

RNAi screening reveals a large signaling network controlling the Golgi apparatus in human cells

Joanne Chia, Germaine Goh, Victor Racine, Susanne Ng, Pankaj Kumar, Frederic Bard

Corresponding author: Frederic Bard, Institute for Molecular and Cell Biology

Review timeline:

Submission date:	20 April 2012
Editorial Decision:	25 May 2012
Revision received:	20 August 2012
Editorial Decision:	04 October 2012
Revision received:	10 October 2012
Accepted:	11 October 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

25 May 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 had positive words for the goals of this work, they each felt that additional work would be needed to support these findings and to analyze the high-content data in more detail. In particular, the first reviewer feels that some of the most novel biological conclusions deserve additional experimental investigation and support (see particularly, points #1 and #6), and the second reviewer has important technical concerns regarding image analysis and SVM based phenotypic classification.

In addition, Molecular Systems Biology generally asks authors to release high-content data in the most unprocessed manner reasonable. Moreover, it does seem that data release is somewhat material to the concerns raised by the last reviewer -- future researchers are may want to use other feature sets or try other methods for assigned phenotypes, and could potentially uncover further insights from this rich dataset. I would ask that you provide the raw numeric data at least at the single cell level with this revision, and it would be useful to know the size of the underlying image data.

If you feel you can satisfactorily deal with 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.

Sincerely,
Editor - Molecular Systems Biology
msb@embo.org

Referee reports:

Reviewer #2 (Remarks to the Author):

In the manuscript "RNAi screening reveals a large signaling network controlling the Golgi apparatus in human cells" by Chia et al, the authors describe a RNAi screen of human kinases and phosphatases (and some additional molecules) on the structural integrity of the Golgi apparatus. After describing the screen procedure, the authors discuss potential effects of some groups of hits and finally end with describing effects on glycan processing. The manuscript is well written, the experiments and in general sound and the topic is within the scope of the journal. The analysis of the hits is very careful and innovative and I am certain that it will be of interest for the scientific community and therefore I support its publication in "Molecular Systems Biology". However, there are several points that need to be addressed.

Major points:

1- Diffuse Golgi morphology: if it reflects relocation to the ER, then a colocalization with an ER marker (PDI, CLIMP63, calnexin, etc...) should be performed for of at least the top five hits in this category. In addition, the authors speculate about ER-to-Golgi trafficking defects. This should be shown by using either VSVG-tsO45 or by performing pulse-chase experiments with endoH digestion. This should not be performed for all hits (as the authors anyway performed a secretion assay), but rather just for a couple of hits.

2- Some papers by colleagues have not been cited:

The recent work of Catherine Rabouille is not mentioned (Zacharogianni M et al, EMBOJ, 2011). This work also focuses on MAPK signaling to the early secretory pathway. In addition, it focuses on MAPK15 (ERK7), which is one of the hits from the current work. Thus, I feel that this work should be cited and discussed.

Discuss the work of Peter McPherson on SCYL1 and Golgi structure. The current work identified SCYL3 and its knockdown leads to a diffuse Golgi phenotype that is more prominent at the cis-Golgi. This might suggest that COPI is affected and such a link is supported by data on SCYL1 (Burman et al, PLoS One, 2010; Burman et al, J Biol Chem, 2008). However, SCYL3 does not seem to have a clear COPI interaction motif and this also should be discussed.

The authors identify PIK4CA as a hit that leads to a diffuse + fragmented Golgi. This is consistent with earlier work that showed that PIK4CA regulated ER exit sites (Farhan et al, EMBOJ, 2008). I noticed that this work is cited, but it should also be mentioned in this context (e.g. on p.12 in the last five lines).

3- The paper would benefit from some type of hit analysis such as functional annotation clustering (DAVID database; also now possible with the STRING database; BINGO analysis using cytoscape, etc...). This would show which signaling pathways or biologic processes are enriched among a certain group of hits. As it is now, it is difficult to state whether for instance hits that produce a condensed Golgi associate with a different set of signaling pathways than hits that cause the Golgi to fragment. In addition, the authors discuss several potential effects of certain groups of hits on the Golgi (e.g. MAPK in Figure 5D). Functional annotation clustering would tell whether the enrichment of MAPK family members is significant. This type of analysis is not very difficult and yields information that is easily accessible to many readers of this journal.

4- The experiment shown in Figure S4C is intended to demonstrate that there is a sorting defect in BUB1 knockdown cells. This statement is in my opinion not backed-up by the result shown, which only shows that some proteins are secreted better than others, but it is not clear whether this is due to sorting or to effects on expression of these proteins or due to increased/decreased degradation etc....

Either this result should be removed, or it should be replaced with data that are clearer (e.g. mis-sorting of apical and basolateral proteins, or mis-sorting during endocytic recycling, etc...).

5- On p.15, the authors hypothesize that Cdk1 and BUB1 "may function together at the TGN". Directly after this sentence, they state that Cdk1 has been shown to phosphorylate GRASP65. However, GRASP65 is a cis-Golgi protein and I am not aware that it localizes to the TGN. If I am wrong then the authors should cite the work that shows GRASP65 at the TGN. If not, then it should be made clear that GRASP65 is not at the TGN.

6- Six sections in the Results and two figures are dedicated to discuss possible effects of a certain group of hits on the Golgi (MAPK, acto-myosin, cell cycle, growth factors, etc...). In general, these discussions are very insightful and clearly written. However, they are mostly speculative and would need to be verified. I think that it is important that the authors should at least back-up some of their claims with experimental evidence. One suggestion is the following:

If DUSPs are negative regulators of ERK signaling, then a knockdown of (for instance) DUSP6 should increase ERK activity. This should be shown. Furthermore, this should be linked to Golgi structure by showing the the effect of DUSP6 depletion on Golgi structure can be rescued by a co-knockdown of ERK or by pharmacologic inhibition of ERK signaling. Finally, it should be shown that the increased ERK signaling in DUSP6 knockdown cells leads to a higher phosphorylation of GRASP65.

This is of course just a suggestion and the authors could just pick any of these speculative links (like the link between mitotic kinases and post-Golgi trafficking) and support it with experimental evidence.

7- A major point of the paper is the investigation of effects on glycan expression. A model is proposed whereby signaling pathways affect the functional organization of the Golgi and thereby affect glycan expression. While this model appears safe, it is in my opinion incomplete. Effects on glycosylation will also affect glycosylation of cell surface receptors and this would alter their expression and function and thereby affect signaling. This should be discussed and an arrow pointing back from "Glycan expression" towards "Signaling cascades" should be included into Figure 8B.

Minor points:

1- The authors state that "genes with a significant effect on cell number" were excluded. This statement is vague. It would be good if a number could be mentioned here.

2- In Figure 4C, DUSP6 is displayed in the nucleus. This is not correct. The majority of DUSP6 is in the cytosol making it one of the most prominent cytosolic DUSPs.

3- On p.12, upper paragraph: The section is not very clearly explained (at least I did not really get the point). As far as I understood, the authors found 854 proteins annotated to the Golgi and of these 413 were phosphoproteins. This appears clear. However, the next selection step is not so clear. How did the authors narrow down the 413 phosphoproteins of the Golgi to 135? They say that these were related to membrane traffic. What Golgi proteins were excluded that are unrelated to membrane traffic? This selection process is not completely clear and requires more details so in order to become more transparent.

4- On p.14: it the citation of Figure 7 is not correct and should be replaced by Figure S5.

5- On p.15: it the citation of Figure S4D is not correct and should be replaced by Figure S4C.

6- When discussing the effect of mitotic kinases the authors state at the bottom of p.15: "Overall, this sub-network suggests an intriguing link between cell cycle kinases and post-Golgi traffic (Figure 5C)." Maybe it should be stated here that these effects are all in a non-mitotic context, unless of course the authors speculate about a link to mitosis. In the latter case, this should also be clearly stated.

Reviewer #3 (Remarks to the Author):

In this manuscript the authors perform a kinome scale high-content RNAi screening using HeLa

cells to identify genes that regulate Golgi architecture. Three markers were used to illustrate different compartments of Golgi (cis-, medial- and trans-) and a straightforward computational workflow was constructed. Automatic segmentation were carried out for each compartment, SVMs were trained to assign one of three typical phenotypes (Diffused, Fragmented, and Condensed) to each Golgi compartment of a single cell, and a vector was generated to quantify cellular response to certain treatment by summarizing the ratio of cells in different phenotypes, clustering analysis and network mapping were then carried out to interpret the screening results. As a result, the authors discovered a large signaling network controlling the Golgi apparatus in human cells. Understanding how signaling networks regulates the morphogenesis of organelles is a fundamental question in biology, and there is desperate need for these types of studies. In many regards this work is of good quality and the findings are potentially interesting. However I feel there are a number of technical concerns prevent this manuscript being published in *Molecular Systems Biology* in its current form.

The biggest issues here are that due to 1) the resolution they are using, 2) the segmentation and feature extraction (e.g. it is not clear they have the right features for Golgi analysis), and 3) the classification techniques (only three reference phenotypes, 2 of which are highly correlated) they do not achieve credible discriminatory power in this screen. I believe they are identifying new direct/indirect players in regulation of Golgi, but compared to what could have been done visually and/or using a simpler analysis tools - it does not look like they are describing a significant advance in high-content analysis. With more advance imaging and classification tools, the authors might be better able to sub-divide different mutant phenotypes and make judgments regarding gene function.

Major-

1. Cell segmentation and image analysis

- The Nuclei of HeLa cells have limited roundness, thus for the definition of cellular area, the application of "a circular regions of 45 microns" on (only) the nuclei channel may bring the issue of under-segmentation and lead to heavy burden for image quality control-especially when it is unclear whether all cells are as nicely spaced out as in Fig 2C (some more crowded images are presented). If the authors wish to rely on this strategy, it may be helpful to favor the direction of major axis of nuclei when defining the region; or information from channels other than nuclei can be introduced to help obtain better segmentation.

2. More features may be necessary for proper classification of the three reference phenotypes

-In Figure S2D, The z-scores for "Top" features in C-classification are low (top out at 0.29). Is it not correct that C phenotype have less "number of objects" or higher "intensity inside the object"?

-In fact the definition of "object" is not exactly clear to me, given the notion like Obj[1] and Obj[2] in Figure S2D (as well as "the number of objects" mentioned in Supplementary), did the authors consider each scattered cluster of high intensity pixels as a single object? Then phenotype D and F can easily have 5-10 such pixel clusters for each cell. Or did the Obj[1] and Obj[2] represent "small objects" and "bigger objects" mentioned in Supplementary? This should be further clarified.

3. Number of phenotypes available in the dataset.

The authors seem to agree that these three phenotypes D, F, and C may contain sub-populations. It would be beneficial to do a clustering of cells and see whether more reasonable sub-populations/phenotypes exists. Then again more comprehensive feature set would be necessary to obtain more information.

4. The usage of SVM

-For feature selection, a "filtering/ranking" method based on separation of z-scores for each single feature was used, and an arbitrary number of features (5) was recruited for almost all SVM on quality control and phenotype classification. This is way too simple: judging from Figure S2D each classification has a quite unique scenario (D-classification has two dominant features, F-classification has a series of so-so features while C-classification has a bunch of very weak features) and thus may require different sizes of feature sets. Some cross-validation should be done on different combinations of top-ranked features, and the number of features should be selected based on the tradeoff of performance vs. set-size. Given that linear SVMs are used for phenotype classification, Recursive Feature Elimination (RFE) may be a reasonable choice.

-It may not be particularly favorable to apply an SVM with only one feature, especially when no proof was shown on the ability of other (larger) feature combinations.

-The performance of linear SVM in phenotype classification is not specifically impressive on training set, as shown in Figure S2G. Again the quality of feature set is a concern, and other type of kernels for SVM may be necessary. Also, there should be a table showing each SVM's performance in cross-validation on training sets.

-Figure S2B shows a strikingly high rate of "moderate agreement" in single cell level among experts

(red bar), I wonder did all the training cells (i.e. the labels of control vs phenotype or red vs blue in Figure S2G) have a "complete agreement" among experts? If yes, then the results from Figure S2G become even less impressive; if no, then the quality of the reference cell set is questionable.

-When it comes to phenotype classification, all classifiers are trained to separate "one phenotype vs control", which is vulnerable to reliability of training cells. "One phenotype vs All others" may be helpful to tackle that overlap between D and F phenotype.

5. Penetrance of the population, and the blank area of the circles in Figure 4C.

-The authors confirmed in page 10 that "scores for three markers does not necessarily correlate." As an illustration, the area of three blank regions within one circle in Figure 4C may not be the same (e.g. SCYL3 in Figure 3D, which is also highlighted in the text). Thus I suggest the addition of a "penetrance" score into the final vector for clustering, which corresponds to the ratio of normal cells (negative for all nine classifiers) for each gene. This may increase the resolution of the analysis and create more meaningful clusters than those in Fig 3C.

-This is also part of the reason that we suggest a "One vs All others" classifier, it is understandable but a little tricky when one cell is "normal for two markers while abnormal for the other."

Minor-

1. The proposed network analysis is not particularly informative. It is not surprising that they would identify a number of phosphosites in kinases and phosphatases, and without doing a more sophisticated method of network inference, I am not sure how useful their network is.

2. For phenotype D and F, HPL dyes spread wider than the other two, and nearly depicted the whole cell body in the case of D phenotype in Figure 1C, which gives phenotypes in HPL channels a quite different look. Is it biologically significant or simply artifact due to characteristics of dye?

3. Exactly what clustering parameters are used for generating Figure 3C? (the clustering method, similarity measurement, similarity cutoff for defining clusters etc.)

4. There are inconsistent claims in the manuscript, for example:

-p11, last line, "111 hit kinases"->Are these 111 kinases part of the 181 in p10? or 159 in p10? or 61+53 in p11?

-the "glycan biosynthesis" section in p18 and p19,

is it 146 (p18, title) or 166 (p19, 6th line from bottom);

is it 180 (p19, 6th line from bottom) or 181 (claims in p10)

-p14, 6th line from bottom, "(Figure 7)"->exactly where in Figure 7?

5. Some typos exist

-p7, 8th line from bottom, the medial Golgi is missed in claiming of stained channels/compartments

-p24, 5th line from bottom, "GalNacT"->it was "GalNac-T" in p6

-p25, 3rd line from bottom, "controlling directly Golgi physiology"->rather strange ordering of words

-the sentence crossing p9 and p10, "we discarded from further analysis genes with..."-> rather strange ordering of words

-Supplementary for SVM design, why choose a cost parameter (28) which "minimizes" the accuracy average of the 3 PF?

1st Revision - authors' response

20 August 2012

Reviewer #2 (Remarks to the Author):

In the manuscript "RNAi screening reveals a large signaling network controlling the Golgi apparatus in human cells" by Chia et al, the authors describe a RNAi screen of human kinases and phosphatases (and some additional molecules) on the structural integrity of the Golgi apparatus. After describing the screen procedure, the authors discuss potential effects of some groups of hits and finally end with describing effects on glycan processing. The manuscript is well written, the experiments and in general sound and the topic is within the scope of the journal. The analysis of the hits is very careful and innovative and I am certain that it will be of interest for the scientific community and therefore I support its publication in "Molecular Systems Biology". However, there are several points that need to be addressed.

We wish to thank reviewer #2 for his/her time, detailed reading of our manuscript and constructive comments. Please find below the answers to the points raised.

Major points:

1- Diffuse Golgi morphology: if it reflects relocation to the ER, then a colocalization with an ER marker (PDI, CLIMP63, calnexin, etc...) should be performed for of at least the top five hits in this category. In addition, the authors speculate about ER-to-Golgi trafficking defects. This should be shown by using either VSVG-tsO45 or by performing pulse-chase experiments with endoH digestion. This should not be performed for all hits (as the authors anyway performed a secretion assay), but rather just for a couple of hits.

We performed the suggested experiment and indeed, diffuse morphology corresponds to ER morphology as we now show in Figure S3. We have added the following paragraph in the main text: Page 11:

“Diffuse morphology is reminiscent of an ER pattern, suggesting that the marker displaying this morphology has been relocated at least partially to the ER. When we co-stained some of the diffuse hits with the ER marker calnexin, we could observe and quantify significant increase in colocalization indeed (Figure S3). Therefore, the diffuse morphology is likely to reveal a perturbation in distribution between ER and Golgi compartment.”

We have also performed a VSV-G transport assay for some diffuse hits and found that in most cases, this does show a significant reduction in traffic.

Page 19: “We further verified the secretion defect for nine hits using the well-established tsO45G transport assay {Zilberstein, 1980, r04446}. By probing the colocalization of tsO45G with the Golgi marker MannII-GFP 15 mins after release from the restrictive temperature (which induces retention of tsO45G in the ER), we could assess more specifically the ER to Golgi trafficking step. In agreement with the Met-Luc secretion data, depletion of six genes induced a significant reduction in ER to Golgi traffic (Figure S9A,B).”

2- Some papers by colleagues have not been cited:

The recent work of Catherine Rabouille is not mentioned (Zacharogianni M et al, EMBOJ, 2011). This work also focuses on MAPK signaling to the early secretory pathway. In addition, it focuses on MAPK15 (ERK7), which is one of the hits from the current work. Thus, I feel that this work should be cited and discussed.

Indeed, the work of Zacharogianni and colleagues is relevant to our study. We have now included the reference.

Page 16: “This is consistent with the results of two recent screens looking at ER exit sites in mammalian cells and Golgi morphology in Drosophila S2 cells {Farhan, 2010, p01667} {Zacharogianni, 2011, p11187}.”

Discuss the work of Peter McPherson on SCYL1 and Golgi structure. The current work identified SCYL3 and its knockdown leads to a diffuse Golgi phenotype that is more prominent at the cis-Golgi. This might suggest that COPI is affected and such a link is supported by data on SCYL1 (Burman et al, PLoS One, 2010; Burman et al, J Biol Chem, 2008). However, SCYL3 does not seem to have a clear COPI interaction motif and this also should be discussed.

Although this appeared somewhat peripheral to the main story in our first draft, this is a good suggestion and we have now included the following in the discussion:

Page 23: “In recent years, it has been well established that a close homolog, SCYL1, binds the COPI coat and regulates both retrograde traffic and Golgi morphology {Burman, 2008, r02518} {Burman, 2010, r02513}. SCYL2, alias CVAK104, has been proposed to mediate clathrin coated vesicles formation at the TGN {Duwel, 2006, r02529}. Together, these data suggest that the SCY-1 Like family of catalytically inactive protein kinases play similar roles in regulating membrane traffic. To note, neither SCYL2 nor SCYL3 contain the COPI binding site identified in SCYL1 {Burman, 2008, r02518}.”

The authors identify PIK4CA as a hit that leads to a diffuse + fragmented Golgi. This is consistent with earlier work that showed that PIK4CA regulated ER exit sites (Farhan et al, EMBOJ, 2008). I noticed that this work is cited, but it should also be mentioned in this context (e.g. on p.12 in the last five lines).

According to this recommendation, we have included the following sentence, Page 12:
 “Specifically, PIK4CA appears to regulate PIP4 at the level of ER exit sites as recently shown {Farhan, 2008, p08026}, while PIK4CB appears to function at the Golgi itself {Godi, 1999, r02594}.”

3- The paper would benefit from some type of hit analysis such as functional annotation clustering (DAVID database; also now possible with the STRING database; BINGO analysis using cytoscape, etc....). This would show which signaling pathways or biologic processes are enriched among a certain group of hits. As it is now, it is difficult to state whether for instance hits that produce a condensed Golgi associate with a different set of signaling pathways than hits that cause the Golgi to fragment. In addition, the authors discuss several potential effects of certain groups of hits on the Golgi (e.g. MAPK in Figure 5D). Functional annotation clustering would tell whether the enrichment of MAPK family members is significant. This type of analysis is not very difficult and yields information that is easily accessible to many readers of this journal.

We agree with the reviewer’s assessment on this point. Indeed, we had already included a KEGG pathway analysis in Fig. S4B, mentioned page 16 of the manuscript.
 “Our screen’s results indicate a significant enrichment in the MAPK signaling pathway (Figure S4B).”

This analysis reveals that the enrichment in the MAPK signaling pathway is indeed highly significant. The analysis using different databases such as DAVID only yields marginal differences, so we did not include it in the results.

4- The experiment shown in Figure S4C is intended to demonstrate that there is a sorting defect in BUB1 knockdown cells. This statement is in my opinion not backed-up by the result shown, which only shows that some proteins are secreted better than others, but it is not clear whether this is due to sorting or to effects on expression of these proteins or due to increased/decreased degradation etc.... Either this result should be removed, or it should be replaced with data that are clearer (e.g. mis-sorting of apical and basolateral proteins, or mis-sorting during endocytic recycling, etc...).

The conclusion of a sorting defect in BUB1 KD conditions is based on the facts that the secretion of some proteins is increased, while others are decreased. This phenotype is not consistent with a general defect in protein expression or degradation. Yet, at this is indeed not a major point of the paper, we have removed the figure and associated comments.

5- On p.15, the authors hypothesize that Cdk1 and BUB1 "may function together at the TGN". Directly after this sentence, they state that Cdk1 has been shown to phosphorylate GRASP65. However, GRASP65 is a cis-Golgi protein and I am not aware that it localizes to the TGN. If I am wrong then the authors should cite the work that shows GRASP65 at the TGN. If not, then it should be made clear that GRASP65 is not at the TGN.

We were aware of this inconsistency, hence our initial sentence relative to this point:
 “However, it is not clear whether this is related to the phenotype we observe.”

We have now replaced it with:

“However, it is not clear whether this is related to the phenotype we observe as GRASP65 is not known to localize at the TGN.”

6- Six sections in the Results and two figures are dedicated to discuss possible effects of a certain group of hits on the Golgi (MAPK, acto-myosin, cell cycle, growth factors, etc...). In general, these discussions are very insightful and clearly written. However, they are mostly speculative and would need to be verified. I think that it is important that the authors should at least back-up some of their claims with experimental evidence. One suggestion is the following: If DUSPs are negative regulators of ERK signaling, then a knockdown of (for instance) DUSP6 should increase ERK activity. This should be shown. Furthermore, this should be linked to Golgi structure by showing the effect of DUSP6 depletion on Golgi structure can be rescued by a co-knockdown of ERK or by

pharmacologic inhibition of ERK signaling. Finally, it should be shown that the increased ERK signaling in DUSP6 knockdown cells leads to a higher phosphorylation of GRASP65. This is of course just a suggestion and the authors could just pick any of these speculative links (like the link between mitotic kinases and post-Golgi trafficking) and support it with experimental evidence.

We have followed the reviewer's suggestions and further confirmed three of the regulatory modules we proposed.

The first one is for the ROCK1-PAK1 signaling module:

Page 15:

“Treatment of cells with IPA3, a PAK1 inhibitor {Deacon, 2008, r04449}, also induced fragmentation of the Golgi (Figure S5A) after six hours of treatment. The effect was dose dependent (Figure S5B). In agreement with our model of opposite action of ROCK1 and PAK1 at the Golgi, IPA3 could rescue at least partially the effect of ROCK1 knockdown while having no effect on PAK1 depleted cells (Figure S5C).”

The second one is for the Golgi fragmentation induced by the DUSP proteins. We could verify that indeed DUSP depletion leads to activation of the ERK kinase. Furthermore, this mediates at least in part the fragmentation of the Golgi apparatus as demonstrated by co-treatment with an Erk inhibitor. Unfortunately, for technical reasons (mostly the lack of specific phospho-GRASP65 antibody), we were unable to test whether GRASP65 is hyper-phosphorylated in conditions of DUSP depletion.

Page 17:

“We verified that depletion of DUSP2, 6 and especially 8 resulted in a hyper-phosphorylation of ERK1/2 (Figure S6A). This suggests that fragmentation of the Golgi apparatus results from the activation of ERK. Indeed, treatment of the DUSP2, 6 or 8 depleted cells with the ERK inhibitor FR180204 reverted the Golgi phenotype (Figure S6B, C). ERK has been shown to phosphorylate the Golgi structural protein GRASP65 during the orientation of the Golgi towards the leading edge {Bisel, 2008, p02418}. This phosphorylation event could be one of the underlying mechanisms of the observed Golgi fragmentation.”

The third is for another MAPK cascade, this time related to JNK:

Page 17:

“Indeed, one of our hits, MECOM, aka Evi1, has been proposed to negatively regulate the JNK by direct binding {Kurokawa, 2000, r04413}. JNK has been linked to Golgi associated proteins, such as AKRL1/2 {Harada, 2003, r04408} or vesicular trafficking proteins such as JIP1/2/3 {Akhmanova, 2010, p04808}. Consistent with the effect of Evi1 depletion being mediated by JNK, we find that treatment of the Evi1 depleted cells with the JNK inhibitor SP600125 reverts the Golgi condensed phenotype (Figure S7A,B).”

7- A major point of the paper is the investigation of effects on glycan expression. A model is proposed whereby signaling pathways affect the functional organization of the Golgi and thereby affect glycan expression. While this model appears safe, it is in my opinion incomplete. Effects on glycosylation will also affect glycosylation of cell surface receptors and this would alter their expression and function and thereby affect signaling. This should be discussed and an arrow pointing back from "Glycan expression" towards "Signaling cascades" should be included into Figure 8B.

We agree with the reviewer's remark and have modified the figure accordingly. We also modified the corresponding text (now in the Discussion section).

“The glycosylation of cell surface receptors is known to affect their stability and signaling potential {Boscher, 2011, r01175}. Therefore, it is likely that the regulation of glycan expression will in turn impact signaling cascades (Figure 8B).”

Minor points:

1- The authors state that "genes with a significant effect on cell number" were excluded. This statement is vague. It would be good if a number could be mentioned here.

We inserted the following: (less than 200 nuclei detected, while average number for hit genes is 526)

2- *In Figure 4C, DUSP6 is displayed in the nucleus. This is not correct. The majority of DUSP6 is in the cytosol making it one of the most prominent cytosolic DUSPs.*

We removed the nucleus in figure 4C as other predicted nuclear localizations were also not correct.

3- *On p.12, upper paragraph: The section is not very clearly explained (at least I did not really get the point). As far as I understood, the authors found 854 proteins annotated to the Golgi and of these 413 were phosphoproteins. This appears clear. However, the next selection step is not so clear. How did the authors narrow down the 413 phosphoproteins of the Golgi to 135? They say that these were related to membrane traffic. What Golgi proteins were excluded that are unrelated to membrane traffic? This selection process is not completely clear and requires more details so in order to become more transparent.*

To clarify our approach, we modified the text as follow as well as the figure.

“To test this, we conducted a systematic search for proteins with a Gene Ontology (GO) Cellular Component (CC) containing the term “Golgi” and found 854 proteins. In the PhosphositePlus database, almost half (413) of these Golgi associated proteins were found to carry at least one phosphorylated residue (Figure 4B). Additionally, to ensure that the network generated would be as stringent as possible, the additional filter of “membrane trafficking” GO Biological Process (BP) term was applied, resulting in 135 out of 413 proteins being retained. Hence 135 proteins are annotated in databases to be localized at the Golgi, to be phosphorylated and to regulate membrane traffic.”

4- *On p.14: it the citation of Figure 7 is not correct and should be replaced by Figure S5.*
Done.

5- *On p.15: it the citation of Figure S4D is not correct and should be replaced by Figure S4C.*
Done.

6- *When discussing the effect of mitotic kinases the authors state at the bottom of p.15: "Overall, this sub-network suggests an intriguing link between cell cycle kinases and post-Golgi traffic (Figure 5C)." Maybe it should be stated here that these effects are all in a non-mitotic context, unless of course the authors speculate about a link to mitosis. In the latter case, this should also be clearly stated.*

Following sentences added:

“Overall, this sub-network suggests an intriguing link between cell cycle kinases and post-Golgi traffic (Figure 5C) that may have to do with the dramatic changes in cell morphology and surface-to-volume ration observed at the onset of mitosis. Whether or not this link is related to the Golgi fragmentation observed during mitosis is hard to establish at present: judging by their DNA, the cells are not arrested in mitosis and their Golgi phenotype appears different from a mitotic Golgi, which would be clearly fragmented: for BUB1 the phenotype is mostly diffuse cis and trans and for CDC2 it is a mixture of diffuse and fragmented (Figure 5C).”

Reviewer #3 (Remarks to the Author):

In this manuscript the authors perform a kinome scale high-content RNAi screening using HeLa cells to identify genes that regulate Golgi architecture. Three markers were used to illustrate different compartments of Golgi (cis-, medial- and trans-) and a straightforward computational workflow was constructed. Automatic segmentation were carried out for each compartment, SVMs were trained to assign one of three typical phenotypes (Diffused, Fragmented, and Condensed) to each Golgi compartment of a single cell, and a vector was generated to quantify cellular response to certain treatment by summarizing the ratio of cells in different phenotypes, clustering analysis and network mapping were then carried out to interpret the screening results. As a result, the authors discovered a large signaling network controlling the Golgi apparatus in human cells.

Understanding how signaling networks regulates the morphogenesis of organelles is a fundamental question in biology, and there is desperate need for these types of studies. In many regards this work is of good quality and the findings are potentially interesting. However I feel there are a number of technical concerns prevent this manuscript being published in Molecular Systems Biology in its current form.

The biggest issues here are that due to 1) the resolution they are using, 2) the segmentation and feature extraction (e.g. it is not clear they have the right features for Golgi analysis), and 3) the classification techniques (only three reference phenotypes, 2 of which are highly correlated) they do not achieve credible discriminatory power in this screen. I believe they are identifying new direct/indirect players in regulation of Golgi, but compared to what could have been done visually and/or using a simpler analysis tools - it does not look like they are describing a significant advance in high-content analysis. With more advance imaging and classification tools, the authors might be better able to sub-divide different mutant phenotypes and make judgments regarding gene function.

We wish to thank the reviewer for her/his extensive review of our manuscript and the techniques we used.

Major-

1. Cell segmentation and image analysis -

The Nuclei of HeLa cells have limited roundness, thus for the definition of cellular area, the application of "a circular regions of 45 microns" on (only) the nuclei channel may bring the issue of under-segmentation and lead to heavy burden for image quality control-especially when it is unclear whether all cells are as nicely spaced out as in Fig 2C (some more crowded images are presented). If the authors wish to rely on this strategy, it may be helpful to favor the direction of major axis of nuclei when defining the region; or information from channels other than nuclei can be introduced to help obtain better segmentation.

The reviewer is correct to state that the segmentation strategy is somewhat limited by the use of the nuclear channel information only. However, our goal was to obtain a multidimensional view of the Golgi apparatus, hence the use of three imaging channels for this organelle. Using Golgi related channels for segmentation would have compounded the analysis for Golgi morphological effects. At any rate, as shown in Figure S2C, our automated image analysis is in near perfect agreement with an expert visual evaluation of the phenotype at the well level. At the individual cell level, (Figure S2B) the agreement is less good because of borderline phenotypes but it is completely comparable to user-to-user agreement.

2. More features may be necessary for proper classification of the three reference phenotypes -In Figure S2D, The z-scores for "Top" features in C-classification are low (top out at 0.29). Is it not correct that C phenotype have less "number of objects" or higher "intensity inside the object"? -In fact the definition of "object" is not exactly clear to me, given the notion like Obj[1] and Obj[2] in Figure S2D (as well as "the number of objects" mentioned in Supplementary), did the authors consider each scattered cluster of high intensity pixels as a single object? Then phenotype D and F can easily have 5-10 such pixel clusters for each cell. Or did the Obj[1] and Obj[2] represent "small objects" and "bigger objects" mentioned in Supplementary? This should be further clarified.

As stated above, our bench-marking against human experts indicate that our classification strategy is robust and as accurate as a human being.

While it is correct that individual features alone are not very effective at classifying Golgi morphologies (as shown in Figure 5D and 5F), the combination of 5 features in an SVM is highly efficient, as shown in Figure 5G.

3. Number of phenotypes available in the dataset.

The authors seem to agree that these three phenotypes D, F, and C may contain sub-populations. It would be beneficial to do a clustering of cells and see whether more reasonable sub-

populations/phenotypes exists. Then again more comprehensive feature set would be necessary to obtain more information.

Although we agree that more phenotypes may exist than the three we defined, it remains unclear how we could validate subtler phenotype definition. In our experience, bench-marking with human users hits its limits when we go beyond these three consensual phenotypes (i.e. two persons have a hard time agreeing on sub-population definitions). Furthermore, it is not completely clear that finer-grained phenotype definition would necessarily result in deeper insight in the biology of the Golgi apparatus.

4. The usage of SVM

-For feature selection, a "filtering/ranking" method based on separation of z-scores for each single feature was used, and an arbitrary number of features (5) was recruited for almost all SVM on quality control and phenotype classification. This is way too simple: judging from Figure S2D each classification has a quite unique scenario (D-classification has two dominant features, F-classification has a series of so-so features while C-classification has a bunch of very weak features) and thus may require different sizes of feature sets. Some cross-validation should be done on different combinations of top-ranked features, and the number of features should be selected based on the tradeoff of performance vs. set-size. Given that linear SVMs are used for phenotype classification, Recursive Feature Elimination (RFE) may be a reasonable choice.

It is indeed possible that using only two features might have sufficed for classification of the diffuse phenotype. However, for simplicity sake, we constrained ourselves to 5 features for each phenotype. As previously noted, this strategy was effective for classifying the three different phenotypes as we amply verified.

-It may not be particularly favorable to apply an SVM with only one feature, especially when no proof was shown on the ability of other (larger) feature combinations.

We have only used a combination of the 5 best features for the three SVMs used.

-The performance of linear SVM in phenotype classification is not specifically impressive on training set, as shown in Figure S2G. Again the quality of feature set is a concern, and other type of kernels for SVM may be necessary. Also, there should be a table showing each SVM's performance in cross-validation on training sets.

-Figure S2B shows a strikingly high rate of "moderate agreement" in single cell level among experts (red bar), I wonder did all the training cells (i.e. the labels of control vs phenotype or red vs blue in Figure S2G) have a "complete agreement" among experts? If yes, then the results from Figure S2G become even less impressive; if no, then the quality of the reference cell set is questionable.

In the experiment in Figure 2B, cells were not selected individually but taken from wells previously labeled with a given phenotype. A significant level of variability exists at the individual cell level and phenotypes tend to follow a continuous distribution, which explains why some cells may be classified "normal" by one expert and "fragmented" by another. As noted in Figure in 2C, these differences tend to even out at the well level, resulting in a reliable classification.

-When it comes to phenotype classification, all classifiers are trained to separate "one phenotype vs control", which is vulnerable to reliability of training cells. "One phenotype vs All others" may be helpful to tackle that overlap between D and F phenotype.

This is a good suggestion. However the goal of our analysis was to obtain a quantitative detection of perturbed phenotypes. It is not absolutely critical at this stage to accurately separate "diffuse" from "fragmented" phenotypes, since the functional clustering based on phenotypes was not extremely informative. However, we will surely take this into account for the analysis of the genome-wide screen in due time.

5. Penetrance of the population, and the blank area of the circles in Figure 4C. -The authors confirmed in page 10 that "scores for three markers does not necessarily correlate." As an

illustration, the area of three blank regions within one circle in Figure 4C may not be the same (e.g. SCYL3 in Figure 3D, which is also highlighted in the text). Thus I suggest the addition of a "penetrance" score into the final vector for clustering, which corresponds to the ratio of normal cells (negative for all nine classifiers) for each gene. This may increase the resolution of the analysis and create more meaningful clusters than those in Fig 3C.

The reviewers are correct that the blank area of phenotypic disks correspond to the penetrance of a particular treatment. While this is an interesting feature, it is also hard to really exploit as it may depend on gene knockdown efficiency as much as on the functional importance of the gene considered.

-This is also part of the reason that we suggest a "One vs All others" classifier, it is understandable but a little tricky when one cell is "normal for two markers while abnormal for the other."

Actually this was part of the rationale of the screen: the Golgi is a composite organelle and its different parts can be regulated independently. The results prove that indeed some genes knock-down affect only one marker, i.e. one compartment of the Golgi apparatus, but not the other two.

Minor-

1. The proposed network analysis is not particularly informative. It is not surprising that they would identify a number of phosphosites in kinases and phosphatases, and without doing a more sophisticated method of network inference, I am not sure how useful their network is.

The phosphosites were identified in Golgi associated proteins not in kinases and phosphatases. This result indicates an abundance of potential substrates at the Golgi for the kinases and phosphatases we identified. More in depth network analysis could possibly be obtained in the future by more specialized teams.

2. For phenotype D and F, HPL dyes spread wider than the other two, and nearly depicted the whole cell body in the case of D phenotype in Figure 1C, which gives phenotypes in HPL channels a quite different look. Is it biologically significant or simply artifact due to characteristics of dye?

Indeed, while the other two are single proteins, HPL is a lectin that binds a specific sugar present on numerous proteins. HPL staining intensity is also more variable than the other two markers, reflecting the regulation of O-GalNAc glycosylation that we previously reported (Gill et al. 2010).

3. Exactly what clustering parameters are used for generating Figure 3C? (the clustering method, similarity measurement, similarity cutoff for defining clusters etc.)

The following description has been added to the M&M section:

“The ‘Agglomerative Hierarchical Clustering’ method was applied using ‘Euclidean Distance’ as distance metric and ‘Complete Linkage’ as the linkage criteria. We picked the clusters manually by visual analysis of heatmap and thus identified 6 different groups of genes.”

4. There are inconsistent claims in the manuscript, for example:

-p11, last line, "111 hit kinases"->Are these 111 kinases part of the 181 in p10? or 159 in p10? or 61+53 in p11?

Indeed, the 111 hit kinases are part of the 181 hits that contain also phosphatases and kinases associated genes. We have clarified this point in the text.

“In fact, mapping the 111 hit kinases (out of 181 primary hits) from the screen on a phylogenetic tree of kinases {Manning, 2002, p08884} reveals that all the major families are involved in Golgi regulation (Figure 4A).”

-the "glycan biosynthesis" section in p18 and p19, is it 146 (p18, title) or 166 (p19, 6th line from bottom); is it 180 (p19, 6th line from bottom) or 181 (claims in p10)

We thank the reviewer for pointing out these inconsistencies, which originated mostly from the difference between validated and non-validated hits. We have now corrected the issue:

Most (146/159) of the validated Golgi morphology hits resulted in a significant change of intensity for at least one lectin (Figure 7F).

-p14, 6th line from bottom, "(Figure 7)"->exactly where in Figure 7?
The intended reference was Fig S5, this is now corrected.

5. Some typos exist

-p7, 8th line from bottom, the medial Golgi is missed in claiming of stained channels/compartments
The medial Golgi is detected thanks to the GFP of the MannII-GFP, therefore does not require staining

-p24, 5th line from bottom, "GalNacT"->it was "GalNAc-T" in p6

Fixed

-p25, 3rd line from bottom, "controlling directly Golgi physiology"->rather strange ordering of words

My "Frenchness" showing up, fixed.

-the sentence crossing p9 and p10, "we discarded from further analysis genes with..."-> rather strange ordering of words

Ditto

-Supplementary for SVM design, why choose a cost parameter (28) which "minimizes" the accuracy average of the 3 PF?

Corrected "minimizes" into "maximizes", thank you for pointing this out.

Reviewer #1

Chia et al., have established an elegant assay to quantitatively monitor Golgi morphology in intact cells by microscopy. This assay is suitable for high throughput analyses and is applied to a low to medium sized siRNA screen targeting kinases and phosphatases. This screen identified 159 targets of which siRNA mediated knock-down caused changes in Golgi morphology. Secondary screens monitoring protein secretion or PM lectin quantification are used to experimentally better characterize the identified hits. Altogether, the experimental screening work appears very reproducible and is of high quality at all fronts, including data analysis, mining and presentation. These data represent some useful information for scientists working in this field and possibly beyond.

The big question that remains largely unaddressed by the present manuscript is what we have learned from this work that we did not know before and how significantly does this knowledge represent a real significant advance in the field.

Large parts of the manuscript describe the confirmation of literature data by the screening data of this manuscript. This is nice and demonstrates the high quality of the screening work, but it does not really extend our knowledge.

It remains elusive to which extent the identified hits cause the Golgi phenotypes directly or indirectly. Knock-down of 110 hits causes also alterations in protein secretion and thus raise the question if the Golgi phenotype is directly or indirectly caused (e.g. by ER to Golgi transport inhibitions, which is well known to have drastic effects on Golgi structure).

The manuscript is loaded with speculations based on the screening data and follow-up bioinformatic analyses. While this represents a good start to formulate hypotheses to understand the phenotypes caused by the gene knock-downs experimental testing is entirely missing, though essential to be conclusive. Demonstrating one or two of these hypothesis experimentally in more depth (e.g. providing experimentally direct evidence that the molecules identified as hits form indeed a network of genetic, physical interactions or phosphorylations) could strengthen the manuscript significantly.

We wished to provide an overview, with a broad yet in depth analysis, providing a resource for the community. Several hits have been confirmed more in depth, including the influence of cell surface receptors (Figure 6), where stimulation with a cognate ligand has an opposite effect to the depletion

of the receptor. We now have also confirmed some regulatory nodes as described in the answer to reviewer 2.

Specific points:

1.

The title of the manuscript is misleading. This reviewer does not see the large signalling network controlling the Golgi apparatus. Some signalling modules are presented, however none of them is experimentally addressed. If the authors mean that they have identified a large number of signalling molecules that could form a network they should specify this. Independent screens targeting kinases and phosphatases have identified similarly high numbers of signalling molecules being involved in e.g. endocytosis (Pelkmans et al.,) and ER exit (Farhan et al.,).

The Golgi apparatus is a complex but physically unified intracellular structure (sometimes referred as the Golgi network). In that sense, any signaling cascade that impacts Golgi organization is connected through the Golgi node to all other Golgi regulating cascades. Another way to phrase it is that all the signaling modules must by definition work in concert to maintain the relatively constant Golgi organization observed. In that sense, it is obvious that signaling modules regulating Golgi organization must be coordinated somehow and therefore form a large regulatory network.

2.

It is not clear if the experiments measuring MET-Luc secretion have been replicated. This should be clearly stated to provide the reader with an idea how strong these data are.

Reproducibility data was provided in Figure S5D in the original manuscript. A mention of reproducibility has now been added in page 19.

“This relatively surprising result was highly reproducible as the rest of the Met-Luc secretion data (Figure S8D).”

2nd Editorial Decision

04 October 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 the Reviewer #2 is now supportive of publication, but the Reviewer #3 has some remaining concerns and makes suggestions for modifications, which we would ask you to carefully address in a final revision of the present work.

The editor feels the concerns raised by Reviewer #3 can likely be addressed with some additional clarification and potentially some additional computational validation. In regard to the concerns related to Figure S2G, the editor agrees that the results seem to indicate a non-trivial false negative rate at the single-cell level, which may deserve more rigorous quantification with cross validation. Regardless, though, the editor agrees that the well-level reliability has been reasonably supported is ultimately more important than the individual cell-level performance. Regardless, some additional description of the reference set construction does seem needed.

In addition, we ask you to address the following format and content issues when preparing your revised work:

1. Thank you for informing us of the approximate size of the single-cell high content dataset. At 5.5 Gb even the numeric results would unfortunately be too large for us to archive as supplementary data. Nonetheless, we do feel that it would be important to offer these data to the scientific community in the event that others wish to test other classification schemes, or reproduce the existing classifications. One solution would be to submit this dataset to a public repository like DataDryad, which should be able to handle at least the numeric single cell features. This would also help assure the long term availability of these data, and possibly encourage reuse of this dataset.

2. The Supplementary Information PDF should begin with a Table of Contents listing page numbers

for all of the Supplementary content, including the separate Dataset (list as "separate file"). The Supp. Figure legends should be immediately below or after the relevant Supp. Figures.

3. Please provide an additional machine-readable version of Supplementary Table 1, ideally in excel or a text-based table format. I agree that the images contained in the current pdf version are potentially valuable, so I would encourage to provide two versions -- one pdf, one text-based table. This again may help encourage reuse and future integration of these results.

Thank you for submitting this paper to Molecular Systems Biology.

Yours sincerely,

Editor - Molecular Systems Biology
msb@embo.org

Referee reports:

Reviewer #2 (Remarks to the Author):

The authors present an improved version of the original manuscript and have addressed all my points. The inclusion of the data with IPA3 that rescues the ROCK1 knockdown and the data on the DUSPs and the rescue with the ERK inhibitor certainly improved the experimental section of the manuscript. This is important as some of the original conclusions of the manuscript are now supported by experimental data and less theoretical in nature.

Reviewer #3 (Remarks to the Author):

We appreciate the authors' efforts during the revision, but some concern still we still have some concerns exist and that prevent our recommendation of this manuscript from being published in Molecular Systems Biology in its current form:

1- Regarding the reply to comment 1 (from reviewer 3, same below), the justification for the use of SVMs relies heavily on the well-level benchmarking shown in Fig S2C, but this can be fragile: -Assume we choose 2 wells from each one of six clusters defined in Fig 3C and pool the well-level data together, can human experts accurately restore the cluster definition (Fig S2C claim more than 90% agreement between human and SVM in well level, then experts should be able to assign two wells randomly picked from a same cluster accurately back together)? And it may be especially difficult for the three clusters in the middle.

2- Regarding the reply to comment 2, the definition of "objects" still needs to be clarified.

3- Regarding the reply to the "Figure S2G" question in comment 4, the reply focuses on Figure S2B (random sample of cells), while the question was on Figure S2G (reference set). And we still believe that 1) the accuracy shown in Fig S2G is not enough for a reference set; 2) it's necessary to show the SVM performance during cross validation when comparing each reference sets with control cells; 3) the generation of reference set should be described.

Thank you for your e-mail, we much appreciate your balanced editorial work. I would like to thank also the reviewers for their valuable time and effort and for their contribution in improving our manuscript.

Regarding the last concerns of Reviewer #3, I believe they are mostly due to some remaining miscommunication as I explain below. We have modified the main text to fix these last issues as explain below.

The editor feels the concerns raised by Reviewer #3 can likely be addressed with some additional clarification and potentially some additional computational validation. In regard to the concerns related to Figure S2G, the editor agrees that the results seem to indicate a non-trivial false negative rate at the single-cell level, which may deserve more rigorous quantification with cross validation. Regardless, though, the editor agrees that the well-level reliability has been reasonably supported is ultimately more important than the individual cell-level performance. Regardless, some additional description of the reference set construction does seem needed.

Reviewer #3 (Remarks to the Author):

We appreciate the authors' efforts during the revision, but some concern still we still have some concerns exist and that prevent our recommendation of this manuscript from being published in Molecular Systems Biology in its current form:

1- Regarding the reply to comment 1 (from reviewer 3, same below), the justification for the use of SVMs relies heavily on the well-level benchmarking shown in Fig S2C, but this can be fragile: -Assume we choose 2 wells from each one of six clusters defined in Fig 3C and pool the well-level data together, can human experts accurately restore the cluster definition (Fig S2C claim more than 90% agreement between human and SVM in well level, then experts should be able to assign two wells randomly picked from a same cluster accurately back together)? And it may be especially difficult for the three clusters in the middle.

Fragile? This is a bit of a gratuitous suggestion. To generate Figure 2C, we used 120 different wells where experimenters independently classified the phenotypes and then compared it to SVM classification. The agreement was excellent (~95%). To make this point more clear, we have now replaced this panel to show more details and modified the main text slightly.

By the way, we have also visually verified all the phenotypes in the kinome screen detected by the SVM (a careful and painstaking process).

For the test described above, I think the reviewer confuses an agreement over a general phenotype versus a precise quantitative measure. Figure S2C demonstrates the overall agreement between human experimenters and the SVM over Golgi phenotypes displayed by cells in a given channel. Fig 3C refers to clustering of gene depletion results using the quantified results of the SVM over three different channels. An experimenter would be hard-pressed to quantify precisely the specific phenotypes for all the cells imaged in all three channels; that is in part why we need the SVM in the first place, it is more precise and quantitative that we can be. Thus, the test proposed, trying to restore the clustering results through visual classification is not only redundant and unwarranted but would also not be very informative because “visual clustering” would be quite difficult to perform and imprecise.

2- Regarding the reply to comment 2, the definition of "objects" still needs to be clarified.

We have added the following sentence in the Figure S2 legend. “BigobjectsSeg” and “SmallObjectsSeg” refer respectively to objects of roughly the size of a normal Golgi or large Golgi fragments and vesicles or small Golgi fragments”

3-Regarding the reply to the "Figure S2G" question in comment 4, the reply focuses on Figure S2B (random sample of cells), while the question was on Figure S2G (reference set). And we still believe that 1) the accuracy shown in Fig S2G is not enough for a reference set; 2) it's it is necessary to show the SVM performance during cross validation when comparing each reference sets with

control cells; 3) the generation of reference set should be described.

I stand by my previous answer. The result in Figure S2G is due to the cell-to-cell variability observed in most experimental conditions, this is also related the difference in SVM performance between cell level and well level results.

What maybe we failed to explain clearly enough is that for the reference set, we selected cell populations from specific wells (listed in Fig S2A) but each cell was not individually selected. As explained in my previous reply, many gene depletions used in the reference sets resulted in variably penetrant phenotypes. Some cells showed a clearly fragmented Golgi while others did not. This explains most of the alleged “misclassification” seen in Figure S2G. In fact, this is in our view not a misclassification but roughly the expected result.

Figure S2C, on the other hand, shows the performance of the SVM at the well level. As mentioned above, we have expanded it to show in greater detail the performance of the SVM for all three phenotypes.

The generation of the reference set was already described in Supplementary Material. We have added a sentence in the figure legend of S2 to explain better the result observed in Figure 2G. “The cells from the reference set that are not classified as fragmented reflect the phenotypic heterogeneity of cell populations.”

In addition, we ask you to address the following format and content issues when preparing your revised work:

1. Thank you for informing us of the approximate size of the single-cell high content dataset. At 5.5 Gb even the numeric results would unfortunately be too large for us to archive as supplementary data. Nonetheless, we do feel that it would be important to offer these data to the scientific community in the event that others wish to test other classification schemes, or reproduce the existing classifications. One solution would be to submit this dataset to a public repository like DataDryad, which should be able to handle at least the numeric single cell features. This would also help assure the long term availability of these data, and possibly encourage reuse of this dataset.

We are currently uploading the single-cell high content dataset onto DataDryad and will be providing a link shortly.

2. The Supplementary Information PDF should begin with a Table of Contents listing page numbers for all of the Supplementary content, including the separate Dataset (list as "separate file"). The Supp. Figure legends should be immediately below or after the relevant Supp. Figures.

Supp material file modified accordingly.

3. Please provide an additional machine-readable version of Supplementary Table 1, ideally in excel or a text-based table format. I agree that the images in contained in the current pdf version are potentially valuable, so I would encourage to provide two versions -- one pdf, one text-based table. This again may help encourage reuse and future integration of these results.

The Table 1 in the two formats, excel and pdf, will be uploaded. We kept the images in the Excel file as we did not think they would impair machine-readability.