

Combined heterogeneity in cell size and deformability promotes cancer invasiveness

Asadullah, Sandeep Kumar, Neha Saxena, Madhurima Sarkar, Amlan Barai and Shamik Sen
DOI: 10.1242/jcs.250225

Editor: Andrew Ewald

Review timeline

Original submission:	16 June 2020
Editorial decision:	28 July 2020
First revision received:	4 January 2021
Accepted:	29 January 2021

Original submission

First decision letter

MS ID#: JOCES/2020/250225

MS TITLE: Combined heterogeneity in cell size and deformability promotes cancer invasiveness

AUTHORS: Asadullah, Sandeep Kumar, Neha Saxena, Madhurima Sarkar, Amlan Barai, and Shamik Sen

ARTICLE TYPE: Research Article

We have now reached a decision on the above manuscript.

To see the reviewers' reports and a copy of this decision letter, please go to: <https://submit-jcs.biologists.org> and click on the 'Manuscripts with Decisions' queue in the Author Area. (Corresponding author only has access to reviews.)

As you will see, the reviewers are excited by the study, especially its integrated experimental / computational nature. They also raise a number of substantial criticisms that prevent me from accepting the paper at this stage. They suggest, however, that a revised version might prove acceptable, if you can address their concerns. If you think that you can deal satisfactorily with the criticisms on revision, I would be pleased to see a revised manuscript. We would then return it to the reviewers.

We are aware that you may currently be unable to access the lab to undertake experimental revisions. If it would be helpful, we encourage you to contact us to discuss your revision in greater detail. Please send us a point-by-point response indicating where you are able to address concerns raised (either experimentally or by changes to the text) and where you will not be able to do so within the normal timeframe of a revision. We will then provide further guidance. Please also note that we are happy to extend revision timeframes as necessary.

Please ensure that you clearly highlight all changes made in the revised manuscript. Please avoid using 'Tracked changes' in Word files as these are lost in PDF conversion.

I should be grateful if you would also provide a point-by-point response detailing how you have dealt with the points raised by the reviewers in the 'Response to Reviewers' box. Please attend to all of the reviewers' comments. If you do not agree with any of their criticisms or suggestions please explain clearly why this is so.

Reviewer 1*Advance summary and potential significance to field*

This is a compelling manuscript, demonstrating the power of combining computational modeling with traditional wet lab experiments to answer a critical biological question. I especially congratulate the authors in their clear explanation of the computational modeling so that it is accessible to non-modelers. This manuscript evaluates the role of cellular phenotypic heterogeneity in invasion with a particular focus on cell size and deformability. The authors present clear in vitro data, modeling of small/large/soft/stiff cells in various amounts of mixing, and present an interesting hypothesis, with limited data, regarding the stem cell nature of these invasive cells. Overall, it is an excellent manuscript, but there are a number of points for improvement and revision that I will outline below:

Comments for the author

- In the introduction and first in vitro studies, the authors focus on the role of cell size and deformability. Throughout the modeling, however, one of the major components is cell-cell adhesion. Simply introducing such an idea part-way through the manuscript is confusing and is not supported by cell line data provided in the manuscript (e.g., not described in the biophysical properties of the breast cancer cell lines). Why is cell-cell adhesion considered in the modeling, but not in the in vitro studies?
- There are interesting observations regarding the “spread area” of MCF7 and MDA-MB-231 cells, and this data is collected from the same time lapse series as the motility data. What is the interpretation of the “spread area” of the cells? Can the authors speculate about the z-axis morphology of the cells (e.g., rounded vs flat) and how it may contribute to invasiveness and motility?
- In a similar line of thought, it is clear, both from the quantitated data in Fig 1 as well as the Supplemental video 1, that there is a great deal of heterogeneity cell-to-cell in “spread area” as well as variation in spread area in the same cell over time. In looking at the tracks included in the SVideo1, movement of the cell appears to be assessed by displacement of the nucleus. In these large cells with non-cuboidal morphology however, this may not be an adequate measure of movement as the rest of the cell body clearly moves in the nearby space (e.g., high cell movement, but minimal movement of the nucleus). It would be valuable to understand cell motility (a surrogate for invasion in the subsequent computational modeling) in the framework of cell morphology and size. Have the authors addressed the changes in 2D morphology in interpreting the differences in motility?
- The authors only plot cell speed vs area for cells with less than or equal to 50% change in cell size. While this experiment certainly aligns with the authors’ hypothesis that small cells (presumably with small changes in size) are the most likely invading cells. By this virtue, however, it is important to present the complete dataset, including cells with changes in cell size. If the same trends hold, it would be valuable information to make a more universal statement of velocity vs size, and if it is not the same trend, it is critical information for the remainder of the manuscript.
- The authors also investigate a publicly available RNA-seq database to assess transcriptional heterogeneity. However, the dataset they selected, GSE75688, is from breast tumors, not from cell lines, and includes both primary tumor as well as LN specimens. While it is valuable to consider human disease directly, an analysis of MCF7 and MDA-MB-231 would be more valuable to directly compare to the biophysical data presented in Figure 1. Such datasets are widely available.
- Again, following the clearly presented in vitro data, the authors begin computational modeling with a focus on “varying the extent of cell-cell adhesion.” This is a non-sequitur and it is unclear the connection of cell-cell adhesion to cell size or deformability. As this apparently is a critical variable for the remainder of the modeling, there should be some discussion of 1) it’s relative importance and 2) the cell-cell adhesion of the in vitro cell lines that inform the computational modeling (MCF7 and MDA-MB-231).

- The authors include cell size in the computational modeling. How is cell size / observed motility adjusted for? For example, a larger cell will necessarily occupy more pixels than a small cell. Is d and D calculated the same regardless of how much space a cell occupies, even with the clear differences in movement of the cell itself as well as distance actually covered (as above in the cell tracking in vitro)? In drawing conclusions about the relationship of cell size and physical displacement, a consideration of the total area covered by the cell is necessary.
- It is unclear why, in the absence of a chemotaxis gradient (as is applied in Fig 3), net translocation (D) is assessed at all. In the absence of a stimulus, the directionality of movement - which is an unstated assumption underlying the measurement of D - is not necessarily expected. Therefore, that there is no correlation with D is unsurprising. The authors should clarify why D was measured (vs only total distance, d), and clarify their definitions of “motility” (d) vs “invasion” (D) in this experiment.
- Figure 2 itself needs revision, including a legends for Fig 2A and a clear statement in the legend/on the figure for what the red and blue circled regions are in the tSNE plots.
- Again, the purpose of focusing on cell-cell adhesion in the computational modeling is unclear. “...these analyses implicate intermediate cell-cell adhesion in promoting invasion efficiency of cells in the presence of directional cues.” Likewise, throughout the manuscript the authors use the terms “clustering” in invasion/motility - a more clear explanation of the hypothesis and how the results support/do not support this framework is needed.
- The authors propose that the small invasive cells are CSCs and state “Collectively, these results suggest that small sized cells of varying deformability are maximally invasive, and may correspond to a mix of mesenchymal and hybrid CSCs.” Based on the data provided, this statement is much too strong.
- It is unclear how the statistics in Fig 5 were determined - “ $n > 190$ cells per condition pooled from $N = 3$ independent experiments.” How many cells were assessed in each experiment?
- The gating strategy for sorting CD44hi/CD24- and CD44hi/CD24+ cells should be included in supplement. Percent of the population in the provided “hCSC” and “mCSC” gates should be included.
- The spheroid assay is creative, but raises a number of questions. First, Fig 6 G figure legend incorrectly states “Temporal evolution of zone-wise distribution...” It is not clear that what is observed is evolution at all - and indeed, in such a short time frame, such evolution would be rapid indeed with limited generation time.
- Because the spheroid assay does not utilize cell tracking, it is impossible to determine whether the observed CD44 intensity at the outer region of the spheroid is due to physical movement of the cells outward or due to transdifferentiation of the cells at the outer edge.
- The CD44 and CD24 expression of the cells in 2D was presented in Fig 5. Is such information available for the cells grown in a spheroid, either by flow cytometry or by IF? In Fig 5, the data shows a clear difference in size and stiffness between mCSC (CD24-) and hCSC (CD24+). To complete their model, therefore, such characterization of the “CD44 intensity” cells in Fig 6 spheroid studies is needed. While in the discussion the authors state “Since the proportion of mCSCs is significantly higher than that of hCSCs, it is likely that a majority” of the CD44+ cells in the spheroid are indeed mCSCs - this is a broad statement not supported by the provided data.
- The concluding in vitro experiments and the presented model fail to include the role of cell-cell adhesion. This is puzzling because of the importance it played in the computational modeling. I advise either removing it from the modeling or including additional rationale and data related to cell-cell adhesion in the final studies.
- Throughout the discussion, I would urge the authors to soften their often too-bold statements. For example,

“...we have shown that cancer invasiveness is maximum when heterogeneity in both cell size and cell deformability are present.” While this was computationally modeled, this was not shown in in vitro or in vivo studies.

Likewise, “Subsequent experiments identify these cells as CSCs.” And later “we experimentally demonstrate to be CSCs.” While the CD44+ cells they assess have some of the same properties, they do not definitively determine that the cells from the computational model are necessarily CSCs.

Specific points for clarification in the materials and methods section:

- The specific protocol used for image analysis (e.g., plugins or workflow) should be included for all measures (cell size, change in cell spread over time, movement).
- Confirm that the MDA-MB-231 cells were lifted with trypsin prior to flow cytometry. This is an issue due to CD44 and CD24 antigen retention following trypsinization vs other cell disassociation methods.
- The specific antibiotic-antimycotic solution should be included.
- Flow cytometry analysis strategy should be detailed.
- Specific information about n per experiment and total number of experiments for in vitro would should be included.

Reviewer 2

Advance summary and potential significance to field

Combined heterogeneity in cell size and deformability promotes cancer invasiveness

Asadullah et al.

The authors present a needed examination of phenotypic heterogeneity and its role during invasion. In particular the authors evaluate the influence of cell size and cell deformability in breast carcinoma invasion. The paper utilizes both experimental and computational approaches (extension of the Cellular Potts Model). Results suggest that combined heterogeneity in size and deformability enhance invasiveness. The authors further suggest that these characteristics are enriched in CSC populations and that CSCs are more robust leaders at the invasive front. There are many interesting advances and innovation in the manuscript. The framework is interesting and the use of the CPM/GGH framework certainly has value for the questions being asked. Enthusiasm, however, is somewhat tempered by a few key issues (outlined below). Yet, it is this reviewer’s opinion that the manuscript could be acceptable for publication in JCS following key changes to the text and likely some additional data or analysis. The following comments are meant as constructive criticisms to help improve the manuscript.

Comments for the author

Specific Comments:

1. There are issues with the Introduction and reference list. First the reference list is not formatted in a manner that allows it be evaluated. Following the JCS formatting would help and make the reference list accessible. Furthermore, the Introduction appears to be under-referenced and key primary literature on CSCs size and behavior is missing (e.g. Li et al., *Seminars in Cancer Biology*, 2015; De Paiva et al., *Stem Cell*, 2016; Ray et al., *Biophysical Journal* 2017 etc.).
2. The decision to use 5 bins of size and deformability conditions appears reasonable, but a much better discussion of the rationale behind the cutoffs would strengthen the manuscript. Did the data in Figure inform these boundaries? Do sensitivity and/or statistical studies need to be performed to better define the cutoffs etc.?

3. The model description is quite incomplete. This reviewer had to study Reference 25 in some detail to interpret many aspects of the current work. Even in Ref 25 the model is not explained in sufficient detail. A much more in depth and robust description of the modeling framework and implementation is needed. It would also greatly strengthen the manuscript, particular for the more general cell biology audience, to connect the model framework to particular cell behaviors or outcomes. For instance, how the framework handles size in terms of the energy formulation; AND for things like the statement “Together, these analyses implicate intermediate cell-cell adhesion in promoting invasion efficiency of cells in the presence of directional cues”, Why? Please connect the model and outcomes - How does this connect to the model parameters and the known cell biology. As presented it is quite abstract.

4. In line with comment 3, the statement “The last term in the energy expression was included to model chemotaxis of cells in the direction of the chemoattractant gradient (15), with (symbol) representing the effective chemical potential, and $v(\text{target})$ and $v(\text{source})$ representing the concentrations of chemoattractant at target and source pixel respectively. This term was used in the energy expression only for simulating cancer invasion in the presence of a chemotactic field.” This seems like it needs to parameterization / sensitivity studies over a range of values etc. The behaviors are likely not linear so difference gradients could produce different outcomes.

5. The data in Figure 1 is already, for the most part, widely available in the literature, and the findings quite expected. That said it is clear how this can be used to set the stage for the manuscript, but it requires better development and connectivity to the subsequent theoretical work. This may have additional relevance since MDA-MB-231 cells do not form any or strong cell-cell adhesion (they are a strong EMT phenotype at baseline). Furthermore, the deformability difference is very small. While statistically significant, is it biologically significant?

6. How do the authors account for size and other phenotype behaviors? For instance, are the smaller cells just more amoeboid in phenotype and with different cell dynamics and physical modes of motility? What is steric and what is migration phenotype? This should be addressed as much as possible in the text and any possible analysis.

7. It is not clear that the model can really handle leader cell, collective invasive front, and single and collective follower dynamics as these processes have been shown to include quasi-ballistic behavior due to crowding/jamming. Some discussion or additional analysis in this area would strengthen the manuscript.

More minor comments:

8. The Figures are very crowded and at time difficult to easily interpret. Figure 4 in particular is difficult to interpret.

9. Translocation and distance are not adequately described.

10. All 231 cell appear to be CD44 high (consistent with work by Massague), suggesting a high CSC baseline state for these cells. As such, is CD44 staining for leaders cells to be CSC robust enough? A bit more discussion here would be useful and also clear analysis of how CD44 levels are being parsed out in immunofluorescence images.

11. The PCA and t-SNE plots are interesting, but also need better connectivity presentation. Why not show the PCA data? Also, the red and blue designations need better description particularly in the figure text.

Reviewer 3

Advance summary and potential significance to field

The authors have examined the role of two axes of phenotypic heterogeneity - cell size and deformability - in driving cancer invasiveness. This integrative computational-experimental study is well-done, very timely and contributes to the growing literature on implications of non-genetic (i.e. phenotypic) heterogeneity in cancer progression.

Comments for the author

I request the authors to address the following minor queries/clarifications:

1. Fig 1, S1:

- a) The authors have tracked individual cells over a period of 24 hour for quantifying cell size and shape changes. Was the cell cycle status considered to be a factor here? If not, the authors should mention this as a limitation of their analysis. Cell cycle heterogeneity can be a confounding factor (Spencer et al. Nature 2009), and cells are expected to change their biophysical traits during division. A recent study quantified the effects on EMT status as well during cell division (Tripathi et al. PLoS Comp Biol 2020).
- b) Fig 1E: The authors should explain how “maximum change in cell size” is calculated. They can consider plotting some distributions of cell size changes.
- c) Fig S1: Can the authors dissect the contribution of intra-tumor and inter-tumor variability here? To strengthen their arguments later about connection of size/ deformability with EMT and/or stemness, they can consider calculating EMT and/or stemness scores/enrichment through single-sample GSEA for these cells.
- d) Fig 1H: The authors mention that there is no correlation between cell size and motility. Can this argument be substantiated by relevant statistical test?
- e) The authors end this section with “breast cancer cell lines have heterogeneity at phenotypical as well as genetic level”. However, there is no genetic status (mutation etc.) analysis done; the authors should replace “genetic” with “gene expression”.

2. Fig 2:

- a) 2F: Can the authors explain for what all parameters the t-SNE is performed? Also, because of its stochastic nature, t-SNE can give different clusters in each iteration. How do the authors gain confidence in the presented trends for t-SNE?
- b) 2D: The authors suggest that combined heterogeneity in deformability and cell size led to increased average net translation, but heterogeneity in only one axis reduced invasiveness. Can the authors offer a biophysical explanation for it? Moreover, relevant statistical test results are needed for a stronger argument.

3. Fig 3:

- a) The authors should include the quantitative details of the gradient used.
- b) The authors mention that chemotaxis is fastest at intermediate cell-cell adhesion. Can they attempt to discuss their results vis-à-vis observations that intermediate cell-cell adhesion can promote migration of cells as clusters (Bocci et al. Cancer Res 2019) and that clusters can be better at chemotaxis as compared to individual cells (Camley et al. PRL 2016).

4. Fig 5:

- a) The authors have identified two different kinds of CSCs (CD44+ CD24+ - hCSCs, and CD44+ CD24- mCSCs) and suggest that mCSCs may be the one enriched at invasive edge. A recent computational modeling effort had similar conclusions (Bocci et al. PNAS 2019); the authors should discuss similarities and differences with this study. Moreover, other relevant studies on CD44+ CD24+ in breast cancer cells should be cited - Grosse-Wilde et al. PLoS ONE 2015, Goldman et al. Nat Comm 2014, Boareto et al. J R Soc Interface 2016, given the predominance of CD44+ CD24- as the established CSC signature in breast cancer.
- b) Experimentally, did the authors observe interconversion among these two subsets of CSCs, as seen in Liu et al. Stem Cell Reports 2014?

First revision

Author response to reviewers' comments

MS ID#: JOCES/2020/250225

MS TITLE: Combined heterogeneity in cell size and deformability promotes cancer invasiveness

AUTHORS: Asadullah ., Sandeep Kumar, Neha Saxena, Madhurima Sarkar, Amlan Barai, and Shamik Sen

ARTICLE TYPE: Research Article

We thank the editor and the reviewer for appreciating our work and providing feedback for further strengthening our paper. Below we detail our responses to the reviewers' criticisms.

Response to Reviewer 1 Comments:

This is a compelling manuscript, demonstrating the power of combining computational modeling with traditional wet lab experiments to answer a critical biological question. I especially congratulate the authors in their clear explanation of the computational modeling so that it is accessible to non-modelers. This manuscript evaluates the role of cellular phenotypic heterogeneity in invasion with a particular focus on cell size and deformability. The authors present clear in vitro data, modeling of small/large/soft/stiff cells in various amounts of mixing, and present an interesting hypothesis, with limited data, regarding the stem cell nature of these invasive cells. Overall, it is an excellent manuscript, but there are a number of points for improvement and revision that I will outline below:

We thank the reviewer for appreciating the merit of our work. We have made a sincere attempt to address the concerns raised by the reviewer. Our responses are detailed below:

Reviewer 1 Comments for the Author:

- In the introduction and first in vitro studies, the authors focus on the role of cell size and deformability. Throughout the modeling, however, one of the major components is cell-cell adhesion. Simply introducing such an idea part-way through the manuscript is confusing and is not supported by cell line data provided in the manuscript (e.g., not described in the biophysical properties of the breast cancer cell lines). Why is cell-cell adhesion considered in the modeling, but not in the in vitro studies?

We have incorporated cell-cell adhesion in our simulations as it is an intrinsic property of cells and relevant to cancer. In our initial experimental data, we have shown biophysical heterogeneity in MCF-7 and MDA-MB-231 cells. While MCF-7 cells are epithelial in nature, MDA-MB-231 cells are mesenchymal in nature.

In our simulation framework, for a fixed value of cell-matrix adhesion energy ($J_{CM} = 16$) the cell-cell adhesion energy (i.e., J_{CC}) can be tuned to simulate collective cell migration ($J_{CC} = 1$) as well as single cell migration ($J_{CC} = 40$). $J_{CC} = 16$ corresponds to intermediate cell-cell adhesion where both single cells and cell clusters are observed.

As per the reviewer's suggestion, we have now added a paragraph in the Introduction on cell-cell adhesion and its implications in cancer invasion. We have also stained cells for the cell-cell adhesion protein E-cadherin and added the data in Figs. 1A, B.

- There are interesting observations regarding the "spread area" of MCF7 and MDA-MB-231 cells, and this data is collected from the same time lapse series as the motility data. What is the interpretation of the "spread area" of the cells? Can the authors speculate about the z-axis morphology of the cells (e.g., rounded vs flat) and how it may contribute to invasiveness and motility?

"Spread area" serves as a measure of cell size. Cells which exhibit low levels of change in spread area may correspond to "phenotypically stable" sub-populations. In comparison to MCF-7 cells, MDA-MB-231 cells are more spread and flatter, i.e., nuclear height is less. Since nucleus is the largest and stiffest organelle inside a cell, nuclear deformation is the rate limiting factor for 3D invasion. MDA-MB-231 cells, which are more motile than MCF-7 cells, are also more invasive. In general, there is a positive correlation between motility and invasiveness in most cancer cells.

- In a similar line of thought, it is clear, both from the quantitated data in Fig 1 as well as the Supplemental video 1, that there is a great deal of heterogeneity cell-to-cell in "spread area" as well as variation in spread area in the same cell over time. In looking at the tracks included in the SVideo1, movement of the cell appears to be assessed by displacement of the nucleus. In these large cells with non-cuboidal morphology, however, this may not be an adequate measure of movement as the rest of the cell body clearly moves in the nearby space (e.g., high cell movement,

but minimal movement of the nucleus). It would be valuable to understand cell motility (a surrogate for invasion in the subsequent computational modeling) in the framework of cell morphology and size. Have the authors addressed the changes in 2D morphology in interpreting the differences in motility?

We acknowledge the reviewer's concern. We have now re-analysed our time-lapse movies based on tracking of the cell centroid. This re-analysis has changed the magnitudes of cell speed, though the trend of MDA-MB-231 cells being more motile than MCF-7 cells is still true.

- The authors only plot cell speed vs area for cells with less than or equal to 50% change in cell size. While this experiment certainly aligns with the authors' hypothesis that small cells (presumably with small changes in size) are the most likely invading cells. By this virtue, however, it is important to present the complete dataset, including cells with changes in cell size. If the same trends hold, it would be valuable information to make a more universal statement of velocity vs size, and if it is not the same trend, it is critical information for the remainder of the manuscript. As suggested by the reviewer, we have now plotted the whole dataset. As before, there seems to be no correlation between cell size and cell speed. We have also added the values of the Pearson correlation coefficient between cell size and cell speed in the Figure.

- The authors also investigate a publicly available RNA-seq database to assess transcriptional heterogeneity. However, the dataset they selected, GSE75688, is from breast tumors, not from cell lines, and includes both primary tumor as well as LN specimens. While it is valuable to consider human disease directly, an analysis of MCF7 and MDA-MB-231 would be more valuable to directly compare to the biophysical data presented in Figure 1. Such datasets are widely available.

We thank the reviewer for this suggestion. Analysis of patient database allowed us to map heterogeneity in biophysical properties, stemness and EMT. We have also established a correlation between biophysical properties with Stemness and EMT scores.

Based on the reviewer's suggestion, we have now incorporated analysis of heterogeneity in biophysical properties, stemness and EMT scores based on single-cell RNAseq of MCF-7 and MDA-MB-231 cells in Supp. Fig. 2.

- Again, following the clearly presented in vitro data, the authors begin computational modeling with a focus on "varying the extent of cell-cell adhesion." This is a non-sequitur and it is unclear the connection of cell-cell adhesion to cell size or deformability. As this apparently is a critical variable for the remainder of the modeling, there should be some discussion of 1) its relative importance and 2) the cell-cell adhesion of the in vitro cell lines that inform the computational modeling (MCF7 and MDA-MB-231).

We acknowledge the reviewer's concern. We have now added a paragraph in the Introduction on cell-cell adhesion and its importance in cell invasion. We have also added E-cadherin staining in Fig. 1.

Incorporating cell-cell adhesion as a parameter in our simulations enables us to model invasive behavior of individually migrating cells as well as collectively migrating cells. While $JCC = 1$ is expected to recapitulate invasive behavior of MCF-7 cells, $JCC = 40$ is expected to recapitulate the invasive behavior of MDA-MB-231 cells which are mesenchymal in nature.

We have now made this more explicit in the second Results section.

Our simulations suggest that the effect of phenotypic heterogeneity on invasiveness is dependent on the extent of cell-cell adhesion.

- The authors include cell size in the computational modeling. How is cell size / observed motility adjusted for? For example, a larger cell will necessarily occupy more pixels than a small cell. Is d and D calculated the same regardless of how much space a cell occupies, even with the clear differences in movement of the cell itself as well as distance actually covered (as above in the cell tracking in vitro)? In drawing conclusions about the relationship of cell size and physical displacement, a consideration of the total area covered by the cell is necessary. Since both d and D are calculated based on changes in the position of the cell centroid, these

measures are insensitive to cell size.

We have now explicitly mentioned this in the Results section.

- It is unclear why, in the absence of a chemotaxis gradient (as is applied in Fig 3), net translocation (D) is assessed at all. In the absence of a stimulus, the directionality of movement - which is an unstated assumption underlying the measurement of D - is not necessarily expected. Therefore, that there is no correlation with D is unsurprising. The authors should clarify why D was measured (vs only total distance, d), and clarify their definitions of “motility” (d) vs “invasion” (D) in this experiment.

In random cell motility experiments, cells always exhibit some degree of directional persistence. To account for this intrinsic polarization, we have assumed cells exhibit self-propelled motion, i.e., the cell velocity vector in a given iteration is defined as the average of the velocity vectors in previous 10 iterations. It is this term which contributes to migration persistence. The presence of a chemotactic gradient further enhances migration persistence.

We have now added this description in the Model section.

- Figure 2 itself needs revision, including a legends for Fig 2A and a clear statement in the legend/on the figure for what the red and blue circled regions are in the tSNE plots.

We have now added Supp. Fig. 3B where we show the close-up of a group of cells surrounded by ECM fibers and fluid pixels. Here we have depicted the different types of adhesive interactions. Also, in Figure 2 legend, we have now described what the red and blue regions represent.

- Again, the purpose of focusing on cell-cell adhesion in the computational modeling is unclear. “...these analyses implicate intermediate cell-cell adhesion in promoting invasion efficiency of cells in the presence of directional cues.” Likewise, throughout the manuscript the authors use the terms “clustering” in invasion/motility - a more clear explanation of the hypothesis and how the results support/do not support this framework is needed.

As mentioned above, incorporation of cell-cell adhesion as a parameter in our simulations enables us to assess the importance of phenotypic heterogeneity in different modes of invasion including single cell migration and collective migration. By cell cluster, we mean a group of cells within which cells are in contact with each other. We have now clarified this in the Results section.

- The authors propose that the small invasive cells are CSCs and state “Collectively, these results suggest that small sized cells of varying deformability are maximally invasive, and may correspond to a mix of mesenchymal and hybrid CSCs.” Based on the data provided, this statement is much too strong.

We have appropriately modified the sentence and also the title of the Result section.

- It is unclear how the statistics in Fig 5 were determined - “ $n > 190$ cells per condition pooled from $N = 3$ independent experiments.” How many cells were assessed in each experiment?

For quantification of cell size and stiffness of mCSCs and hCSCs, we have pooled measurements from 3 independent experiments. To obtain the distribution of cell size/stiffness, comparable number of cells were taken from each independent experiment. We have added a sentence in Methods section stating this.

- The gating strategy for sorting CD44^{hi}/CD24⁻ and CD44^{hi}/CD24⁺ cells should be included in supplement. Percent of the population in the provided “hCSC” and “mCSC” gates should be included. We have now added the gating strategy in Supp. Fig. 5A. We have also indicated the proportion of mCSCs and hCSCs in this figure.

- The spheroid assay is creative, but raises a number of questions. First, Fig 6 G figure legend incorrectly states “Temporal evolution of zone-wise distribution...” It is not clear that what is observed is evolution at all - and indeed, in such a short time frame, such evolution would be rapid indeed with limited generation time.

We acknowledge the reviewer’s concern. We have now replaced the word “evolution” with “rearrangement”.

- Because the spheroid assay does not utilize cell tracking, it is impossible to determine whether the observed CD44 intensity at the outer region of the spheroid is due to physical movement of the cells outward or due to transdifferentiation of the cells at the outer edge.

We agree that without single cell tracking experiments, it is impossible to ascertain the exact cause of increase in CD44 intensity at the outer region. To get some insight into this, we have now quantified inter-conversion between non-CSCs (nCSCs), mCSCs and hCSCs by isolating the three sub-populations, culturing them independently upto 48 hours and then checking the relative proportions in each of the independent cultures.

Our data suggests that both nCSCs and hCSCs give rise to mCSCs (Fig. 7). In comparison, mCSCs primarily exhibit symmetric division giving rise to mCSCs. The increase in CD44 intensity may thus be partly attributed to increase in the proportion of mCSCs.

We have now described this data in the manuscript and in the Discussion.

- The CD44 and CD24 expression of the cells in 2D was presented in Fig 5. Is such information available for the cells grown in a spheroid, either by flow cytometry or by IF? In Fig 5, the data shows a clear difference in size and stiffness between mCSC (CD24-) and hCSC (CD24+). To complete their model, therefore, such characterization of the “CD44 intensity” cells in Fig 6 spheroid studies is needed. While in the discussion the authors state “Since the proportion of mCSCs is significantly higher than that of hCSCs, it is likely that a majority” of the CD44+ cells in the spheroid are indeed mCSCs - this is a broad statement not supported by the provided data. We thank the reviewer for this suggestion. The proportion of mCSCs and hCSCs within spheroids was comparable to that in MDA-MB-231 cells grown in 2D cultures.

To assess the mechanism underlying enrichment of CD44^{hi} cells at the invasive front, mCSCs, hCSCs and non-CSCs (i.e., nCSCs) were isolated from parental cells, and expanded separately to assess their rates of interconversion. Our data suggests that both nCSCs and hCSCs convert to mCSCs. In comparison, mCSCs primarily exhibit symmetric division. Thus the enrichment of CD44^{hi} cells at the invasive front can be partly attributed to conversion of nCSCs (i.e., CD44^{low} cells) to mCSCs (i.e., CD44^{hi} cells).

We have now added these results in the newly added Figs. 7A, B and Supp. Fig. 6B, C. We have also discussed other papers which have made similar conclusions.

- The concluding in vitro experiments and the presented model fail to include the role of cell-cell adhesion. This is puzzling because of the importance it played in the computational modeling. I advise either removing it from the modeling or including additional rationale and data related to cell-cell adhesion in the final studies.

We have now performed FACS analysis in MCF-7 cells. These cells contain ~30% hCSCs, but no mCSCs. Subsequently, we performed spheroid invasion experiments using MCF-7 spheroids and tracked the spatial distribution of CD44^{hi} cells.

[NOTE: We have removed a figure which was provided to the reviewers in confidence]

Since the laser scanning confocal microscope at our institute that was used for imaging MDA-MB-231 spheroid invasion has been down for the past six months, we had to resort to regular epifluorescence imaging of MCF-7 spheroids. As a consequence, the cell boundaries were not clearly visible making it difficult to measure CD44 intensities of individual cells. Hence, line intensity of CD44 was tracked along three different directions (red, blue and green dotted lines) for each time-point. The line plots suggest there is no enrichment of CD44^{hi} cells at the invasive front.

Since this quantification may be error prone, and the method of analysis varies from the one presented in Fig. 6, we have chosen not to include this data in the manuscript. Instead, we have added a sentence in the fourth paragraph of the Discussion section where we state the need to perform these experiments with cells of varying extent of cell-cell adhesion.

Though we have not added this data to the manuscript, our modelling results make important predictions that are likely to be tested by experimentalists in the near future. Hence, we strongly feel that our simulations for different values of J_{cc} add value to our manuscript.

In our proposed model, we have not explicitly shown how the scenario will change depending on the extent of cell-cell adhesion, as this would make the schematic more complicated and difficult to follow. However, in the Figure legend, we have stated that the relative proportions of phenotypically distinct sub-populations at the invasion front is dependent on the extent of cell-cell adhesion.

- Throughout the discussion, I would urge the authors to soften their often too-bold statements.

For example, "...we have shown that cancer invasiveness is maximum when heterogeneity in both cell size and cell deformability are present." While this was computationally modeled, this was not shown in in vitro or in vivo studies.

Likewise, "Subsequent experiments identify these cells as CSCs." And later "we experimentally demonstrate to be CSCs." While the CD44+ cells they assess have some of the same properties, they do not definitively determine that the cells from the computational model are necessarily CSCs.

We thank the reviewer for this feedback. We have now gone through the Discussion section and appropriately modified sentences which seemed too strong.

Specific points for clarification in the materials and methods section:

- The specific protocol used for image analysis (e.g., plugins or workflow) should be included for all measures (cell size, change in cell spread over time, movement).

Image analysis pertaining to quantification of cell area as a function of time was performed in Image J by manually outlining each cell. Similarly, quantification of membrane-cytoplasmic E-cadherin ratio was performed using Image J.

We have now provided these details in the Methods section.

- Confirm that the MDA-MB-231 cells were lifted with trypsin prior to flow cytometry. This is an issue due to CD44 and CD24 antigen retention following trypsinization vs other cell disassociation methods.

It is usual practice to do mild trypsinization of adherent cells before flow cytometry (0.05% Trypsin-EDTA is used for 1-3 mins). This treatment does not appreciably alter antigen expression on cells.

We have followed the protocol based on the following references:

1. CD44+/CD24-breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis; doi. 10.1186/bcr1610
2. A "No-Touch" Antibody-Staining Method of Adherent Cells for High-Throughput Flow Cytometry in 384-Well Microplate Format for Cell-Based Drug Library Screening; doi: 10.1002/cyto.a.23956

- The specific antibiotic-antimycotic solution should be included.

We have now added the catalog number of the item in the Methods section.

- Flow cytometry analysis strategy should be detailed.

We have now added the gating strategy in Supp. Fig. 5A.

- Specific information about n per experiment and total number of experiments for in vitro would should be included.

This information about 'n' has been provided in Figure Legend of each Figure. On an average, equal number of measurements were considered from each independent experiment. We have now added a sentence in the Experimental Methods section stating this.

Response to Reviewer 2 Comments:

Combined heterogeneity in cell size and deformability promotes cancer invasiveness

Asadullah et al.

The authors present a needed examination of phenotypic heterogeneity and its role during invasion.

In particular the authors evaluate the influence of cell size and cell deformability in breast carcinoma invasion. The paper utilizes both experimental and computational approaches (extension of the Cellular Potts Model). Results suggest that combined heterogeneity in size and deformability enhance invasiveness. The authors further suggest that these characteristics are enriched in CSC populations and that CSCs are more robust leaders at the invasive front. There are many interesting advances and innovation in the manuscript. The framework is interesting and the use of the CPM/GGH framework certainly has value for the questions being asked. Enthusiasm, however, is somewhat tempered by a few key issues (outlined below). Yet, it is this reviewer's opinion that the manuscript could be acceptable for publication in JCS following key changes to the text and likely some additional data or analysis. The following comments are meant as constructive criticisms to help improve the manuscript.

We thank the reviewer for appreciating the value of our work and for providing constructive criticism for further strengthening our paper. We list below our responses to the concerns raised by the reviewer.

**Reviewer 2 Comments for the Author:
Specific Comments:**

1. There are issues with the Introduction and reference list. First the reference list is not formatted in a manner that allows it be evaluated. Following the JCS formatting would help and make the reference list accessible. Furthermore, the Introduction appears to be under-referenced and key primary literature on CSCs size and behavior is missing (e.g. Li et al., *Seminars in Cancer Biology*, 2015; De Paiva et al., *Stem Cell*, 2016; Ray et al., *Biophysical Journal* 2017 etc.). In this version, all references have been formatted as per JCS guidelines. We had cited Li. Et al. in the earlier version of the manuscript in the Discussion section. We apologize for missing out the other two references suggested by the reviewer.

We have now added two sentences in the Introduction citing the abovementioned references.

2. The decision to use 5 bins of size and deformability conditions appears reasonable, but a much better discussion of the rationale behind the cutoffs would strengthen the manuscript. Did the data in Figure inform these boundaries? Do sensitivity and/or statistical studies need to be performed to better define the cutoffs etc.?

The rationale behind the cut-offs is that we tried to study the population data close to average value, so we have taken one standard deviation from the average value both on the positive and the negative x-axis.

We have now described this in greater detail in the Results section.

3. The model description is quite incomplete. This reviewer had to study Reference 25 in some detail to interpret many aspects of the current work. Even in Ref 25 the model is not explained in sufficient detail. A much more in depth and robust description of the modeling framework and implementation is needed. It would also greatly strengthen the manuscript, particularly for the more general cell biology audience, to connect the model framework to particular cell behaviors or outcomes. For instance, how the framework handles size in terms of the energy formulation; AND for things like the statement "Together, these analyses implicate intermediate cell-cell adhesion in promoting invasion efficiency of cells in the presence of directional cues", Why? Please connect the model and outcomes - How does this connect to the model parameters and the known cell biology. As presented it is quite abstract.

Based on the reviewer's feedback, we have described our modelling framework in greater detail in Results Section 2. Based on the inputs of Reviewer 1, we have also described our choice of JCC and connected with experiments.

4. In line with comment 3, the statement "The last term in the energy expression was included to model chemotaxis of cells in the direction of the chemoattractant gradient (15), with (symbol) representing the effective chemical potential, and $v(\text{target})$ and $v(\text{source})$ representing the concentrations of chemoattractant at target and source pixel respectively. This term was used in the energy expression only for simulating cancer invasion in the presence of a chemotactic field." This seems like it needs to be parameterized / sensitivity studies over a range of values etc. The

behaviors are likely not linear so difference gradients could produce different outcomes.

We had initially performed simulations for $\lambda = \{200, 500, 1000, 2000, 5000\}$. While the maximum number of simulation steps was limited to 1800 MCS, simulations were stopped as soon as a cell reached the right boundary. For $\lambda = \{2000, 5000\}$, across all independent simulations for different values of J_{cc} , at least one cell reached the right boundary within 1800 MCS. In contrast, for $\lambda \leq 1000$, there were multiple simulations where no cell reached the right boundary within 1800 MCS.

Since our aim was to study repositioning of cells within the cluster, for comparing between independent simulations for a given value of J_{cc} , as well as comparing across simulations for different values of J_{cc} , it was important that the criteria of “one cell reaching the right boundary” was satisfied in all independent simulations. Only then, can we compare t^* , δ and spatial distribution of cells across the different cases. This is the reason, why we have reported our data for $\lambda = 5000$.

We have now added two sentences in the Methods section describing our parametric studies and that effective chemotaxis was observed for $\lambda \geq 2000$.

5. The data in Figure 1 is already, for the most part, widely available in the literature, and the findings quite expected. That said it is clear how this can be used to set the stage for the manuscript, but it requires better development and connectivity to the subsequent theoretical work. This may have additional relevance since MDA-MB-231 cells do not form any or strong cell-cell adhesion (they are a strong EMT phenotype at baseline). Furthermore, the deformability difference is very small. While statistically significant, is it biologically significant?

To better connect our experiments with our simulations, we have now performed E-cadherin staining in MCF-7 and MDA-MB-231 cells. It is difficult to comment on the biological significance of the difference in cell deformability. However, it is well known that MDA-MB-231 cells are highly metastatic. Since metastasis is a multi-step process, in addition to deformability, synergy with other factors may render greater advantage to a sub-population of cells. For example, softer cells which are also proteolytically more active are likely to possess higher metastatic potential.

6. How do the authors account for size and other phenotype behaviors? For instance, are the smaller cells just more amoeboid in phenotype and with different cell dynamics and physical modes of motility? What is steric and what is migration phenotype? This should be addressed as much as possible in the text and any possible analysis.

In our simulations, individual cells are assigned sizes from the size distribution in Fig. 2B. These sizes serve as “target size” in the energy formalism, i.e., whenever size of a given cell deviates from its target size, there is an energetic cost. This is accounted for while calculating the total system energy. Thus, sizes of individual cells fluctuate around their target sizes (Supp. Fig. 3C).

In our model, all cells possess identical MMP activity. For a given cell, steric hindrance is provided by neighbouring cells and the surrounding ECM fibers. The extent of ECM degradation is given by the MMP secretion rate multiplied by the number of pixels in contact between a given cell and an ECM fiber.

7. It is not clear that the model can really handle leader cell, collective invasive front, and single and collective follower dynamics as these processes have been shown to include quasi-ballistic behavior due to crowding/jamming. Some discussion or additional analysis in this area would strengthen the manuscript.

Since our model is a coarse grained model, it is unlikely to capture complex dynamics of collectively migrating clusters. Beyond reporting our observations, any additional analysis might be error prone. For this purpose, we have now removed the leader cell data from the manuscript.

More minor comments:

8. The Figures are very crowded and at times difficult to easily interpret. Figure 4 is particularly difficult to interpret.

We understand that in Figure 4, given the number of combinations, the data can be difficult to understand. For ease of following, we have expanded the Figure Legend describing what each point represents.

9. Translocation and distance are not adequately described.

These terms have now been clearly described in the Results as well as in Figure 2 Legend.

10. All 231 cell appear to be CD44 high (consistent with work by Massague), suggesting a high CSC baseline state for these cells. As such, is CD44 staining for leaders cells to be CSC robust enough? A bit more discussion here would be useful and also clear analysis of how CD44 levels are being parsed out in immunofluorescence images.

All MDA-MB-231 cells are CD44+. Both mCSCs and hCSCs as CD44^{hi} cells.

For quantification of CD44 intensity, images were acquired under identical gain and exposure settings for all conditions. Integrated CD44 intensity of individual cells within a given zone obtained by manually outlining the cells in Image J, were averaged to obtain the average integrated CD44 intensity of each zone. Subsequently, the average integrated intensities of individual zones were divided by the average integrated intensity of the innermost zone to determine the normalized CD44 intensity distribution.

We have now added this detail in the Methods section.

11. The PCA and t-SNE plots are interesting, but also need better connectivity presentation. Why not show the PCA data? Also, the red and blue designations need better description particularly in the figure text.

We have performed principal component analysis for directed migration.

[NOTE: We have removed a figure which was provided to the reviewers in confidence]

Since the above data is similar to that of tSNE plot, we have decided not to include the PCA data in the paper. In Figure 2 legend, we have now clearly mentioned what the red and blue regions represent.

Reviewer 3 Advance Summary and Potential Significance to Field:

The authors have examined the role of two axes of phenotypic heterogeneity - cell size and deformability - in driving cancer invasiveness. This integrative computational-experimental study is well-done, very timely and contributes to the growing literature on implications of non-genetic (i.e. phenotypic) heterogeneity in cancer progression.

We thank the reviewer for appreciating our paper.

Reviewer 3 Comments for the Author:

I request the authors to address the following minor queries/clarifications:

1. Fig 1, S1:

a) The authors have tracked individual cells over a period of 24 hour for quantifying cell size and shape changes. Was the cell cycle status considered to be a factor here? If not, the authors should mention this as a limitation of their analysis. Cell cycle heterogeneity can be a confounding factor (Spencer et al. Nature 2009), and cells are expected to change their biophysical traits during division. A recent study quantified the effects on EMT status as well during cell division (Tripathi et al. PLoS Comp Biol 2020).

We have not taken cell cycle status into consideration as cells were not synchronized in our experiments. We have added a sentence in the last paragraph on Discussion indicating the need to account for this factor.

b) Fig 1E: The authors should explain how “maximum change in cell size” is calculated. They can consider plotting some distributions of cell size changes.

Cell size fluctuation measurements were done by tracking cells for 12 hours. Maximum change in cell size was calculated with respect to initial cell size. We have now stated this explicitly in the Figure and the Figure legend.

c) Fig S1: Can the authors dissect the contribution of intra-tumor and inter-tumor variability here?

To strengthen their arguments later about connection of size/ deformability with EMT and/or stemness, they can consider calculating EMT and/or stemness scores/enrichment through single-sample GSEA for these cells.

We thank the reviewer for this suggestion. We have now re-analyzed the dataset using GSEA approach and have added tSNE-based 2D-projection of this dataset that allowed visualization of heterogeneity in Stemness/EMT and Biophysical property within samples from specific patient as well as in the entire dataset. Also, to strengthen the argument that biophysical properties of cells are important for Stemness and EMT potential, we curated a custom geneset BiophysicalGeneset that comprises of MYH9, MYH10, MYO10, MYO5B, MYO5C, RAC1, RAC2, ACTB, ACTG1, ACTN1, ACTN4, VIM, ROCK1, ROCK2, and CDC42, and calculate enrichment score for this geneset. Our analysis then showed a strong correlation between enrichment score for BiophysicalGeneset and that for EMT/Stemness. This is indeed a very interesting and important observation and we believe that inclusion of this quantification has improved our manuscript further.

We have added this Figure in Supp. Fig. S1 and described the results in Results Section.

d) Fig 1H: The authors mention that there is no correlation between cell size and motility. Can this argument be substantiated by relevant statistical test?

We have now provided the values of the correlation coefficients for both cell-types. For both the cell types, the value is close to zero.

e) The authors end this section with “breast cancer cell lines have heterogeneity at phenotypical as well as genetic level”. However, there is no genetic status (mutation etc.) analysis done; the authors should replace “genetic” with “gene expression”.

We acknowledge the reviewer’s concern. We have appropriately modified the sentence.

2. Fig 2:

a) 2F: Can the authors explain for what all parameters the t-SNE is performed? Also, because of its stochastic nature, t-SNE can give different clusters in each iteration. How do the authors gain confidence in the presented trends for t-SNE?

We have used following parameter values for performing tSNE:

Perplexity: 30 (default), Number of iterations: 1000, Number of PCA used: 2.

To obtain a more stable 2D tSNE embedding, position on tSNE space was computed 1000 times until algorithm stops further adjusting the cell positioning. We also performed analysis for different values of perplexity and obtained similar results.

b) 2D: The authors suggest that combined heterogeneity in deformability and cell size led to increased average net translation, but heterogeneity in only one axis reduced invasiveness. Can the authors offer a biophysical explanation for it? Moreover, relevant statistical test results are needed for a stronger argument.

Combined heterogeneity allows greater extent of rearrangement within the cell cluster, which is lesser when there is heterogeneity only in one axis. We have now performed statistical tests of significance for Fig. 2D and Supp. Fig. 3E.

3. Fig 3: a)The authors should include the quantitative details of the gradient used.

For our study, as shown in Fig. 3A, the chemokine concentration is 0 units at the left of the simulation lattice and 1 unit at the right of the simulation lattice. The gradient is linear. We have indicated this in the Results section.

b)The authors mention that chemotaxis is fastest at intermediate cell-cell adhesion. Can they attempt to discuss their results vis-à-vis observations that intermediate cell-cell adhesion can promote migration of cells as clusters (Bocci et al. Cancer Res 2019) and that clusters can be better at chemotaxis as compared to individual cells (Camley et al. PRL 2016).

We thank the reviewer for this suggestion. We have discussed these papers in the Discussion section.

4. Fig 5:

a) The authors have identified two different kinds of CSCs (CD44+ CD24+ - hCSCs, and CD44+ CD24-

mCSCs) and suggest that mCSCs may be the one enriched at invasive edge. A recent computational modeling effort had similar conclusions (Bocci et al. PNAS 2019); the authors should discuss similarities and differences with this study. Moreover, other relevant studies on CD44+ CD24+ in breast cancer cells should be cited - Grosse-Wilde et al. PLoS ONE 2015, Goldman et al. Nat Comm 2014, Boareto et al. J R Soc Interface 2016, given the predominance of CD44+ CD24- as the established CSC signature in breast cancer.

We thank the reviewer for this suggestion. We have now discussed and/or cited these papers in the manuscript.

b) Experimentally, did the authors observe interconversion among these two subsets of CSCs, as seen in Liu et al. Stem Cell Reports 2014?

We have now performed this experiment. Our data suggests that both non-CSCs and hybrid CSCs both are capable of converting to mCSCs. In comparison, mCSCs primarily exhibit symmetric division. We have added this data in Fig. 7.

Second decision letter

MS ID#: JOCES/2020/250225

MS TITLE: Combined heterogeneity in cell size and deformability promotes cancer invasiveness

AUTHORS: Asadullah, Sandeep Kumar, Neha Saxena, Madhurima Sarkar, Amlan Barai, and Shamik Sen

ARTICLE TYPE: Research Article

I am happy to tell you that your manuscript has been accepted for publication in Journal of Cell Science, pending standard ethics checks.

Reviewer 1

Advance summary and potential significance to field

The authors have addressed all of my concerns. No further revisions are requested.

Comments for the author

The authors have addressed all of my concerns. No further revisions are requested.

Reviewer 2

Advance summary and potential significance to field

The authors have a done a satisfactory job of addressing the concerns of each of the reviewers. This reviewer has no additional comments.

Comments for the author

The authors have a done a satisfactory job of addressing the concerns of each of the reviewers. This reviewer has no additional comments.

Reviewer 3

Advance summary and potential significance to field

The authors have examined the role of two axes of phenotypic heterogeneity - cell size and deformability - in driving cancer invasiveness. This integrative computational-experimental study is well-done, very timely and contributes to the growing literature on implications of non-genetic (i.e. phenotypic) heterogeneity in cancer progression.

Comments for the author

The authors have now addressed the questions I had in the first round of review.