molecular systems biology

Global analysis of genome, transcriptome and proteome reveals the response to aneuploidy in human cells

Silvia Stingele, Gabriele Stoehr, Karolina Peplowska, Jürgen Cox, Matthias Mann, Zuzana Storchova

Corresponding author: Zuzana Storchova, Max Planck Institute of Biochemistry

Review timeline:

Submission date:	03 April 2012
Editorial Decision:	18 May 2012
Revision received:	03 July 2012
Editorial Decision:	26 July 2012
Revision received:	31 July 2012
Accepted:	01 August 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 18 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 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 some important concerns on your work, which, I am afraid to say, preclude its publication in its present form.

The most fundamental concerns are raised by Reviewer #1 who lists two important factors that could potentially confound your results to some degree: 1) variation in statistical power to detect differential expression, either between different gene classes or due to differences in RNA vs protein quantification; and 2) off-target genomic aberrations in the cell lines. These issues need to be considered in detail, and additional statistical analyses and/or supporting control experiments appear needed to convincingly allay concerns that these issues do not undermine the main conclusions of this work.

In addition, when submitting your revised work, please provide passcodes for the proteomic datasets, and include the additional items listed in the revision checklist (below).

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):

Stingele et al report gene expression and proteomic analysis of human cell lines carrying extra chromosomes. They create isogenic strains, strengthening their ability to link observed changes to the extra chromosomes. They find, as others have shown in other systems, that aneuploid cells have growth deficiencies, that gene expression largely scales with gene copy number, and that protein levels also generally scale but are buffered to some extent. The proteins most subject to buffering are those in complexes. Also, particular pathways are perturbed in these strains.

The experiments are largely well designed and well done, and it is an advance to have this sort of data in human cells. However, I have a few concerns, detailed below:

The introduction glosses over the diversity of findings in the literature about whether all systems show gene expression changes proportional to copy number. Some studies in plants and flies, and even some naturally aneuploid yeast strains (e.g. Kvitek 2008), for example, have found this to not be the case.

Given the large growth delays and other phenotypes, it is difficult to determine which gene expression and protein changes may result from the phenotype as a secondary effect rather than being a direct result of the karyotype.

The proteomics data look to be of high quality, but proteomics still suffers from a lack of sensitivity versus nucleic acid detection. Could the fraction of proteins detectable be different in kind from the undetected proteins (especially given the other aneuploidies mentioned below)? Also, the power to detect a change in levels, especially when those are expected to be <2X, should be stated. A related issue, it wasn't clear what background set was being used for the enrichment analyses.

In Figure 2, the CGH shows an apparent decrease in copy number on chromosome 17, which seems to possibly also be detectable in the RNA data. Is this real? If so, to what degree are the other cell lines also potentially aneuploid at other loci? Looking in the supplementary Fig. 1, several of the other lines seems to also show apparent "off target" changes in copy number, and, even worse, the starting cell line also appears to have some segmental changes that appear in the mFISH karyogram as well. This complicates the ability to detect the totality of any aneuploidy signature. If the starting cells already express such a pattern, the experiments are really only detecting whatever worsening may be caused by adding more chromosomes. Although I don't think this added complication upsets the major conclusions, the manuscript should be clearer in describing the lines used. These controls are described as completely diploid in the manuscript.

I was unable to actually download the data from the TRANCHE database since it asked for a passphrase.

The paper requires minor copy editing.

Reviewer #3 (Remarks to the Author):

The manuscript by Stingele, Stoehr et al. describes an analysis of proteomic and transcriptomic changes in response to adding one or two additional chromosomes to two different human cell lines via micronuclear transfer. This is the first such study in human cells, and appears to be a very useful model to complement previous experiments done in yeast. In general, mRNA levels usually corresponded closely to gene copy number. However, the level of many proteins coded by genes on aneuploid chromosomes, notably protein kinases and members of protein complexes (which usually

also include proteins coded on euploid chromosomes), were down-regulated toward levels that would be expected from diploid chromosomes. In addition, aneuploid cell lines consistently showed down-regulation of proteins involved in DNA and RNA metabolism, and up-regulation of energy and membrane metabolism pathways, lysosomal pathways and p62-selective autophagy regardless of which chromosomes coded the pathway proteins (i.e. euploid proteins were also regulated). Detailed mechanisms for these changes in response to aneuploidy were not proposed, however the manuscript presents a fairly comprehensive global proteomic analysis of aneuploidy in human cells and should be of high general interest for the readers of Molecular Systems Biology. The findings might also prove useful for helping to understand some important human diseases such as Down's syndrome and cancer in which aneuploidy plays an important role.

Specific comments:

1. Starting with the abstract on page 2, I suggest the word "deregulated" be changed to something like "altered" or "changed" since deregulation implies lack of regulation, but in fact there is consistent but altered regulation in response to aneuploidy. For example, Supp. Fig. 4A legend includes "Co-regulation of significantly deregulated pathways" which is a contradiction.

2. page 5 line 10 is confusing. Perhaps should be "generated a cell line trisomic for chromosomes 5 and 12, and another cell line trisomic for chromosome 21..."

3. Figure 2 and supplementary methods: Were data from all biological and technical replicates including reversed label experiment included in Fig. 2? In fact, except for the switched label experiment all diploid cells were grown in heavy medium and aneuploid cells grown in light (normal) medium. While unlikely, it is possible that the medium, in addition to or instead of chromosome copy number, had a systematic effect on cell growth and/or some expression levels. While ideally the labels would have been switched for half of the biological replicates, or perhaps between the experiments with the 2 different cell types, the reverse label experiment with HCT-116 5/4 cells can be very useful for allaying these concerns. While I realize all data are available in the tables, it would be very helpful to show in a Supplemental Figure a scatter plot comparison between the single forward and reverse label experiment that was done to demonstrate this point.

4. Supp Fig. 2B legend is confusing. "Categories" can mean many things (functional categories, etc.). I believe in this case it means "genes or proteins", if so, those words should be used in place of "categories". Also, which 2 populations are significantly different in all cases: comparisons of genes or proteins on aneuploid chromosomes vs. euploid chromosomes or ratio of genes on aneuploid chromosomes vs. ratio of proteins coded by aneuploid chromosomes? The latter would be more interesting and more in line with the conclusions of the paper, but in either case this should be explicitly stated.

5. Supp. Fig. 3B typo should be "one subunit"

6. Page 8 line 2 typo should be "complexes"

7. Supp. Fig. 6B bottom left labels should be "12/3" not "5/12" in both figures.

8. Page 13 line 19 should be "human"

9. References should be in alphabetical order

1st Revision - authors' response 03 July 2012

Author Name: Stingele, Silvia; Stoehr, Gabriele

Detailed response to reviewer comments

The reviewer comments are shown in their entirety below (black font) with our responses in red. Additional data not shown in the manuscript are presented as Figure R1. Citations from the modified main text are written in bold.

Reviewer #2 (Remarks to the Author):
Stingele et al report gene expression and proteomic analysis of human cell lines carrying extra chromosomes. They create isogenic strains, strengthening their ability to link observed changes to the extra chromosomes. They find, as others have shown in other systems, that aneuploid cells have growth deficiencies, that gene expression largely scales with gene copy number, and that protein levels also generally scale but are buffered to some extent. The proteins most subject to buffering are those in complexes. Also, particular pathways are perturbed in these strains.
The experiments are largely well designed and well done, and it is an advance to have this sort of data in human cells. However, I have a few concerns, detailed below:

The introduction glosses over the diversity of findings in the literature about whether all systems show gene expression changes proportional to copy number. Some studies in plants and flies, and even some naturally aneuploid yeast strains (e.g. Kvitek 2008), for example, have found this to not be the case.
>We are glad that the reviewer found the experiments well designed and well done. We thank the reviewer for suggesting additional manuscripts on gene expression changes in response to aneuploidy in other organisms. The relationship between gene copy number and transcription levels is complex and it might differ in different species. In order to better reflect this ongoing discussion, we have modified the specific paragraph (Introduction, page 4, row 2 - 8) and cited additional publications.

Given the large growth delays and other phenotypes, it is difficult to determine which gene expression and protein changes may result from the phenotype as a secondary effect rather than being a direct result of the karyotype.
>Indeed, it is possible that some of the gene expression changes are a secondary effect of the growth delay. Some of the pathways that are down-regulated in aneuploid cells (e.g. DNA replication, chromatin remodeling or cell cycle) might be altered due to slower growth. However, this is likely not the main reason as, for example, the HCT116 5/4 and HCT116 3/3 grow similarly slowly (Fig. 1B), but e.g. the category “cell cycle” is significantly down-regulated only in HCT116 3/3 (Supplementary Fig. 5A). Furthermore, we have recently obtained aneuploid cell lines derived from HCT116 H2B-GFP and RPE-1 H2B-GFP that grow nearly as good as the parental cell lines, yet they show the “aneuploid” pattern of pathway regulation changes. We prefer to leave out these new cell lines from the manuscript, because the aneuploids were generated in a different way and analyzed only by microarrays, but we include some of the data for the reviewers (Figure R1). We believe that even secondary changes are interesting and possibly relevant to the studies of consequences of aneuploidy. To acknowledge the possibility that the effects might be secondary, we have discussed this issue in the Discussion (Page 13, bottom):
It should be noted that some of the observed changes in pathway regulation might be an indirect consequence of the altered phenotypes of aneuploid cells, such as the growth defect. We propose that the effect is rather modest, because the changes in pathway regulation are not proportional to the growth defects.

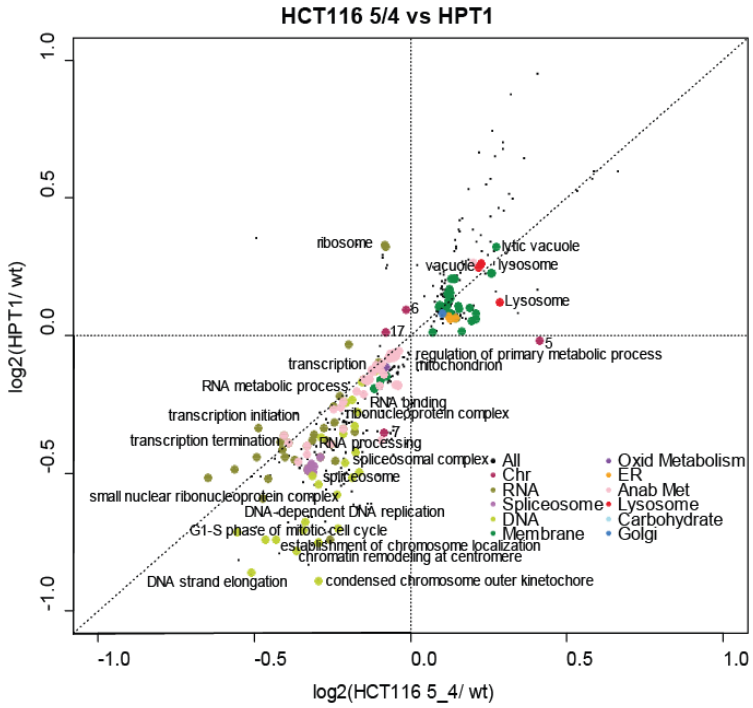

Figure R1: 2-D comparison of the transcriptome changes in HCT116 5/4 (growth delay in comparison to HCT116 approximately 10 hrs) and HPT1 (an aneuploid cell line derived from HCT116 with several chromosomal changes and a growth delay ~1.5 hrs in comparison to HCT116)

The proteomics data look to be of high quality, but proteomics still suffers from a lack of sensitivity versus nucleic acid detection. Could the fraction of proteins detectable be different in kind from the undetected proteins (especially given the other aneuploidies mentioned below)? Also, the power to detect a change in levels, especially when those are expected to be <2X, should be stated.

>Although it is true that current proteomics does not reach the sensitivity of nucleic acid detection, we reason that in our experimental setting this does not have a major effect on the results. First, as we always measure the aneuploid-to-diploid ratio, any possible bias for detecting different protein classes in aneuploids should be eliminated, because the ratio for these proteins could not be calculated. The segmental aneuploidies should not affect the result either as they are present in the diploid controls as well (see below). Secondly, we have identified similar pathway alterations when comparing the complete mRNA datasets in the 2-D pathway analysis or when comparing the matched mRNA and proteome datasets. These results have now been added to the manuscript in Supplementary figure 5B and are mentioned in the main text (Page 10, line 2-4):

Remarkably, the 2-D enrichment analysis of the transcriptome data determined similar changes in the pathway regulation, suggesting that the pathway response cannot be caused by a bias in protein detection (Supplementary Figure S5B, S5C).

The power to detect a change in level depends on number of measurements, their standard deviation, false positive rate α and the difference we wish to detect (the effect size). We have performed three biological replicates for HCT116 5/4 (six measurements in total, including label swap). The RPE-1 derived samples were measured in biological triplicates. The SD in these experiments ranges from 0.06 to 0.15. We calculated that the power to detect a change 1.6-fold exceeds 0.8 in all these measurements, which is considered sufficient. However, we prefer not to include these calculations for following reasons. First, they were performed *a posteriori*, which is often misleading. These calculations might be rather used to estimate *a priori* the sample size (= number of measurements) that should be used in order to determine the changes with high confidence. We followed the standard of the field where three measurements are considered reasonably sufficient and typically used; four measurements provide an increase in both accuracy and precision (e.g. Pavelka et al., Mol. Cell. Proteomics 2008). Finally, and most importantly, we determined the shared global trends within protein classes rather than

identifying changes in abundance of individual proteins, and significance herein was additionally tested with Wilcoxon rank sum test.
A related issue, it wasn't clear what background set was being used for the enrichment analyses.
>We have performed the 2D annotation enrichment analysis for all analyzed cell lines. This software allows comparison of two different dataset by extracting significantly altered pathways within the individual dataset. Thus, each subcategory (= a specific pathway) from one dataset was compared to the complete dataset of proteins quantified in the same experiment. A comparative t-test was performed between the two different populations, i.e. subcategory versus the complete dataset. This is now additionally specified in the main text as well as in the supplementary data (page 9, line 12 - 14):
For each clone, we identified all significantly altered pathways (as defined by Gene Ontology categories, KEGG pathways and CORUM database) and ranked the relative abundance changes of proteins within the category compared to the complete measured dataset of the respective cell line.

In Figure 2, the CGH shows an apparent decrease in copy number on chromosome 17, which seems to possibly also be detectable in the RNA data. Is this real? If so, to what degree are the other cell lines also potentially aneuploid at other loci? Looking in the supplementary Fig. 1, several of the other lines seems to also show apparent "off target" changes in copy number, and, even worse, the starting cell line also appears to have some segmental changes that appear in the mFISH karyogram as well. This complicates the ability to detect the totality of any aneuploidy signature. If the starting cells already express such a pattern, the experiments are really only detecting whatever worsening may be caused by adding more chromosomes. Although I don't think this added complication upsets the major conclusions, the manuscript should be clearer in describing the lines used. These controls are described as completely diploid in the manuscript.
>The cell line HCT116 contains some numerical (lacking chromosome Y) and segmental aneuploidies (e.g. on chromosome 17 and chromosome 10), as can be seen in the Supplementary Figure S1B. These alterations are carried over to the HCT116-derived tri- and tetrasomic cell lines and “neutralized” by the analysis in which we calculate aneuploid-to-diploid ratio. The only apparent exception is the segmental aneuploidy of chromosome 17 in HCT116 5/4. In fact, the copy number of a small part of chromosome 17 is increased in HCT116 (as well as in HCT116 3/3, HCT116 H2B-GFP 5/4 and 5/3, Supplementary Figure S1B). This amplification is partially lost in HCT116 5/4, which can be observed as a “decrease” in the plot of aneuploid-to-diploid ratios (Figure 2A). We do not expect this change to affect the results since we identified the same pathway response in all analyzed cell lines derived from HCT116, both with and without the copy number change of part of chromosome 17. Additionally, we used tri- and tetrasomic cell lines derived from the immortalized primary cell line RPE-1 that is a near perfect diploid, with only one amplification of a small part of chromosome 10 (also evident in the aneuploid clones, Sup. Fig. 1B). Thus, a *de novo* response to aneuploidy should be detected in these cell lines. Results obtained with the RPE-1 derived cell lines are very similar to the results from HCT116 derived aneuploids. It should be noted that most of the available cell lines contain at least small segmental chromosomal aberrations. We now clarify the description of the used cell lines (page 5, first paragraph of the Result section):
HCT116 is a transformed cell line with several previously identified chromosomal changes such as the chromosome Y loss and amplified regions of chromosomes 8, 10 and 17 (Masramon et al., 2000). These aberrancies are mostly present in the new aneuploid cell lines (Supplementary Figure 1B) and thus likely do not affect the results. Nevertheless, to strengthen our analysis and to overcome this possible drawback, we generated cell lines trisomic for chromosomes 5 and 12, and another cell line trisomic for chromosome 21, both derived from the diploid primary epithelial cell line RPE-1 that was immortalised by expression of hTert and that lacks substantial chromosomal aberrancies.

I was unable to actually download the data from the TRANCHE database since it asked for a passphrase.
>We apologize for omitting the passphrases. They were now added to the table with the hashcodes for the TRANCHE database (Supplementary Information, page 19). The passphrase for the proteomics data of the HCT116 derived samples is "StorchovAneu10"; for all other datasets "Aneuploidy".

The paper requires minor copy editing.
>We carefully edited the manuscript.

Reviewer #3 (Remarks to the Author):
The manuscript by Stingele, Stoehr et al. describes an analysis of proteomic and transcriptomic changes in response to adding one or two additional chromosomes to two different human cell lines via micronuclear transfer. This is the first such study in human cells, and appears to be a very useful model to complement previous experiments done in yeast. In general, mRNA levels usually corresponded closely to gene copy number. However, the level of many proteins coded by genes on aneuploid chromosomes, notably protein kinases and members of protein complexes (which usually also include proteins coded on euploid chromosomes), were down-regulated toward levels that would be expected from diploid chromosomes. In addition, aneuploid cell lines consistently showed down-regulation of proteins involved in DNA and RNA metabolism, and up-regulation of energy and membrane metabolism pathways, lysosomal pathways and p62-selective autophagy regardless of which chromosomes coded the pathway proteins (i.e. euploid proteins were also regulated). Detailed mechanisms for these changes in response to aneuploidy were not proposed, however the manuscript presents a fairly comprehensive global proteomic analysis of aneuploidy in human cells and should be of high general interest for the readers of Molecular Systems Biology. The findings might also prove useful for helping to understand some important human diseases such as Down's syndrome and cancer in which aneuploidy plays an important role.
>We are glad that the reviewer finds the manuscript of high general interest and similar to the reviewer we hope that these data will be useful for understanding human pathologies associated with aneuploidy.

Specific comments:
1. Starting with the abstract on page 2, I suggest the word "deregulated" be changed to something like "altered" or "changed" since deregulation implies lack of regulation, but in fact there is consistent but altered regulation in response to aneuploidy. For example, Supp. Fig. 4A legend includes "Co-regulation of significantly deregulated pathways" which is a contradiction.
>We have used the word "deregulated" as a description for a regulation different from the standard regulation including both up- and down-regulation. However, we understand that this might be confusing and we describe it now as "alterations" and "altered regulations" through the text, as for example in the abstract (on page 2):
This analysis revealed specific and uniform alterations in pathway regulation in cells with extra chromosomes.

2. page 5 line 10 is confusing. Perhaps should be "generated a cell line trisomic for chromosomes 5 and 12, and another cell line trisomic for chromosome 21..."
>We have changed the sentence according to the reviewers suggestions and it now reads:
To expand our analysis, we generated cell lines trisomic for chromosomes 5 and 12, and another cell line trisomic for chromosome 21, both derived from the diploid primary epithelial cell line RPE-1 that was immortalized by expression of hTert and that lacks substantial chromosomal aberrancies."

3. Figure 2 and supplementary methods: Were data from all biological and technical replicates including reversed label experiment included in Fig. 2? In fact, except for the switched label

experiment all diploid cells were grown in heavy medium and aneuploid cells grown in light (normal) medium. While unlikely, it is possible that the medium, in addition to or instead of chromosome copy number, had a systematic effect on cell growth and/or some expression levels. While ideally the labels would have been switched for half of the biological replicates, or perhaps between the experiments with the 2 different cell types, the reverse label experiment with HCT-116 5/4 cells can be very useful for allaying these concerns. While I realize all data are available in the tables, it would be very helpful to show in a Supplemental Figure a scatter plot comparison between the single forward and reverse label experiment that was done to demonstrate this point.
>We agree with the reviewer that this is an important point. We have performed the reverse label experiment with the biological replicates of HCT116 5/4 and did not see any significant difference in the quantified levels; thus, we did not repeat the label swap with the other cell lines. To illustrate the correlation between the individual measurements, we included the scatter plot comparisons. The Pearson correlation of individual experiments was between 0.6 and 0.9. The plots that show comparison of technical replicates, biological replicates and label-swapped replicates are now part of Supplementary figure 2A.

4. Supp Fig. 2B legend is confusing. "Categories" can mean many things (functional categories, etc.). I believe in this case it means "genes or proteins", if so, those words should be used in place of "categories". Also, which 2 populations are significantly different in all cases: comparisons of genes or proteins on aneuploid chromosomes vs. euploid chromosomes or ratio of genes on aneuploid chromosomes vs. ratio of proteins coded by aneuploid chromosomes? The latter would be more interesting and more in line with the conclusions of the paper, but in either case this should be explicitly stated.
>Indeed, the term “Categories” was used here for genes or proteins and we have now changed the figure legend appropriately (please note that these plots are now in a new Supplementary Fig 3) :
. . . The numbers in parenthesis represent the total number of quantified genes and proteins, respectively. . .
Regarding the comparisons, we had compared the quantitative changes of either proteins or genes on the euploid chromosomes vs. the quantitative changes on the aneusomes (red line). This was stated in the headline of each plot (see “DNA density” or “Protein density”) and in the plot insets (“All w/o chr. 5” and “Chr. 5”). However, we agree with the reviewer that the more interesting comparison is the ratio of genes (mRNAs) on aneuploid chromosomes vs. ratio of proteins coded by aneuploid chromosomes (as it is in the Fig. 2B) and we exchanged these plots accordingly. The new plots with corrected figure legend are now in Supplementary Figure 3.

5. Supp. Fig. 3B typo should be "one subunit"
6. Page 8 line 2 typo should be "complexes"
7. Supp. Fig. 6B bottom left labels should be "12/3" not "5/12" in both figures.
8. Page 13 line 19 should be "human".
9. References should be in alphabetical order
>We appreciate the careful reading. We have corrected all these typos.

2nd Editorial Decision 26 July 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 this revised study. As you will see, the referees the felt that the revisions made to this work had largely satisfied their previous concerns. The last reviewer, however, has some final suggestions for modifications (see below), which we would ask you to address in a final revision of the present work.

In addition, please address the following format and content issues when preparing your revised work:

1. Thank you for providing the underlying large-scale datasets via the Tranche repository. In general, we ask that microarray-based datasets are deposited at a community repository specializing in these datatypes, such as GEO or ArrayExpress (expression and aCGH data). Please contact if this will pose a problem. The proteomic data, of course, is fine to provide via Tranche.

2. In addition to our capacity to host datasets in our supplementary information section, we provide a new functionality that allows readers to directly download the 'source data' associated with selected figure panels (e.g. <http://tinyurl.com/365zpej>). This sort of figure-associated data may be particularly appropriate for some of the figures presented in this work, especially in cases were key data are not covered by the large-scale datasets (e.g. Fig. 1BC). Please see our Instructions of Authors for more details on preparation and formatting of figure source data (<http://www.nature.com/msb/authors/index.html#a3.4.3>).

3. Please add a few sentences concisely describing the contents of the Supp. Tables to the Supplementary Information PDF, to assist readers in interpreting this files.

Please resubmit your revised manuscript online, with a covering letter listing amendments and responses to each point raised by the referees. Please resubmit the paper **within one month** and ideally as soon as possible. If we do not receive the revised manuscript within this time period, the file might be closed and any subsequent resubmission would be treated as a new manuscript. Please use the Manuscript Number (above) in all correspondence.
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):

I am satisfied with the authors' responses to the reviews.

Reviewer #3 (Remarks to the Author):

I believe that for the most part the authors have adequately addressed the reviewers' concerns. I would request however that the authors address 2 minor issues before publication:

1. As written on page 10, line 10 the figure with data excluding chromosome 5 is no longer Fig S5C but is now Fig. S5D.

2. The plot in Supplementary Fig. 2A showing correlation between biological replicates with

swapped labels is very useful and shows that there is not a major, global effect of label isotopes in the experiments. However, there may be a minor effect as indicated by lower r value for that experiment (0.62) compared to the other biological replicate experiment without swapped labels (0.79). Are the proteins with outlying ratios in the upper left and lower right quadrants whose values may depend more on the label than on the cell's gene copy number functionally related, i.e. could they explain one of the protein functional groups with ratio changes thought to be explained by aneuploidy, or are they non-related, random proteins? A sentence describing the general characteristics of these proteins, or lack of general characteristics, would be helpful.

2nd Revision - authors' response 31 July 2012

We have now addressed the final comments regarding the manuscript. As requested, we have created an account in the GEO database and will upload the data from transcription analysis following the MIAME guidelines shortly. We have also improved the description of the Supplementary tables so that it clearly explains the content of the respective tables. We were pleased that both reviewers are satisfied with our responses. We addressed the final remarks from Reviewer #3 on the next page. We believe that the final comments were addressed sufficiently and adequately.

We hope that you will find the revised manuscript improved and suitable for publishing in Molecular Systems Biology.

Detailed response to reviewer comments

Reviewer #3 (Remarks to the Author):

I believe that for the most part the authors have adequately addressed the reviewers' concerns. I would request however that the authors address 2 minor issues before publication:

1. As written on page 10, line 10 the figure with data excluding chromosome 5 is no longer Fig S5C but is now Fig. S5D.

We thank the reviewer for carefully reading the manuscript. The figure number has been changed now in the text.

2. The plot in Supplementary Fig. 2A showing correlation between biological replicates with swapped labels is very useful and shows that there is not a major, global effect of label isotopes in the experiments. However, there may be a minor effect as indicated by lower r value for that experiment (0.62) compared to the other biological replicate experiment without swapped labels (0.79). Are the proteins with outlying ratios in the upper left and lower right quadrants whose values may depend more on the label than on the cell's gene copy number functionally related, i.e. could they explain one of the protein functional groups with ratio changes thought to be explained by aneuploidy, or are they non-related, random proteins? A sentence describing the general characteristics of these proteins, or lack of general characteristics, would be helpful.

We selected the population of proteins whose abundance changes did not correlate after label swap and performed a Fisher Exact Test (Benjamini-Hochberg FDR < 2%) on these proteins. No enrichment of any pathway or protein category has been identified, thus confirming that this group

consists of non-related, random proteins. The text explaining this finding was now added to the figure legend of the respective figure:
Supplementary Figure 2A: . . . . A small population of proteins that do not follow the same trend can be identified when comparing forward and reverse labeling experiments (right panel). These proteins are not significantly enriched for any pathway (Fisher exact test, Benjamini-Hochberg FDR < 2%), suggesting that the effect of label swap is random and does not affect the pattern of pathway down- or up-regulations.

