chamber."; stating "... with response to all *384* points in the same chamber." would remove any of the ambiguities that were the source of my criticisms in the first version of the manuscript.
We have revised the text according to the referee's suggestion.

3.6 p.25, (Statistical analysis of siRNA screening results): Adding the explanation (stated in A3.3) of why the median was chosen instead the mean within this section would allow the authors to justify their choice of metric for the readers of the paper.
We have now added this explanation to the "Statistical analysis of siRNA screening results" section.

3.7 p.27 (Simulating up- and down-regulation of each node): The new brief description of the logic-based modeling strategy is very helpful. However, what the authors describe is a three-state (or multistate) logic-based model, not, strictly speaking, a fuzzy logic approach. A fuzzy approach would add the possibility of a partial membership in each state. Please clarify further.
We have now revised the text, according to the referee suggestion.

3.8 Figure 2D: Use "siINCENP" instead of "siINCEMP". Figure 2E: labeling the x-axes as "z-score" would be helpful.
As suggested by the referee, we have modified Figure 2.

3.9 Figure 6A and 6B: In the simulated data chart (right side), the column labels still appear misaligned.
As suggested by the referee, we have aligned the columns labels.

3.10 Figure S4: The figure is very helpful, although bar graphs seem more appropriate as there are categorical variables.
According to the referee suggestion, we have now changed Figure S4 into a dispersion graph, since the bar graph was less clear.

Acceptance letter 10 July 2012

Thank you again for sending us your revised manuscript. We are now satisfied with the modifications made and I am pleased to inform you that your paper has been accepted for publication.

Before we can send this work to production, we have some minor format and content issues that will need to be addressed:

1. The legend for Fig. 7 mentions standard deviation although the bars where removed from the figure. (thanks for revising these figures and providing the figure source files)

2. The single_cell_values.zip has been uploaded as "Dataset 1" please reference this dataset at least once in the main manuscript, and list it in the Table of Contents at the beginning of the Supplementary Information pdf (as "separate file").

Revised files addressing these minor points can sent as attachments to a reply email.

Thank you very much for submitting your work to Molecular Systems Biology.

Sincerely,

Editor - Molecular Systems Biology
msb@embo.org

