

Distinct Cellular States Determine Calcium Signaling Response

Jason Yao, Anna Pilko and Roy Wollman

Corresponding author: Roy Wollman, UCLA

Review timeline:

Submission date:	15 June 2016
Editorial Decision:	27 July 2016
Revision received:	21 October 2016
Editorial Decision:	11 November 2016
Revision received:	16 November 2016
Accepted:	18 November 2016

Editor: Maria Polychronidou

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

27 July 2016

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 study. Unfortunately, after several reminders we have not managed to obtain a report from reviewer #3. Since the recommendations of the other two referees are similar, I prefer to make a decision now rather than further delaying the process. As you will see below, the reviewers acknowledge that the study is interesting and that the presented methodology is likely to be useful for the field. However, they raise a number of concerns, which should be carefully addressed in a major revision of the manuscript.

Without repeating all the points listed below, some of the more fundamental issues are the following:

- The conclusion regarding the existence of distinct states needs to be strengthened by additional analyses.
- Providing further experimental evidence related to the mechanisms giving rise to the three cells states would significantly enhance the impact of the study.

REFeree REPORTS

Reviewer #1:

Review of 'Distinct Cellular States Determine Calcium Signaling Response' by Jason Yao and Roy

Wollman.

In this manuscript, the authors perform timelapse imaging of calcium response to ATP in single MCF10A cells to obtain single-cell timetraces. They then perform mathematical modelling combined with a Bayesian Inference-based fitting approach to identify and fit sets of differential equation systems to each single cell, to determine parameter values for each single cell that result in models that describe pre-processed single-cell traces best. Finally, they calculate the divergence between the probability distributions of these parameters of two cells in a pairwise manner across the whole dataset using a method from information theory (Kullback Leibler divergence), and use this, in combination with the actual distributions, to cluster single cells, revealing three major cell clusters. These three clusters are then interpreted as providing possible mechanistic insight into the nature of these three clusters, namely the existence of strong positive feedback of calcium on the IP3 receptor, weak positive feedback, and negative feedback.

Generally, I found the manuscript easy to follow, it is well written, explains the ideas and methods well, and communicates the results clearly. The overall impression is that this is a simple and elegant approach, which appears to work well. The findings provide support for one of the two main views on how cell-to-cell variability arises in isogenic mammalian cell populations exposed to the same conditions, namely from the existence of distinct cellular states (or state distributions) which are determined by differences in cellular physiology and not caused by random or stochastic drifts. As such, the work is an important contribution and adds to the increasing body of literature that favors this more deterministic view. It also provides a potentially very useful and relatively straightforward method to a large number of researchers that are analyzing single-cell time traces of dynamic cellular responses, which may yield surprising insights, notably in the cell signaling field where the prevailing view still is that most variability is of a stochastic nature.

While the above is all positive, I do have some major criticisms, which I summarize below:

1. The overall impression of the paper is that it is a bit 'thin'. There is no experimental validation of the three states. The paper would significantly strengthen if such work would be included. Could the authors come up with measurements of the strength of positive feedback in single cells? Could they explore the use of IP3R inhibitors or agonists to dissect this? I am not an expert on calcium signaling, but it appears to me that many experimental tools are available for this system to obtain experimental evidence for the proposed mechanisms that underlie the three major cell states. With such data included, the paper would be much stronger.
2. As the authors state, they pre-process the raw time-traces before they attempt model fits, and they also somewhere mention low-pass filtering of the data. Clearly, this removes some of the variability in the data. An alternative approach is to ask how much of the total variability in the data (without any pre-processing) can be explained by the models. And after removing that, is the amount of unexplained variability approaching the Poisson limit, plus some experimental noise? Can the amount of experimental noise in their setup be estimated?
3. Finally, while the distinct cellular states support the more deterministic view, the authors provide no insight in what is the source of these different states? In theory these could still emerge from stochastic events occurring in growing cell populations which then become stabilized. The authors suggest from Figure 1f-g and extended view Figure 3 that it is not correlated with the spatial positioning of cells or with the amount of DNA in cells (used as a proxy for position in the cell cycle). This analysis is however underwhelming. Merely glancing at an image of a couple of hundred cells does not provide any quantitative insight about whether there are correlations with local cell density, cell size, or cell shape. Furthermore, Hoechst staining can be tricky to interpret (depending on amount and time of incubation). Do the authors see two nice peaks of G1 and G2 cells in Hoechst intensity distributions? Providing more insight into what could be the sources that determine the 3 major cellular states would also be a major improvement of the manuscript.

Reviewer #2:

Yao & Wollman address an interesting and timely question: to what degree the experimentally observed single-cell variability results from purely stochastic biochemical events and/or whether discrete states exist in a population. They address it by measuring single-cell calcium responses to ATP and by further fitting the responses to a previously published model.

Overall, the methodology presented in the paper is a valuable addition to the field, however, the conclusions about the existence of discrete states in the parameter space are not convincing given the current presentation. Thus, before recommending the paper for publications the authors should address the following.

Major issues:

1. The clustering of distance measure based on KL divergence shown in Fig.4 implies 3 distinctive groups in the parameter space. At the same time, authors conclude that hierarchical clustering of measured responses shown in EV Fig.1A doesn't yield good grouping. One issue I have with this claim is that just by looking at the heat-map/waterfall plot EV Fig.1A I can discern ~6 clusters, based mainly on the time to reach the final steady-state value.

Regardless of authors' claim about the lack of "natural grouping" of experimental time courses, it is unclear to me why they insist there wouldn't be a (decent) agreement between explicit clustering of experimental time courses and clustering of model parameters fitted to those very same time courses. After all, the model with parameter sets from different clusters shown in Fig.4 should generate time courses different enough to classify/assign them to a couple of distinct clusters. If that wasn't the case, the robustness of fitting would be in question.

In other words, distinct states in the parameter space that authors attribute to different feedback mechanisms should relate to distinct dynamic responses. This is partly addressed in Fig.4C but the "green" cluster contains all kinds of dynamic responses, from weak to strong activation. How does it relate to explicit clustering of measured responses (based on entire time courses or their features)?

Additionally to addressing the above concerns, I ask the authors to:

(a) provide a confusion matrix of clustering of raw (or smoothed) measured responses against classification of fitted parameters. How similar percentage wise are those two classifications?

(b) expand Fig. 4 with a panel including time course features (according to Fig. 1D). Are those features distinct between 3 clusters?

(c) add labels designating feedback mechanism to the 3 clusters.

2. Related to the above is the question about response dynamics within three clusters that were associated with different feedback strengths and types. For example, it is known that due to the negative feedback the dynamic responses are transient, while strong positive feedback can induce sustained responses. Thus, are responses belonging to the "Neg. Feedback" cluster more transient (decay faster) than those in strong and weak PF?

If no such differences exist in time courses that belong to any of 3 clusters, then the mechanistic explanation of these clusters is questionable.

3. The following sentence is unclear:

"For every two parameter vector populations corresponding to each of two cells, the distance measure was defined as the KL divergence from a null model where both samples were obtained from identical probability distributions."

Please rephrase/expand the definition of the distance. In particular, please comment on the validity of KLdiv usage in this context, given that KLdiv isn't a true metric, i.e. (usually) $KLdiv(a || b) \neq KLdiv(b || a)$.

Also, the explanation in M&M - "Clustering of the Parameter Ensemble" is not clear and in its current form one can easily get lost in terminology. The entire section would greatly benefit from a

scheme of the work-flow, or an itemised step-by-step explanation.

4. Fig. 1E

The authors have excluded the influence of spatial localisation and cell cycle. However, part of the observed variability could still result from pre-existing cellular states. For example high basal levels (F_0) could imply high peak value (F_m). To further exclude those effects authors should provide scatter plots for pair-wise comparison of 6 main time course features. Are they correlated?

Minor issues:

1. Fig. 1A.

How many cells have been measured in a single experiment?

What is the colour coding scale associated with this panel?

2. Fig. 1E.

The range of Y-axis scale makes it difficult to discern the boxes. Consider removing two outliers in the last two rows, or skipping the last two columns with ratios altogether.

Please add the equation used to calculate the Z value.

3. EV Fig.1B

The visualisation doesn't seem to be appropriate. 6 features along the x-axis transition smoothly; it's hard to recognise that the axis is discrete.

4. EV Fig.3

Current box plot visualisation makes it difficult to see whether G1, S, G2/M are indeed equally populated across three clusters. Thus, I suggest to change this plot or to add a panel with a more familiar cell cycle plot, i.e. line plot of the probability density of Hoechst intensity for 3 different feedback types/strengths.

5. Fig. 3D

What kind of normalisation was used? Area, count, width? Please describe in the caption. The sentence from the caption: "The width of each violin shows the probability density of the underlying distribution." is unclear.

6. EV Fig.4

What does it mean that whiskers do not extend from the box? E.g. "eta1 (ER)" for WP.

What kind of box-plot is that? What do whiskers signify?

7. Sentence, page 9, par. 1:

"To test if the nearest neighbours probability that we calculated shows a statistically significant different..." -> difference

8. Table S1 was missing from my copy of the manuscript.

9. Please explicitly define what "parameter vector" means.

10. Since the model was simplified compared to the original, please deposit the model into BioModels database.

We are excited to re-submit a revised manuscript for your consideration. The paper was revised based on the reviewer comments and now includes more detailed analysis and new experimental data. Below is point by point response to reviewer comments. For clarity we show reviewer comments in serif fonts italics and our response in san serif below.

Editorial comments

- The conclusion regarding the existence of distinct states needs to be strengthened by additional analyses.

We enhanced the basic cluster analysis presented in this work in the following way. In the original analysis the final cluster “cutoff” was determined based on what we subjectively considered to be “natural clusters”. In the revised submission we now use formal criteria for cluster number identification based on modified Calinski-Harabasz criteria. The use of formal criteria improves the analysis in two important ways: 1. it removes some of the subjectivity in the procedure. 2. It allows formal comparison between clustering based on parameters inference to clustering based on calcium responses directly. These comparisons show that while there is some similarity between the groupings based on different clustering approaches, the quality of clustering based on parameter inference is much higher than any other approach.

- Providing further experimental evidence related to the mechanisms giving rise to the three cell states would significantly enhance the impact of the study.

Our analysis identified IP3R activity as a key factor that distinguishes the three cell states and determines calcium response. Therefore this analysis predicts that response IP3 should explain the majority of cellular response variability. This is a non-trivial prediction as it was not clear a-priori that this is where majority of the variability is and not for example in receptor level. We experimentally tested this prediction with a new experiment where we stimulate cells with ATP, allow them to recover for 30 min and then uncage using light a caged IP3. We show that the amplitude of calcium response to IP3 explains ~70% of the variance in amplitude of calcium response to ATP. The new data (shown in Figure 5d) provides a validation of our key finding in way that is independent on any parameter fitting.

Reviewer #1:

Generally, I found the manuscript easy to follow, it is well written, explains the ideas and methods well, and communicates the results clearly. The overall impression is that this is a simple and elegant approach, which appears to work well. The findings provide support for one of the two main views on how cell-to-cell variability arises in isogenic mammalian cell populations exposed to the same conditions, namely from the existence of distinct cellular states (or state distributions) which are determined by differences in cellular physiology and not caused by random or stochastic drifts. As such, the work is an important contribution and adds to the increasing body of literature that favors this more deterministic view. It also provides a potentially very useful and relatively straightforward method to a large number of researchers that are analyzing single-cell time traces of dynamic cellular responses, which may yield surprising insights, notably in the cell signaling field where the prevailing view still is that most variability is of a stochastic nature.

We thank the reviewer for his/her appreciation for the conceptual and methodological contribution of this work.

While the above is all positive, I do have some major criticisms, which I summarize below:

1. The overall impression of the paper is that it is a bit 'thin'. There is no experimental validation of the three states. The paper would significantly strengthen if such work would be included. Could the authors come up with measurements of the strength of positive feedback in single cells? Could they explore the use of IP3R inhibitors or agonists to dissect this? I am not an expert on calcium signaling, but it appears to me that many experimental tools are available for this system to obtain experimental evidence for the proposed mechanisms that underlie the three major cell states. With such data included, the paper would be much stronger.

We now provide experimental validation of key prediction that stems from our analysis. Our analysis identified IP3R activity as key parameter that we predict should explain the majority of calcium response variability. We now include data from new experiment to directly test this prediction (new panel Figure 5d). In this new experiment we performed longitudinal analysis where we measured for the each cells the response to extracellular ATP and intracellular increase in IP3. We show that ~70% of the response variability in calcium response magnitude is explained by the magnitude of the response to IP3. This new data validates the key prediction of the computational analysis regarding the key mechanisms that explain IP3 response as shown in Figure 5d.

2. As the authors state, they pre-process the raw time-traces before they attempt model fits, and they also somewhere mention low-pass filtering of the data. Clearly, this removes some of the variability in the data. An alternative approach is to ask how much of the total variability in the data (without any pre-processing) can be explained by the models.

We now include new expanded view panel (Expanded View Figure 2e) that show that our initial pre-processing removes only a small fraction of the observed variability.

And after removing that, is the amount of unexplained variability approaching the Poisson limit, plus some experimental noise? Can the amount of experimental noise in their setup be estimated?

The observed variability between cells was significantly higher than experimental variability. We now included new expanded view panel (EV Figure 3a) that quantify and compared the experimental variability (that our pre-processing is aimed to remove) and the observed cell to cell variability. As the process is dominated by cellular heterogeneity is kinetic parameters, the observed variability is much higher than would be expected by simple Poisson process with fixed parameters.

3. Finally, while the distinct cellular states support the more deterministic view, the authors provide no insight in what is the source of these different states? In theory these could still emerge from stochastic events occurring in growing cell populations which then become stabilized.

We agree with the reviewer and now added the following text in the discussion paragraph that discusses the three distinct IP3R mechanism:

“What causes different cells to have different IP3R responses is still unclear. Like many cases of “symmetry breaking” it is possible that separation of cells into the three clusters of IP3R activity start with stochastic events that get stabilized to create the emergence of three distinct populations”

The authors suggest from Figure 1f-g and extended view Figure 3 that it is not correlated with the spatial positioning of cells or with the amount of DNA in cells (used as a proxy for position in the cell cycle). This analysis is however underwhelming. Merely glancing at an image of a couple of

hundred cells does not provide any quantitative insight about whether there are correlations with local cell density, cell size, or cell shape. Furthermore, Hoechst staining can be tricky to interpret (depending on amount and time of incubation). Do the authors see two nice peaks of G1 and G2 cells in Hoechst intensity distributions? Providing more insight into what could be the sources that determine the 3 major cellular states would also be a major improvement of the manuscript.

We expanded the analysis and analyzed 12 different cellular features that include cell size, morphology, and local neighborhood. The analysis is shown in expanded figure 3b. We did not identify any simple factor that can explain the separation into three clusters. Any “negative result” in this type of analysis is inconclusive as it is always possible that we did not include the “right” feature. What causes cells to separate into the observed clusters is a key open question. And while a fascinating one, we consider it to be outside the scope of the current work.

Reviewer #2:

Yao & Wollman address an interesting and timely question: to what degree the experimentally observed single-cell variability results from purely stochastic biochemical events and/or whether discrete states exist in a population. They address it by measuring single-cell calcium responses to ATP and by further fitting the responses to a previously published model.

Overall, the methodology presented in the paper is a valuable addition to the field, however, the conclusions about the existence of discrete states in the parameter space are not convincing given the current presentation. Thus, before recommending the paper for publications the authors should address the following.

We thank the reviewer for his/her constructive comments.

Major issues:

1. The clustering of distance measure based on KL divergence shown in Fig.4 implies 3 distinctive groups in the parameter space. At the same time, authors conclude that hierarchical clustering of measured responses shown in EV Fig.1A doesn't yield good grouping. One issue I have with this claim is that just by looking at the heat-map/waterfall plot EV Fig.1A I can discern ~6 clusters, based mainly on the time to reach the final steady-state value. Regardless of authors' claim about the lack of "natural grouping" of experimental time courses, it is unclear to me why they insist there wouldn't be a (decent) agreement between explicit clustering of experimental time courses and clustering of model parameters fitted to those very same time courses. After all, the model with parameter sets from different clusters shown in Fig.4 should generate time courses different enough to classify/assign them to a couple of distinct clusters. If that wasn't the case, the robustness of fitting would be in question. In other words, distinct states in the parameter space that authors attribute to different feedback mechanisms should relate to distinct dynamic responses. This is partly addressed in Fig.4C but the "green" cluster contains all kinds of dynamic responses, from weak to strong activation. How does it relate to explicit clustering of measured responses (based on entire time courses or their features)?

We completely agree with the reviewer comments that:

1. That the use of “natural grouping” is subjective.
2. It is possible to find clusters in the hierarchical clusters based on calcium response alone.
3. That there should be some agreement between clustering based on calcium response and parameters that are fitted to calcium response.

To enhance our analysis we now address these three points specifically. We replaced the subjective last step that determines the cutoff within the hierarchical tree with an objective criterion that is

based on modified Calinski-Harabasz criteria. We apply this criterion for the clustering based on the inferred parameters and the clustering based on calcium response directly. We then compared the quality of the identified clusters based on two approaches based on a measure that looks at the ratio of variance between clusters divided by variance within cluster. We now show (Expanded Figure 1 c,d) that the clustering based on parameter is ~150 times better than clusters based on calcium response directly. Therefore, while it is possible to identify clusters in any hierarchical tree, the parameter inference we performed substantially increase the quality of the identified clusters.

Additionally to addressing the above concerns, I ask the authors to:

(a) provide a confusion matrix of clustering of raw (or smoothed) measured responses against classification of fitted parameters. How similar percentage wise are those two classifications?

The confusion matrix is now shown in Expanded View Figure 1e.

(b) expand Fig. 4 with a panel including time course features (according to Fig. 1D). Are those features distinct between 3 clusters?

The requested data is now shown in Expanded View Figure 4b. As pointed out by the reviewer, given that the three clusters have different underlying mechanism, and given that different mechanism should result in different time course features, one would expect to see that the time course features are distinct between three different clusters. This is indeed what we see as shown in Expanded View Figure 4b

(c) add labels designating feedback mechanism to the 3 clusters.

Labels that show how temporal features are distributed by cluster is now shown in Expanded View Figure 4b

2. Related to the above is the question about response dynamics within three clusters that were associated with different feedback strengths and types. For example, it is known that due to the negative feedback the dynamic responses are transient, while strong positive feedback can induce sustained responses. Thus, are responses belonging to the "Neg. Feedback" cluster more transient (decay faster) than those in strong and weak PF? If no such differences exist in time courses that belong to any of 3 clusters, then the mechanistic explanation of these clusters is questionable.

We clarify in the main text that we would expect to see, and indeed are seeing, differences in response dynamics between different clusters.

"As expected, each of the identified feedback structure had distinct temporal features associated with it (Expanded View Figure 4b)"

3. The following sentence is unclear:

"For every two parameter vector populations corresponding to each of two cells, the distance measure was defined as the KL divergence from a null model where both samples were obtained from identical probability distributions."

The sentence was rewritten to be:

"For every two cells we constructed a probability distribution of the likelihood of nearest neighbor parameter vector to be from the same cells. The parameter vector populations corresponding to each of two cells, the distance measure was defined as the KL divergence from the constructed distribution from a null model where the parameter vector sets for the two cells are both samples were fully mixed obtained from identical probability distributions."

Please rephrase/expand the definition of the distance. In particular, please comment on the validity of KLdiv usage in this context, given that KLdiv isn't a true metric, i.e. (usually) $KLdiv(a || b) \neq KLdiv(b || a)$.

As explained above, the KL distance was not between the two cells multivariate parameter distribution but between a probability distribution that capture the level of “mixing” between cells and a null model where the parameter distributions are fully mixed. This was done for two reasons: 1. as noted by reviewer, KL it not a true metric as it is not symmetric. 2. The dimensionality of the multivariate distribution was too high for direct comparison. The constructed statistics does capture the essence as it take into account the spread of parameters within cell as well as the separation between cells.

Also, the explanation in M&M - "Clustering of the Parameter Ensemble" is not clear and in its current form one can easily get lost in terminology. The entire section would greatly benefit from a scheme of the work-flow, or an itemised step-by-step explanation.

The section was completely revised for clarity.

4. Fig. 1E

The authors have excluded the influence of spatial localisation and cell cycle. However, part of the observed variability could still result from pre-existing cellular states. For example high basal levels (F_0) could imply high peak value (F_m). To further exclude those effects authors should provide scatter plots for pair-wise comparison of 6 main time course features. Are they correlated?

To clarify, we do not claim anywhere that calcium response dynamic is the same between the different clusters. This will be impossible as the clusters are based on kinetic parameters that are inferred from calcium response dynamics. Our key point, is that the parameters provide a better view of these differences that both shows higher degree of separation and is easier and more intuitive to analyze. We include new quantification that shows this point (Expanded View Figure 1b).

Minor issues:

1. Fig. 1A.

How many cells have been measured in a single experiment?

Added to figure legend.

What is the colour coding scale associated with this panel?

Colorbar included.

2. Fig. 1E.

The range of Y-axis scale makes it difficult to discern the boxes. Consider removing two outliers in the last two rows, or skipping the last two columns with ratios altogether.

Corrected.

Please add the equation used to calculate the Z value.

3. EV Fig. 1B

The visualisation doesn't seem to be appropriate. 6 features along the x-axis transition smoothly; it's hard to recognise that the axis is discrete.

Additional labels were added to clarify that the axis is discrete.

4. EV Fig.3

Current box plot visualisation makes it difficult to see whether G1, S, G2/M are indeed equally populated across three clusters. Thus, I suggest to change this plot or to add a panel with a more familiar cell cycle plot, i.e. line plot of the probability density of Hoechst intensity for 3 different feedback types/strengths.

We now include the Hoescht analysis together with additional 11 other features in a different panel.

5. Fig. 3D

What kind of normalisation was used? Area, count, width? Please describe in the caption. The sentence from the caption: "The width of each violin shows the probability density of the underlying distribution." is unclear.

Figure legend was extended to explain this point:

"The violin plot of all kinetic parameters from the fit results to a single cell calcium trajectory. The plot values are in terms of the log of 10 of the parameter values normalized according to the reference values determined either through literature or values within reasonable biological ranges. The meanings of the kinetic parameter symbols are listed in Table S1"

6. EV Fig.4

What does it mean that whiskers do not extend from the box? E.g. "eta1 (ER)" for WP.

What kind of box-plot is that? What do whiskers signify?

The confusion stems from a technical issue in figure export from Matlab to illustrator where whiskers had a dashed line. This was fixed.

7. Sentence, page 9, par. 1:

"To test if the nearest neighbours probability that we calculated shows a statistically significant difference..." -> difference

Corrected

8. Table S1 was missing from my copy of the manuscript.

Table S1 is now included.

9. Please explicitly define what "parameter vector" means.

We added the following definition to the methods section:

“A parameter vector is defined by the collection of parameters that are needed to fully specify the model shown in figure 2”

10. Since the model was simplified compared to the original, please deposit the model into BioModels database.

Model submission to BioModels database is in progress.

2nd Editorial Decision

11 November 2016

Thank you for sending us your revised manuscript. We have now heard back from reviewer #2 who agreed to evaluate the study. As you will see below, this reviewer thinks that the major issues have been satisfactorily addressed. However, s/he raises a remaining concern, which we would ask you to address in a minor revision.

REFeree REPORT

Reviewer #2:

The clarity of the manuscript has improved significantly and the authors have addressed all of my concerns, except one.

At the beginning of page 4 they write:

"We initially attempted to cluster the single cell calcium trajectories based on either the entire calcium responses or through their representative time series features. However, no natural grouping of cells was observed using these two simple cell response signatures (Expanded View Figure 1)..."

I analysed the supplied data (74919_1_additional_figure_data_1020082_jfjt5j.mat) and obtained very good results with one of the most commonly used Ward's linkage method. Also k-means clustering (conceptually similar to Ward's) produces good clustering with Calinski criterion = 3. The results of my analysis along with R script to generate the plots is in the online folder:

<https://www.dropbox.com/sh/436hd2ptnoeodvv/AACLuFaVYN3nnDTrWUavWszva?dl=0>

The authors should therefore rewrite this paragraph and the caption of EV Fig. 1. In its current form it leaves a reader with an impression that the only way to classify such dataset would be through a model-based approach.

I am not questioning this approach, however. It certainly gives a valuable mechanistic insight into the source of heterogeneous responses as evidenced by clustering of model parameters and the experiment. However, the presentation should account for the fact that it is in fact possible to cluster experimental time courses. Additionally it would be interesting to compare such clustering with 3 clusters based on model parameters. Calinski criterion calculated from clustered time courses hints at the existence of 3 clusters, which is also the number of clusters calculated for model parameters.

Once the authors address this issue, I am happy to recommend it for publication.

2nd Revision - authors' response

16 November 2016

Thank you for the opportunity to re-submit the paper after minor revisions. Below is point by point response to reviewer comments. For clarity we show reviewer comments in serif fonts italics and our response in san serif below.

Reviewer #2:

We thank the reviewer for his/her in depth analysis and for taking the time to reanalyze some of our data. The presented insights are meaningful and we are revising the paper as requested:

The authors should therefore rewrite this paragraph and the caption of EV Fig. 1. In its current form it leaves a reader with an impression that the only way to classify such dataset would be through a model-based approach.

We agree with reviewer #2 that cluster analysis is definitely possible based on timeseries data. To prevent any confusion and possible false impression that cluster based on timeseries is valid we edited the text as follows:

Main text:

We revised the statement that no natural clusters were found to the following:

The obtained clusters did not identify clear response archetypes (**Expanded View Figure 1**) potentially as a result of the fact that the ratio of variance between clusters to the intracellular variance was low (**Expanded View Figure 1d**). Therefore, while clustering based on timeseries data is a valid and useful approach, in this case it did not lead to identification of cell states. To further explore the possibility that cell states exist, we hypothesize that insights regarding cellular states could be uncovered by studying the mechanistic structure of pathway using mathematical modeling that incorporate prior knowledge on the underlying signaling network.

Legend of Expanded View Figure 1.

We removed the potentially confusing sentence: "The dendrogram shows that there was no clear clustering of the time course data."

3rd Editorial Decision

18 November 2016

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.

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Roy Wollman

Journal Submitted to: Molecular Systems Biology

Manuscript Number: MSB-16-7137

Reporting Checklist For Life Sciences Articles (Rev. July 2015)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures**1. Data****The data shown in figures should satisfy the following conditions:**

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if $n < 5$, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions**Each figure caption should contain the following information, for each panel where they are relevant:**

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/changed/perturbed in a controlled manner.
- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
 - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ^2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
 - are tests one-sided or two-sided?
 - are there adjustments for multiple comparisons?
 - exact statistical test results, e.g., P values = x but not P values < x;
 - definition of 'center values' as median or average;
 - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

Please ensure that the answers to the following questions are reported in the manuscript itself. We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

In the pink boxes below, provide the page number(s) of the manuscript draft or figure legend(s) where the information can be located. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable).

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	The sample size was chosen such that it is much higher than the number required to identify the distinct clusters.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	Samples were excluded only for technical reason based on preestablished criteria. Specifically, single cell data was fitted using San Diego Super Computer Cluster and in many cases the fitting software crashed and no fit was reached. Therefore the parameter values for these cells should be considered missing data and therefore the cells were excluded from the analysis.
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	All treatment were done on cells in random fashion so there was no bias in that selection.
For animal studies, include a statement about randomization even if no randomization was used.	NA
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	NA
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	Yes
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Our key statistical analysis is descriptive. We do not perform any hypothesis testing.
Is there an estimate of variation within each group of data?	Yes, variance analysis was done and shown in Expanded Figure.
Is the variance similar between the groups that are being statistically compared?	NA

C- Reagents**USEFUL LINKS FOR COMPLETING THIS FORM**

<http://www.antibodypedia.com>
<http://1degreebio.org>
<http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repo>

<http://grants.nih.gov/grants/olaw/olaw.htm>
<http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm>
<http://ClinicalTrials.gov>
<http://www.consort-statement.org>
<http://www.consort-statement.org/checklists/view/32-consort/66-title>

<http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tun>

<http://datadryad.org>

<http://figshare.com>

<http://www.ncbi.nlm.nih.gov/gap>

<http://www.ebi.ac.uk/ega>

<http://biomodels.net/>

<http://biomodels.net/miriam/>
<http://jij.biochem.sun.ac.za>
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
<http://www.selectagents.gov/>

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	No antibodies were used.
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	We used MCF10A cells from ATCC and a derived cell line that expresses mCherry-GCaMP.

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	NA
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	NA
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	NA

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	NA
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	NA
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	NA
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	NA
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	NA
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	NA
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	NA

F- Data Accessibility

18. Provide accession codes for deposited data. See author guidelines, under 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	NA
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	Source data files for all key figures is provided.
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. As far as possible, primary and referenced data should be formally cited in a Data Availability section. Please state whether you have included this section. Examples: Primary Data Wetmore KM, Deutschbauer AM, Price MN, Arkin AP (2012). Comparison of gene expression and mutant fitness in <i>Shewanella oneidensis</i> MR-1. Gene Expression Omnibus GSE39462 Referenced Data Huang J, Brown AF, Lei M (2012). Crystal structure of the TRBD domain of TERT and the CR4/5 of TR. Protein Data Bank 4O26 AP-MS analysis of human histone deacetylase interactions in CEM-T cells (2013). PRIDE PXD000208	NA
22. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as BioModels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	Model was deposited in SBML format to BioModels.

G- Dual use research of concern

23. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC (see link list at top right)). According to our biosecurity guidelines, provide a statement only if it could.	No
---	----