THE EMBO JOURNAL

Manuscript EMBO-2010-75671

Novel asymmetrically localizing components of human centrosomes identified by complementary proteomics methods

Lis Jakobsen, Katja Vanselow, Marie Skogs, Yusuke Toyoda, Emma Lundberg, Ina Poser, Lasse Falkenby, Jens Westendorf, Erich Nigg, Mathias Uhlen, Tony Hyman, Jens Andersen

Corresponding author: Jens Andersen, University of Southern Denmark

Review timeline:

Submission date:	17 August 2010
Editorial Decision:	16 September 2010
Revision received:	15 December 2010
Editorial Decision:	17 January 2011
Revision received:	09 February 2011
Accepted:	11 February 2011

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 16 September 2010

Thank you for submitting your manuscript for consideration by The EMBO Journal. It has now been evaluated by three expert referees, whose comments you will find copied below. As you will see, all three reviewers acknowledge the importance and interest of your identification strategy for centrosomal protein components, and would thus in principle support publication in The EMBO Journal. On the other hands, they raise major concerns with the presentation of the study in its current form, regarding e.g. background introduction, rationalization of different used strategies, or analysis, mining and comparisons of the resulting datasets. I will not expand on this further here, as these criticisms are laid out quite clearly in the reports below; but as I am confident that you should be able to improve the manuscripts along the lines suggested in the reports, I am happy to give you an opportunity to address these points in the form of a revised version of the manuscript. I should however remind you that it is EMBO Journal policy to allow a single round of major revision only, and that it will thus be important to diligently answer to all the various major and minor points raised at this stage. When preparing your revision, please also bear in mind that your letter of response will form part of the Peer Review Process File, and will therefore be available online to the community in the case of publication (for more details on our Transparent Editorial Process initiative, please visit our website: http://www.nature.com/emboj/about/process.html). Finally, please also pay attention to editorial issues such as the EMBO Journal format guidelines (especially for referencing), and briefly specify the individual author contributions, either in the acknowledgements section or in an adjacent separate section. In any case, please do not hesitate to get back to us should you need feedback on any issue regarding your revision.

Thank you for the opportunity to consider your work for publication. I look forward to your revision.

Yours sincerely,
Editor
The EMBO Journal

_____

REFEREE REPORTS:

Referee #1 (Remarks to the Author):

The manuscript "Complementary proteomics methods identify novel components localizing asymmetrically to centrosomes" shows different methodologies to identify novel centrosome components and get some estimates of their biochemical properties. The authors did:
1. Mass spectrometry from centrosome preparations using PCP- SILAC (protein correlation profiling with stable isotope labeling by amino acid in cell culture) which :
ï improved sensitivity regarding previous work and resulted in a list of 167 proteins that co-fractionate with centrosomes (118 already known; 32 novel consistently observed; 12 candidates with data supported by data from 1/2 experiments).
ï Led to rough estimates of abundance of each candidate protein- which span at least two orders of magnitude
ï Led to rough estimates of turnover rates of centrosomal proteins.
ï The authors tested the candidates using BacTRansgenomics of GFP-tagged candidates and/or using immunocytochemistry with HPA antibodies:
-20/32 tested known centrosomal protens localized to centrosomes
-22/34 candidates localized to centrosome
2. Antibody based screen using HPA antibodies- the consortia screened 4000 proteins in 3 different cell lines. This screen used alfa tubulin as a marker for microtubules (which is not particularly good for identifying centrosome localization).
ï From what I understood 99 of those indicated a centrosome staining on at least one cell line. 21 candidates had a more obvious centrosome signal and were further tested in this manuscript with gamma tubulin colocalization (centrosome marker). 10 localised to centrosomes and 4 to structures around the centrosomes (perhaps satellites).
ï The authors picked a range of some of the proteins identified above and searched for the ones that might be differentially present in mother vs daughter centriole, including ones already known such as ODF2 and Centrobin. They found 7 of those using acetylated tubulin as a label for cilia and therefore a proxy for the mother centriole.
In general the data present in the manuscript is very useful. However, I found the manuscript very difficult to read. Two strategies were employed in this manuscript in order to find novel centrosome components. Clearly they are complementary as there is only some overlap. Why is that the case? Could the authors estimate something regarding the REAL total centrosome proteome? How many proteins are likely to localize there and which are their biochemical properties (abundance & lifetime at centrosome) and which methods are likely to produce a better result in their identification?
As a whole I felt the authors need to better explain the strategy followed- for example the fact that there were two separate screens, but that the same antibodies were used in both for screening and validation. Perhaps a diagram with all the techniques used and which ones were used for validation of others would help. Would be also useful to have an idea of how many candidates were retrieved from each one of those stages. I found it difficult to fully understand the order of events, and which candidates from which screen were validated, using which methodology. I feel the manuscript may not have an audience if it is not simplified. I feel there is a lot of validation but most of the data is scattered. Perhaps the manuscript would gain from streamlining and focusing on a particular protein complex as a validation strategy.
To make the manuscript more predictive perhaps would be interesting to cluster candidate proteins according to their abundance and stability (separately) to suggest protein complexes; and then check whether those proteins display similar localization to centrosomes at different cell cycle stages.

Figure legends should be more detailed. As a non expert in the proteomics field I felt it was difficult to understand the figures related to that technique.
Figures with cells would gain from magnified insets from centrosome region as in some of them it is difficult to see the centrosomes.

Referee #2 (Remarks to the Author):

This manuscript describes a study which uses a variety of techniques to identify novel components of the centriole. Of the novel proteins assigned to this sub cellular location, five are shown to be implicated in cell polarity. The manuscript also describes an elegant approach to investigate the range of stabilities within this set of centriole proteins.
This represents a thorough piece of work. The manuscript is mostly clearly written and well supported by experimental data. It falls down in terms of data analysis, where only choice proteins are annotated (figure 2G) and leaves the reader wondering why the method of analysis was chosen and how well the approach differentiates between organelles which have very similar densities and separation profiles on sucrose gradients. Before it is suitable for publication in this journal I suggest the following be considered which involves some re-analysis of the data:

1. The introduction launches straight into centrosome biology without really describing what a centrosome is. I think a sentence describing the function of the centrosome would not go amiss at the beginning of this section.

2. The authors chose one fraction which contains the highest abundance of centrosome proteins plus five neighbouring fractions which contained some amount of these proteins for analysis using protein correlation profiling (PCP). It is not clear if the five fractions were from either side of the peak fraction; figure 1c would suggest only one fraction was more dense than the peak fraction, and data is not shown supporting that these fractions contained centrosomal proteins or indeed which centrosomal proteins they contained. I think this is somewhat of an oversight by the authors and it would be good to see some evidence of for example, a Western blot of the fraction depicting the total distribution of key centrosomal markers down the sucrose gradient.

3. The authors talk about variable incorporation rates of SILAC amino acids in several proteins previously implicated as being part of the centrosome. Can the authors therefore be sure that the distribution patterns of proteins are not skewed by variable incorporation rates leading to false assignments? This should be clarified in the text.

4. The authors claim that the approach using a pooled internal standard coupled with SILAC labelling gives more accurate data than the traditional PCP. This point would be better supported if direct comparison of the two approaches was presented. This should require no more data collection, but analysis of distributions of proteins within the separate fractions without utilizing the internal standard.

5. The rationale behind carrying out the third experiment is not clear. The design of this experiment was not only to demonstrate the reproducibility of the distribution patterns of the centrosomal proteins in separation gradients but also to use correlated distributions between different gradients to add further evidence for similar behaviour between centrosomal candidates. This message was lost a little in the manuscript and should be stated more clearly.

6. Why do they choose 8 as a threshold? The authors could elaborate on this. Why are certain proteins in figure 2G not marked as candidates, even if the Mahalanobis distance is < 8 for the two replicates?

7. Figure 2A-E - consider using thinner lines to make the figures clearer.

8. Can the authors also map onto this figure other organelles? There are plenty of grey dots some of which seem to cluster with the proteomeal and ribosomal proteins, but no attempt at annotation of these is made.

9. On page 8 the authors talk about comparison of the Mahalonobis distances for proteins in 10

fractions in experiment three, but it is not clear what these then fractions refer to. This should be clarified.

10. At the bottom of page 10, the authors describe in-vivo localization studies of 34 tested candidates, but it's not clear what this number refers as on page 8, second paragraph, 32 candidates present in all experiments and 12 additional candidates found only in a single experiment were described (also table 1). If only a selection of candidates were taken for analysis using this method, then the reasons behind the selection should also be stated.

11. In the section describing antibody based screening again not all candidates seem to have been tested.

12. Further in this section differential staining patterns are described for the new centrosomal proteins, from large diffuse structures to distinct satellites. Within the correlation profiling did these proteins which had similar staining patterns have closer correlation with each other than with those who staining patterns was very different?

13. How specific are the antibodies utilized? Could any of these antibodies have been used in Westerns to verify the distribution of proteins down the gradient. If some of the centrosomal proteins were present in more than one sub cellular location of differing densities, the alternative localizations would be missed in this study (see point 2 above).

Referee #3 (Remarks to the Author):

This is a novel and convincing article, combining advanced mass spectrometry and high-throughput immunofluorescence microscopy, providing an important addition to the proteome of centrosomes. The methodologies used are novel, and themselves worth reporting, with the technique being firmly validated by expression of fusion proteins at endogenous levels, resulting in the identification of previously unidentified centrosomal proteins. That the analysis of these proteins is not taken further is justified by the advances made in the methodology.

Major Points:
(i) The major criticism I have is regarding the clarity of some of the text and the consistency of the referrals in the text to the Figures themselves (which is inconsistent in a number of cases and therefore makes it difficult to fully interpret some of the results):
Specifically, on p6, para2, we are told "profiles of the centrosomal proteins closely followed the CEP135 profile" and are pointed to Fig 1D. However, in the Figure legend, this appears to correspond to Fig 1E - with D and F pertaining to a different experiment. Either there is mislabeling of Fig 1, or there needs to be a clearer description in the associated text.

Secondly, the text explaining Figure 3A and B is not clear enough. The two profiles are "two separate centrosome preparations" (Legend of Fig2) - and yet the profiles are markedly different (the ratio of A reduces as the fraction number increases, while B shows a profile similar to the original experiments in Fig1). I think I must be confused - in which case, the text needs to be clarified.

Thirdly, the labels of Figure 3D-H do not seem to match up to the associated text on p9. Again, clarify or correct.

Fourthly, the text states 44 candidate centrosomal proteins were isolated using PCP-SILAC "32 likely candidates observed in all PCP-SILAC experiments and 12 additional candidates supported by data in one or two of the experiments" (p8 para2). However, 47 are mentioned in the Discussion on p13, and only 43 are present in Table 1!

These may seem like minor points, but together it suggests the article was submitted without getting the details right - which are essential in order to fully review the science behind them.

(ii) On a positive note, I applaud the Authors in their representation of the Mass Spectrometry data. It is very well laid out, and contains all the required information to justify their results. It's not often proteomics articles display this degree of ludicity!

(iii) On p10, at the end of the first paragraph, the authors suggest that their data regarding turnover rates of centrosomal proteins might be a useful parameter to estimate the time required for RNAi knockdown. I would like to see this hypothesis tested - either by looking through existing literature, or by a small-scale experiment.

Minor Points:
(i) Introduction, p3; para2; line 7 - References pertaining to the roles of centrosomes in cellular processes need to be added here.
(ii) Results p6; para1; last line - It would be appropriate to include at least one representative picture of a 1D gel in the Supplementary data.
(iii) Results p6; para2 (also Legend for Fig 2) - I could not see the triangles and asterisks supposed to distinguish the different isotope signals.
(iv) Results p8; para 3; line 11 - "several of the newly identified centrosomal proteins were of relatively low abundance". It would be helpful to clarify this statement
(v) Results p8; para 3; last line - BB54 and NPHP4 do not appear in Table 1, but in the text it suggests they are. Were they not identified as candidates? Or have they been shown to be centrosomal by other research groups subsequent to the authors' previous study? It's a bit confusing.
(vi) Results p9; para 1 - The text and the Figure 3 should be consistent - either call/label the proteins TUBG1, 2, 3 etc. or TUBGCP1,2,3 etc. Currently there are different nomenclatures in the text and in Figure 3.
(vii) There is a rather spurious reference to the retinal atrophy protein CCdC66 at the end of p9 - without discussing the significance of this example.
(viii) Results p10; p2 - We are told in the final sentence on this page that 34 candidate proteins were tested for their localization using GFP-tagging. But there were 44 candidates in total. Which proteins were not tagged, and why (I'm assuming the cloning or expression did not work - but we need that information for completeness)
(ix) Results p11; para 2 - the conjecture regarding the function of CCDC21 should be in the Discussion, not results.
(x) Table 1 - need to include what "AB" stands for in the Legend (presumably localization tested using antibodies?)

1st Revision - authors' response 15 December 2010

Point-by-point response to the referees

We would like to thank all the referees for constructive and useful comments. A common concern raised by the referee was related to the clarity of presentation of the data, the rationalization of the different used strategies, and the analysis, mining and comparisons of the resulting datasets. To address this concern, we have (i) added an overview diagram of the different strategies used, (ii) added a dynamic protein localization network summarizing most of our data in a single figure, (iii) rearranged the order of appearance of the experiments in a more logical order within the manuscript, (iv) added schematic drawings to illustrate the asymmetric centrosome localization and to simplify reading the paper, and (v) added additional supplemental figures with the further statistical analysis of the data. The point-by-point response to the individual referees are provided below.

Referee #1 (Remarks to the Author):

The manuscript "Complementary proteomics methods identify novel components localizing asymmetrically to centrosomes" shows different methodologies to identify novel centrosome components and get some estimates of their biochemical properties. The authors did:

1. Mass spectrometry from centrosome preparations using PCP- SILAC (protein correlation profiling with stable isotope labeling by amino acid in cell culture) which : (i) improved sensitivity regarding previous work and resulted in a list of 167 proteins that co-fractionate with centrosomes

(118 already known; 32 novel consistently observed; 12 candidates with data supported by data from 1/2 experiments).(ii) Led to rough estimates of abundance of each candidate protein- which span at least two orders of magnitude. (iii) Led to rough estimates of turnover rates of centrosomal proteins. (iv) The authors tested the candidates using BacTransgenomics of GFP-tagged candidates and/or using immunocytochemistry with HPA antibodies:
- 20/32 tested known centrosomal protens localized to centrosomes
- 22/34 candidates localized to centrosome

2. Antibody based screen using HPA antibodies- the consortia screened 4000 proteins in 3 different cell lines. This screen used alfa-tubulin as a marker for microtubules (which is not particularly good for identifying centrosome localization). (i) From what I understood 99 of those indicated a centrosome staining on at least one cell line. 21 candidates had a more obvious centrosome signal and were further tested in this manuscript with gamma tubulin colocalization (centrosome marker). 10 localised to centrosomes and 4 to structures around the centrosomes (perhaps satellites). The authors picked a range of some of the proteins identified above and searched for the ones that might be differentially present in mother vs daughter centriole, including ones already known such as ODF2 and Centrobin. They found 7 of those using acetylated tubulin as a label for cilia and therefore a proxy for the mother centriole.

In general the data present in the manuscript is very useful. However, I found the manuscript very difficult to read. Two strategies were employed in this manuscript in order to find novel centrosome components. Clearly they are complementary as there is only some overlap. Why is that the case? Could the authors estimate something regarding the REAL total centrosome proteome? How many proteins are likely to localize there and which are their biochemical properties (abundance & lifetime at centrosome) and which methods are likely to produce a better result in their identification?

As a whole I felt the authors need to better explain the strategy followed - for example the fact that there were two separate screens, but that the same antibodies were used in both for screening and validation. Perhaps a diagram with all the techniques used and which ones were used for validation of others would help. Would be also useful to have an idea of how many candidates were retrieved from each one of those stages. I found it difficult to fully understand the order of events, and which candidates from which screen were validated, using which methodology. I feel the manuscript may not have an audience if it is not simplified. I feel there is a lot of validation but most of the data is scattered. Perhaps the manuscript would gain from streamlining and focusing on a particular protein complex as a validation strategy.

We have followed the advice of the referee and added a diagram of the experimental strategies followed to identify and characterize centrosomal proteins (Fig. 1A). The number of candidate proteins retrieved from each of the experimental stages is included in this diagram. We have introduced the terms ‘MS-screen’ and the ‘HPA-screen’ to clarify which candidate proteins from which of the two screen were validated, using which methodology. The two screens are by design non-overlapping: antibodies used for validating candidates identified by the MS-screen were preselected based on proteins identified by mass spectrometry. Antibody-staining evaluated in the HPA-screen represented the remaining subset of proteins for which HPA-antibodies were available. Positive candidates from the HPA screen were further evaluated using the same antibodies but using different assays. Comments about the performance differences between the different used strategies have been included in the text.

To better communicate our results, we have added a new figure (Fig. 1B), which summarizes the identified proteins in the form of a dynamic localization network. This figure indicates which proteins were identified by which screen and whether or not they are novel or previously reported in the literature. The nodes of the network are color coded according to the pulsed-SILAC measurements (protein turnover).

Currently, we might know most of the components of the centrosome. However, the large number of additional candidates identified by the antibody-screen suggests that additional components remain to be identified. Thus, we would prefer to be rather cautious estimating the REAL total centrosome proteome when considering the exceptional dynamic distribution of proteins visiting or residing at centrosomes/spindle poles/basal bodies in various cells and tissues, during the cell cycle, during

development, and under different physiological and disease conditions. Antibodies representing the majority of human proteins are expected to be characterized by the Human Protein Altlas Project in 2015. This resource will certainly help to provide a realistic estimate. Our study provides some ideas about which proteins can be detected using which strategy. This type of information might prove useful for the design of future experiments.

To make the manuscript more predictive perhaps would be interesting to cluster candidate proteins according to their abundance and stability (separately) to suggest protein complexes; and then check whether those proteins display similar localization to centrosomes at different cell cycle stages.

We have clustered the turnover data and analyzed the resulting data by Gene Ontology enrichment. These data are shown in the supplemental Fig. S6. We argue that high turnover is likely to predict regulated proteins, a hypothesis that is further supported by protein localization data during the cell cycle.

Figure legends should be more detailed. As a non expert in the proteomics field I felt it was difficult to understand the figures related to that technique.

We have added additional information to all figure legends.

Figures with cells would gain from magnified insets from centrosome region as in some of them it is difficult to see the centrosomes.

Magnified insets from the centrosome region are now added to the images of Fig. 4, 5, and S7. Images in Fig. 6 are already magnified.

Referee #2 (Remarks to the Author):

This manuscript describes a study which uses a variety of techniques to identify novel components of the centriole. Of the novel proteins assigned to this sub cellular location, five are shown to be implicated in cell polarity. The manuscript also describes an elegant approach to investigate the range of stabilities within this set of centriole proteins.

This represents a thorough piece of work. The manuscript is mostly clearly written and well supported by experimental data. It falls down in terms of data analysis, where only choice proteins are annotated (figure 2G) and leaves the reader wondering why the method of analysis was chosen and how well the approach differentiates between organelles which have very similar densities and separation profiles on sucrose gradients. Before it is suitable for publication in this journal I suggest the following be considered which involves some re-analysis of the data.

1. The introduction launches straight into centrosome biology without really describing what a centrosome is. I think a sentence describing the function of the centrosome would not go amiss at the beginning of this section.

We have added a sentence about what a centrosome is before describing its biological role.

2. The authors chose one fraction which contains the highest abundance of centrosome proteins plus five neighboring fractions which contained some amount of these proteins for analysis using protein correlation profiling (PCP). It is not clear if the five fractions were from either side of the peak fraction; figure 1c would suggest only one fraction was more dense than the peak fraction, and data is not shown supporting that these fractions contained centrosomal proteins or indeed which centrosomal proteins they contained. I think this is somewhat of an oversight by the authors and it would be good to see some evidence of for example, a Western blot of the fraction depicting the total distribution of key centrosomal markers down the sucrose gradient.

To avoid simplify matters; we now refer to centrosome containing fractions rather than a peak centrosomal fraction and its neighboring fractions. In supplemental table S4 we provide detailed information about the identity of all detected proteins and their relative enrichment in each fraction from the double PCP-SILAC experiments. These data are equivalent to more than 1200 accurately measured Western blots comprising more than 100 known centrosomal proteins. Information about the 32 known centrosomal proteins associated with Fig. 2C is included in Table S1-3. We have added a reference in the Fig. 2 legend to clarify this point.

To address the total distribution of proteins, we have added supplemental Fig. S1C, which depicts the centrosome profile derived from the signal intensity of 1158 peptides from 42 centrosomal marker proteins measured by LC-MS from aliquots taken from 10 fractions down the sucrose gradient.

3. The authors talk about variable incorporation rates of SILAC amino acids in several proteins previously implicated as being part of the centrosome. Can the authors therefore be sure that the distribution patterns of proteins are not skewed by variable incorporation rates leading to false assignments? This should be clarified in the text.

We observed in our first test experiments that variable isotope incorporation rates indeed can lead to skewed distribution patterns in PCP-SILAC experiments possible leading to false assignments. Testing for complete isotope incorporation is therefore recommended. We have inserted a note about this issue in the material and method section.

4. The authors claim that the approach using a pooled internal standard coupled with SILAC labelling gives more accurate data than the traditional PCP. This point would be better supported if direct comparison of the two approaches was presented. This should require no more data collection, but analysis of distributions of proteins within the separate fractions without utilizing the internal standard.

We have performed this analysis based on the double PCP-SILAC experiment. The data is presented in supplemental Fig. S3 and indeed support our claim by a large margin.

5. The rationale behind carrying out the third experiment is not clear. The design of this experiment was not only to demonstrate the reproducibility of the distribution patterns of the centrosomal proteins in separation gradients but also to use correlated distributions between different gradients to add further evidence for similar behavior between centrosomal candidates. This message was lost a little in the manuscript and should be stated more clearly

The double PCP-SILAC experiment is unique. Thus, we have carefully explained the rationale behind carrying out this experiment in the manuscript as suggested.

6. Why do they choose 8 as a threshold? The authors could elaborate on this. Why are certain proteins in figure 2G not marked as candidates, even if the Mahalanobis distance is < 8 for the two replicates?

The threshold values for the organelle classification score was based on a value which included more than 95% of the known centrosomal proteins and less than 5% of obvious false positives in each experiment. This is now better explained in the main text and we have adjusted the plot.

7. Figure 2A-E - consider using thinner lines to make the figures clearer.

We have used thinner lines for the centrosomal proteins which indeed make the error bars more visible.

8. Can the authors also map onto this figure other organelles? There are plenty of grey dots some of which seem to cluster with the proteomeal and ribosomal proteins, but no attempt at annotation of these is made.

We have clustered all the complete protein profiles from the double PCP-SILAC experiments and performed Gene Ontology enrichment analysis. These data are presented in supplemental figure S4 and reveals that other protein complexes and cell organelles can be resolved. See also comments below.

9. On page 8 the authors talk about comparison of the Mahalonobis distances for proteins in 10 fractions in experiment three, but it is not clear what these then fractions refer to. This should be clarified.

The 10 fractions refer to the 2 x 5 fractions utilized in the double PCP-SILAC experiment. This is now clarified in the main text.

10. At the bottom of page 10, the authors describe in-vivo localization studies of 34 tested candidates, but it's not clear what this number refers as on page 8, second paragraph, 32 candidates present in all experiments and 12 additional candidates found only in a single experiment were described (also table 1). If only a selection of candidates were taken for analysis using this method, then the reasons behind the selection should also be stated.

Throughout the text we have carefully checked all the numbers of proteins retrieved from the different used methods. These numbers are now also summarized in Fig. 1. Proteins identified in the MS-screen were selected for further analysis based on the availability of antibodies, BAC clones, or the successful generation of cell pools stable expressing GFP-tagged proteins.

11. In the section describing antibody based screening again not all candidates seem to have been tested.

We did not aim for complete coverage of all candidates retrieved from the HPA-screen. The subset of proteins selected was based on a best candidate approach, i.e. proteins where immunefluorescence microscopy revealed two fine dots indicative of centrosome localization.

12. Further in this section differential staining patterns are described for the new centrosomal proteins, from large diffuse structures to distinct satellites. Within the correlation profiling did these proteins which had similar staining patterns have closer correlation with each other than with those who staining patterns was very different?

We did not see a significant difference between the profiles of centrosomal proteins observed with more diffuse staining patterns around the centrosomes and those with a more defined centrosome staining pattern (tested by hierarchical clustering). Since the profiles from the PCP-SILAC experiment reflect how the entire centrosome migrates in the sucrose gradient rather than individual proteins, we did not expect to see such differences unless (i) distinct subpopulations of biochemically purified centrosomes were resolved or unless (ii) the identified proteins distribute to partially resolved structures besides the centrosome. The few proteins observed with different profiles include alpha and beta tubulin, CEP55, and EB1. These proteins bind microtubules and are known to associate with other structures than the centrosome. Optimization of the gradient to separate such structures is important to avoid false negative identifications. See also comment to point 13.

13. How specific are the antibodies utilized? Could any of these antibodies have been used in Westerns to verify the distribution of proteins down the gradient? If some of the centrosomal proteins were present in more than one sub cellular location of differing densities, the alternative localizations would be missed in this study (see point 2 above).

It would indeed be interesting to assess the distribution of centrosomal proteins between different subcellular compartments or the cytosol using spatial proteins (Lamond lab) or global PCP-SILAC experiments. The starting material for such experiments should ideally be representative for all cellular proteins. This is not the case in the present study designed to profile the core components of one structure. When considering protein distribution type experiments we have been concerned about the detection of low abundant centrosomal proteins in an ocean of cellular proteins.

Referee #3 (Remarks to the Author):

This is a novel and convincing article, combining advanced mass spectrometry and high-throughput immunofluorescence microscopy, providing an important addition to the proteome of centrosomes. The methodologies used are novel, and themselves worth reporting, with the technique being firmly validated by expression of fusion proteins at endogenous levels, resulting in the identification of previously unidentified centrosomal proteins. That the analysis of these proteins is not taken further is justified by the advances made in the methodology.

Major Points:
(i) The major criticism I have is regarding the clarity of some of the text and the consistency of the referrals in the text to the Figures themselves (which is inconsistent in a number of cases and therefore makes it difficult to fully interpret some of the results):
Specifically, on p6, para2, we are told "profiles of the centrosomal proteins closely followed the CEP135 profile" and are pointed to Fig 1D. However, in the Figure legend, this appears to correspond to Fig 1E - with D and F pertaining to a different experiment. Either there is mislabeling of Fig 1, or there needs to be a clearer description in the associated text.

We have throughout the manuscript carefully checked for inconsistency and changed the text or figure legends if required.

Secondly, the text explaining Figure 3A and B is not clear enough. The two profiles are "two separate centrosome preparations" (Legend of Fig2) - and yet the profiles are markedly different (the ratio of A reduces as the fraction number increases, while B shows a profile similar to the original experiments in Fig1). I think I must be confused - in which case, the text needs to be clarified.

To better explain the PCP-SILAC experiment we have added a more detailed description in the figure legends. The profile in experiment 3A and 3B differ because the abundance of centrosomal protein in the fraction collected ahead of the peak centrosomal fraction in 3A was very low and therefore not included in the experiment. However, the shape of the profile in-between experiments is not critical. What matters is the relative enrichment of protein in each fraction making the PCP-SILAC methods rather robust.

Thirdly, the labels of Figure 3D-H do not seem to match up to the associated text on p9. Again, clarify or correct.

Corrected

Fourthly, the text states 44 candidate centrosomal proteins were isolated using PCP-SILAC "32 likely candidates observed in all PCP-SILAC experiments and 12 additional candidates supported by data in one or two of the experiments" (p8 para2). However, 47 are mentioned in the Discussion on p13, and only 43 are present in Table 1!

We would rather prefer not to mention numbers at all because they change whenever a candidate protein is published. We have carefully revised all numbers reported in the manuscript.

These may seem like minor points, but together it suggests the article was submitted without getting the details right - which are essential in order to fully review the science behind them.

Agree!

(ii) On a positive note, I applaud the Authors in their representation of the Mass Spectrometry data. It is very well laid out, and contains all the required information to justify their results. It's not often proteomics articles display this degree of ludicity!

(iii) On p10, at the end of the first paragraph, the authors suggest that their data regarding turnover rates of centrosomal proteins might be a useful parameter to estimate the time required for RNAi knockdown. I would like to see this hypothesis tested - either by looking through existing literature, or by a small-scale experiment.

Although we agree that it would indeed be interesting to test this hypothesis, we feel that it is beyond the scope of this paper to include a proper set of data, which could easily constitute a paper on it¥s own.

Minor Points:
(i) Introduction, p3; para2; line 7 - References pertaining to the roles of centrosomes in cellular processes need to be added here

We have reduced part of the previous introduction and added some additional references

(*ii) Results p6; para1; last line - It would be appropriate to include at least one representative picture of a 1D gel in the Supplementary data.*

A picture of the 1D gel from the 4-fraction PCP-SILAC experiments has been included as supplemental figure S1B.

(iii) Results p6; para2 (also Legend for Fig 2) - I could not see the triangles and asterisks supposed to distinguish the different isotope signals.
Labels have been added

(iv) Results p8; para 3; line 11 - "several of the newly identified centrosomal proteins were of relatively low abundance". It would be helpful to clarify this statement.

For simplicity, we have removed this statement.

(v) Results p8; para 3; last line - BB54 and NPHP4 do not appear in Table 1, but in the text it suggests they are. Were they not identified as candidates? Or have they been shown to be centrosomal by other research groups subsequent to the authors' previous study? It's a bit confusing.

These proteins have indeed been identified. The corresponding data is shown in Table S4 rather than Table 1.

(vi) Results p9; para 1 - The text and the Figure 3 should be consistent - either call/label the proteins TUBG1, 2, 3 etc. or TUBGCP1,2,3 etc. Currently there are different nomenclatures in the text and in Figure 3.

The proteins have now been labeled consistently in the manuscript.

(vii) There is a rather spurious reference to the retinal atrophy protein CCdC66 at the end of p9 - without discussing the significance of this example.

This reference has been placed in context in the discussion section.

(viii) Results p10; p2 - We are told in the final sentence on this page that 34 candidate proteins were tested for their localization using GFP-tagging. But there were 44 candidates in total. Which proteins were not tagged, and why (I'm assuming the cloning or expression did not work - but we need that information for completeness).

The number of proteins tested is summarized in Fig. 1A, B and in Table S4. The choice of proteins tested is primarily based on the availability of BAC clones and antibodies.

(*ix) Results p11; para 2 - the conjecture regarding the function of CCDC21 should be in the Discussion, not results.*

Relocated as suggested

(x) Table 1 - need to include what "AB" stands for in the Legend (presumably localization tested using antibodies?)

This information has been added.

2nd Editorial Decision 17 January 2011

Thank you for submitting your revised manuscript for our consideration. It has now been assessed once more by two of the original reviewers, who both consider the manuscript significantly improved in response to the initial comments, and thus in principle now suited for The EMBO Journal pending adequate revision of some remaining minor concerns.

I am therefore returning the study to you once more for a final round of revision, kindly inviting you to modify the manuscript text in response to the referee comments detailed below. When sending us the final version, please make sure to also include the following, in order to avoid any unnecessary further delays with the processing of the manuscript:
- an 'Author Contribution' statement in the text, to be included adjacent to the 'acknowledgements' section

Once we will have received your re-revised manuscript, we should then be able to swiftly proceed with its formal acceptance and publication.

Yours sincerely,
Editor
The EMBO Journal

———

REFEREE REPORTS:

Referee #1 (Remarks to the Author):

The manuscript "Complementary proteomics methods identify novel components localizing asymmetrically to centrosomes" by Jakobsen et al has improved substantially in this revised version and therefore should be published in EMBO J. I just have some minor comments:
-In legend of Figure 1, explain what the asterisk means.
-Figure 5, explain what is labeled in the nasopharynx tissue
-Figure 6-there is no legend regarding Mphosph9- what is in which colour? Does it localize to cilia? Figure 6E has no legend.
-Some references in the text are incomplete, in particular regarding the year.

Referee #2 (Remarks to the Author):

This re submitted manuscript describes a study which uses a variety of techniques to identify novel components of the centriole. Of the novel proteins assigned to this sub cellular location, five are shown to be implicated in cell polarity. The manuscript also describes an elegant approach to investigate the range of stabilities within this set of centriole proteins.

The revised manuscript is greatly improved over the first version I read. It is much clearer, and the data analysis is much more thorough. In terms of my initial comments to the authors, I'm satisfied they have addressed the majority of my concerns. The one concern I still have centres round the interpretation of data generated by use of the HPA antibodies. My original point was as follows:

1. How specific are the antibodies utilized? Could any of these antibodies have been used in Westerns to verify the distribution of proteins down the gradient. If some of the centrosomal proteins were present in more than one sub cellular location of differing densities, the alternative localizations would be missed in this study.
Answer from authors: It would indeed be interesting to assess the distribution of centrosomal proteins between different subcellular compartments or the cytosol using spatial proteins (Lamond lab) or global PCP-SILAC experiments. The starting material for such experiments should ideally be representative for all cellular proteins. This is not the case in the present study designed to profile the core components of one structure. When considering protein distribution type experiments we have been concerned about the detection of low abundant centrosomal proteins in an ocean of cellular proteins.
My concern is still that if these antibodies cross react with unknown proteins, this could give rise to false positives in the HPA screen. The researchers might think they know the identity of the protein in the centrosome as it picks this up in an immunofluorescence screen, but if the centrosomal protein is cross reacting, a protein could be wrongly assigned. There is no mention in the text of how the scoring was done in terms deciding whether the bulk of the fluorescence was centrosomally associated. I would suggest some additional text is added to the manuscript describing the stringency employed in the immunofluorescence scoring methods.

Minor points:
1. I don't think the numbering scheme in the new figure 1 is particularly clear. I think a slightly clearer description of what they mean should be added to the legend for this figure.
2. The Mahalanobis distance equation in the supplemental data looks a little odd. The formatting of this should be checked.
3. The legend for supplementary figure 3 needs a color coding key for the ROC curves
4. Supplemental figure 2 b is unclear - please clarify what dish 3/1 and 2/1 refer to in the legend

Print Email

2nd Revision - authors' response 09 February 2011

Point-by-point response to the referees

Once again, we would like to acknowledge the reviewers for their constructive comments and suggestions. Please find below our response to each point:

Referee #1 (Remarks to the Author):

The manuscript "Complementary proteomics methods identify novel components localizing asymmetrically to centrosomes" by Jakobsen et al has improved substantially in this revised version and therefore should be published in EMBO J. I just have some minor comments:
-In legend of Figure 1, explain what the asterisk means.
-Figure 5, explain what is labeled in the nasopharynx tissue -Figure 6-there is no legend regarding Mphosph9- what is in which colour? Does it localize to cilia? Figure 6E has no legend.
-Some references in the text are incomplete, in particular regarding the year.

ï We have extended the Figure 1 legend to more carefully explain the meaning of the various numbers in the figure, including the asterisk: "The number of 'profiled' proteins refers to those quantified in all fractions out of those quantified in at least one fraction. The number of 'MS-candidate' and 'MS-known' refers to those annotated as novel or known centrosomal proteins, respectively, out of those scored as likely centrosomal proteins based on the PCP-SILAC method (*) or those tested by microscopy. References to the relevant tables (T) and figures are shown in

brackets.

ï We have added information to the legends of Fig. 6 to address the questions related to this figure: "(D) Antibody staining of GFP-C3orf34 in HeLa cells at the G1/S phase and MPHOSPH9 in RPE cells at the G0 phase. Co-staining with the mother centriolar marker protein ODF2 (Cy5) indicates that GFP-C3orf34 associate with the mother centriole. Co-staining with anti-acetylated tubulin (red) in ciliated hTERT-RPE1 cells suggests that MPHOSPH9 (green) localize proximal at the mother centriole and distal and proximal at the daughter centriole. DNA were stained with DAPI, yellow indicates coincidence of green and red signals. Bars, 5 m. (E, F) Interpretation of the staining patterns of tubulin (red) and asymmetrically localized candidate proteins (green) in duplicated centrosomes and basal bodies."

ï Novel references where the year of publication was not extracted by EndNote have now been corrected.

Referee #2 (Remarks to the Author):

This re submitted manuscript describes a study which uses a variety of techniques to identify novel components of the centriole. Of the novel proteins assigned to this sub cellular location, five are shown to be implicated in cell polarity. The manuscript also describes an elegant approach to investigate the range of stabilities within this set of centriole proteins.
The revised manuscript is greatly improved over the first version I read. It is much clearer, and the data analysis is much more thorough. In terms of my initial comments to the authors, I'm satisfied they have addressed the majority of my concerns. The one concern I still have centres round the interpretation of data generated by use of the HPA antibodies. My original point was as follows:

1. How specific are the antibodies utilized? Could any of these antibodies have been used in Westerns to verify the distribution of proteins down the gradient. If some of the centrosomal proteins were present in more than one sub cellular location of differing densities, the alternative localizations would be missed in this study.
Answer from authors: It would indeed be interesting to assess the distribution of centrosomal proteins between different subcellular compartments or the cytosol using spatial proteomics (Lamond lab) or global PCP-SILAC experiments. The starting material for such experiments should ideally be representative for all cellular proteins. This is not the case in the present study designed to profile the core components of one structure. When considering protein distribution type experiments we have been concerned about the detection of low abundant centrosomal proteins in an ocean of cellular proteins.
My concern is still that if these antibodies cross react with unknown proteins, this could give rise to false positives in the HPA screen. The researchers might think they know the identity of the protein in the centrosome as it picks this up in an immunofluorescence screen, but if the centrosomal protein is cross reacting, a protein could be wrongly assigned. There is no mention in the text of how the scoring was done in terms deciding whether the bulk of the fluorescence was centrosomally associated. I would suggest some additional text is added to the manuscript describing the stringency employed in the immunofluorescence scoring methods.

We agree with the concerns of reviewer 2 that the antibodies generated and used in this study potentially cross-react with non-specific proteins leading to false positive identification/validation. On a positive note, the HPA antibodies are epitope purified, which significantly reduce the background of non-specific antibodies present in pre-immune sera having the potential to stain centrosomal proteins. Fortuitous cross-reactions of antibodies are, however, not eliminated in this way. In the previous version of the manuscript we have been openly cautious about this problem (see notes 1-3 below). To underscore the importance of this issue we have added additional sentences in the result and method sections (see notes 4-5 below) together with the following reference: Nigg EA, Walter G, Singer SJ. (1982) On the nature of cross reactions observed with antibodies directed to defined epitopes. Proc Natl Acad Sci U S A. 79, 5939-43. We would also like to note that all antibodies and immunofluorescence images will be publicly available for further investigation and that our manuscript illustrates how complementary experimental approaches might contribute to more confident subcellular localization assignment.

To address the stringency employed in the immunofluorescence scoring of centrosomal proteins we have inserted the following sentence: "The immunofluorescence images were scored in comparison to a negative control. Any staining stronger than the control was assigned one or multiple location as indicated in Table S5. Centrosome association were assigned when at least two cells displayed the characteristic two-dot staining pattern of disengaged centrioles or duplicated centrosomes and/or when the protein staining clearly co-localized with -tubulin at microtubule organising centres."

We have confirmed the antibody specificity for one of the more remarkable candidates by RNAi: "The antibody specificity was supported by reduced staining in cells depleted for MPHOSPH9 by esiRNA (data not shown)".

1. "Inventory analysis by fluorescence microscopy requires the availability of antibodies or cell pools expressing GFP-fusion proteins and is dependent on the quality, in-vivo behaviour and detectability of these reagents. Inventory analysis by mass spectrometry is challenged by lists of proteins crowded with unrelated entries in the midst of genuine components, even when analyzing high-purity preparations. As a consequence, the identified proteins provide corroborating rather than unequivocal evidence of organelle association."

2. "Conversely, proteins identified by the HPA-screen might represent cases of incorrect annotation caused by antibodies cross reacting with epitopes on different proteins (Nigg et al., 1982)."

3. Although we confirmed -tubulin co-localization for a subset of the identified HPA-candidates in the presence and absence of nocodazole, further experiments are necessary to validate antibody specificity and to annotate accurate subcellular localization for all the HPA-candidates. Obvious experiments are staining before and after knockdown and counterstaining with a more diverse set of centrosomal markers such as PCM1 staining centrosomal satellites.

4. "Although the HPA-antibodies were epitope purified and evaluated for specificity and assigned reliability score (see methods), we cannot exclude false assignments due to fortuitous cross-reactions (Nigg et al., 1982). Hence, caution needs to be exerted when interpreting the results of the antibody screen and additional experiments are required to confirm these candidates.

5. "The antibodies were epitope purified and evaluated for specificity and assigned a reliability score (supportive, uncertain and not supportive) in protein arrays, Western blot, immunohistochemistry and immunofluorescence microscopy. The validation data as well as the antigen sequence is available in the Human Protein Atlas. See www.proteinatlas.org/about/assays+annotation and www.proteinatlas.org/about/quality+scoring for more information."

Minor points:
1. I don't think the numbering scheme in the new figure 1 is particularly clear. I think a slightly clearer description of what they mean should be added to the legend for this figure.
2. The Mahalanobis distance equation in the supplemental data looks a little odd. The formatting of this should be checked.
3. The legend for supplementary figure 3 needs a color coding key for the ROC curves
4. Supplemental figure 2 b is unclear - please clarify what dish 3/1 and 2/1 refer to in the legend

ï We have extended the figure legends to address these comments:

ï #1: "The number of ‘profiled’ proteins refers to those quantified in all fractions out of those quantified in at least one fraction. The number of ‘MS-candidate’ and ‘MS-known’ refers to those annotated as novel or known centrosomal proteins, respectively, out of those scored as centrosomal proteins based on the PCP-SILAC method (*) or those tested by microscopy. References to the relevant tables (T) and figures are shown in brackets."

ï #2: The formatting has been corrected.

ï #3: "Blue ROC-curve: PCP-SILAC, red ROC-curve: PCP based on the median of profiles for all peptides representing the same protein, green ROC-curve: PCP based on profiles calculated from the sum of intensity of all peptides representing the same protein."

ï #4: "The visualization window display protein enrichment profiles calculated from the isotope signal intensity ratios medium/light (dish 2/1) and heavy/light (dish 3/1) in the double PCP-SILAC experiment.

