THE
EMBO
JOURNAL

3D model for CAR-mediated cytotoxicity using patient-derived colorectal cancer organoids

Theresa E. Schnalzger, Marnix H.P. de Groot, Congcong Zhang, Mohammed H. Mosa, Birgitta E. Michels, Jasmin Röder, Tahmineh Darvishi, Winfried S. Wels and Henner F. Farin

Review timeline:

Submission date:	16th Oct 2018
Editorial Correspondence:	20th Nov 2018
Authors' correspondence	30th Nov 2018
Editorial Decision:	6th Dec 2018
Revision received:	17th Feb 2019
Editorial Decision:	12th Mar 2019
Revision received:	24th Mar 2019
Editorial Decision:	25th Mar 2019
Revision received:	26th Mar 2019
Accepted:	27th Mar 2019

Editor: Daniel Klimmeck

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 20th Nov 2018

Thank you for the submission of your manuscript (EMBOJ-2018-100928T) to The EMBO Journal. Your manuscript has been sent to three referees, and we have received reports from all of them, which I enclose below.

As you will see, the referees acknowledge the potential high interest and novelty of your work, although they also express a number of issues that will have to be addressed before they can support publication of your manuscript in The EMBO Journal. Referee #3 states that the in vivo function of FRIZZLED -CAR cells is not sufficiently supported by the data and requests additional analyses to corroborate this point. Referee #2 agrees in that the FRIZZLED CAR cell experiments should be expanded and further is concerned that the effects of nicotinamide are not conclusively addressed. In addition, the referees point to issues related to better discussion of the results, missing controls and methods annotation that would need to be conclusively addressed to achieve the level of robustness needed for The EMBO Journal.

I judge the comments of the referees to be generally reasonable and given their overall interest, we are in principle happy to invite you to revise your manuscript experimentally to address the referees' comments.
I do agree that demonstrating in vivo function of the FRZD CAR cells would significantly strengthen the study.

Thus I would ask you to let me know about your view on the revisional work shortly.

REFEREE REPORTS

Referee #1:

This paper describes a method to study the activity of CAR-NK cells against colorectal cancer organoids using engineered NK-92 cells. Having first identified nicotinamide in the medium as inhibitory to cytotoxic activity of NK-92 cells they find that CAR-NK cells cannot penetrate the Matrigel in which organoids are usually grown but they will kill organoids that are place on a thin layer of Matrigel. To measure cytotoxicity they introduced luciferase/GFP into the organoids allowing measurement of cytotoxicity in these multi-cellular organoids.
This model system allowed them to test effector:target cell ratios, the effect of a layer of fibroblasts and to study cell killing in real time, they established a confocal live cell imaging protocol well-illustrated in the accompanying movies with image analysis to measure loss of fluorescence in the organoids. The assay was also able to demonstrate potential off-target activity of one neo-antigen that was a potential target.
Overall this work is well-performed and this model system could have some utility as a pre-clinical screen for engineered NK and T cells.

1. The movies are impressive and illustrate the main findings of the paper very well. It would help the reader if they were referred to individually at the appropriate point in the text rather than just as S1-6 and S7-9.
2. While this paper convincingly shows that multi-cellular structures can be killed by CAR-NK cells and the organoids can be classified as 3D this model does not fully replicate a 3D tumor microenvironment even if the organoids are grown on a fibroblast monolayer. The experiments described here also illustrate that neither the NK-92 cells or the CAR-NK cells can penetrate the Matrigel that normally would enclose the organoids. Presumably would also be unable to penetrate the ECM that will surround malignant cell islands of human colorectal tumors. This finding should addressed in the discussion. Would this be the same for CAR-T cells? Is the absence of NK cells in many tumors at least in part due to their inability to move along ECM matrices.

Referee #2:

Off-tumor toxicity of CAR-cells is a significant problem requiring a personalized approach. The authors report an elegant image-based approach for screening efficacy and safety of CAR-cell based therapies. A strength of the platform is that it is based on 3D organoid culture, yet in a format that essentially corresponds to a 2D system. This makes the model potentially widely accessible for labs with basic cell culture equipment and access to confocal or spinning disc microscopy. The approach should be useful to screen other targets, tumors types, or CAR-mediated strategies.
Overall, the study is systematically designed and experimentally sound. The manuscript is well written and of interest to the readership of the EMBO Journal. The following points should be addressed to render the manuscript suitable for publication:

• Nicotinamide was identified as a non-essential factor for organoid culture. It seems that this conclusion was made based on bright-field microscopy and a qualitative assessment of growth (based on size and morphological hallmarks)? It would be important to conduct a more in-depth analysis of organoid culture in the presence and absence of nicotinamide (effect on maintenance, cell type composition, etc.) to substantiate this conclusion.
• Figure 1G shows a size reduction of organoids in both Matrigel and suspension. If this is linked to a paracrine effect of the NK, it would be good to show proper controls in which no NK are present, and without NK but conditioned medium from NK-tumor co-cultures. Moreover, how do the authors explain that the specific lysis only occurs on Matrigel, and not in suspension? A proper explanation is required as well as showing related controls with only organoids on Matrigel or in suspension to assess organoid death without NK. Finally, what is the effector to target ratio for the experiments shown in this figure?
• Figure 2B: I suggest to show the percentage of organoid death alone, without NK, overtime, or show normalized data based on that. The E:T ratio should also be mentioned.
• In general, image analysis based on mean fluorescence of GFP was used to assess cell death. This read-out may also depend on morphological changes of organoids, the presence of dead or non-

fluorescent cells, or photo-bleaching. It would be good to also assess organoid lysis by staining for live and dead cells to validate the GFP-based measurements.
• Figure 3B: What is the E:T ratio?
• CAR targeting of tumor neoantigens: What cell type was used for engineering the organoids here? Was this from the same patient than in the first experiments? Tumor derived or healthy tissue?
• Figure 5B: Why are NK cells now GFP positive? This would affect the image analysis as it is based on green fluorescence and so any NK overlapping the organoids or in the 50um range influences the overall green intensity.
• The evaluation of CAR targeting Frizzled receptors was based on mouse organoids. In my view, the impact of the work would be strengthened substantially if the authors could conduct these experiments with patient-derived samples.
• Please indicate if the patient samples are fresh or frozen.

Referee #3:

The authors have developed a novel experimental system that allows to assess the cytotoxic effect of chimeric antigen receptor (CAR)-engineered NK-92 cells on colon epithelial organoids. They first identified a co-culture condition where CAR cells target organoid cells on a matrigel-coated culture plate. They then established a quantitative assay for CAR-mediated cytotoxicity toward colon organoids and also a live imaging system for visualization of dynamic behaviors of CAR cells. The well-designed methodology demonstrated in this study may have potential to become a novel platform to assess the interaction between colonic organoids and CAR-engineered NK-92 cells under particular circumstances. However, the data presentation, especially those on the FRIZZLED CAR cells, is insufficient and this makes it hard to understand the significance of the system as a tool to evaluate efficacy and safety of CAR therapy against colorectal carcinoma in vivo.

Major

The authors have generated a novel line of CAR NK-92 cells so that they bind to and react with frizzled proteins, based on the sequence data of the previously described pan-FRIZZLED monoclonal antibody OMP-18R5 (Gurney et al. PNAS 2012). The FRIZZLED CAR cells showed cytotoxic effects on FZD5-expressing MDA-MB453 cells but not on parental cells, demonstrating that the engineered NK-92 cells are capable of killing organoids involving CAR-dependent mechanisms. By using this line of CAR cells, the authors showed that there was no significant difference in the CAR-mediated cell lysis between wild-type, APC-KO and Rnf43/Znrf3-DKO organoids.
Based on this observation, the authors claimed that the strategy to target Frizzled proteins by CAR systems may have potential risk of cytotoxicity against normal colonic tissue. Furthermore, it appears to me that, on the grounds of these data, the authors claim that the co-culture system presented in this study may serve as a suitable preclinical assay to foresee the risk of therapy. However, this is not a persuasive argument unless the data on in vivo effects of the FRIZZLED CAR cells are presented.
The pan-FRIZZLED antibody OMP-18R5 was previously shown to inhibit the growth of a variety of tumor cells including colon tumors in xenograft models (Gurney et al. PNAS 2012). Gurney et al. showed in this previous paper that administration of OMP-18R5 did not show any apparent side effects but resulted in colitis, presumably due to suppression of the Wnt pathway, when it was used at very large doses. The present study should test the in vivo effect of FRIZZLED CAR cells. Do these CAR cells have anti-tumor activity together without any lytic effects on Frizzled-expressing normal tissues? Or do they induce severe colitis as in the previously reported xenograft models treated with high-dose OMP-18R5? If we see a safe therapeutic window in which colon tumor xenografts can be targeted without any other tissue damage, this may suggest that the co-culture model presented in this study is too sensitive to predict the safety of CAR systems appropriately.
The authors should also perform flow cytometry or some quantitative assays to show the expression levels of Frizzled proteins on wild-type, APC-KO and Rnf43/Znrf3-DKO organoids cultured in the presence and absence of Rspo1. This will help readers understand how DKO of Rnf43/Znrf3 or Rspo1 treatment induces the change of Frizzled expression relative to the basal expression levels.

Minor

The authors described that EGFRviii, which was introduced by using the lentiviral system, was undetectable by immunoblot assays due to low expression levels (Fig. 4C). Are the organoids analyzed in Fig. 4A-E the same as those assayed in Fig. 4F-I, which co-expressed detectable levels of EGFP or RFP? If so, what do the authors think is the reason for different levels of expression between EGFRviii and fluorescent proteins? If not, the immunoblot data of the EGFRviii expression in organoids doubly positive for EGFRviii and EGFP should be presented. The authors will be able to try immunohistochemistry to detect truncated EGFRviii on sections of fixed organoids.

Authors' correspondence 30th Nov 2018

On behalf of all co-authors, I want thank for the constructive comments and the opportunity to prepare a revised version of our manuscript. I am very happy that the 3 reviewers have acknowledged our study as 'impressive', 'well-designed' and 'potentially widely accessible' to assess CAR-mediated responses towards patient derived organoids. As per your request, I would like to communicate upfront our experimental plans and argumentation:

The reviewers' comments converge on 4 points: A) the culture setup, B) the live imaging read-out, C) the characterization of the organoids and D) the validity of our platform with respect to safety assessment of the Frizzled-CAR line. We think that we can adequately address all these points with the following new experiments:

A1) Document that nicotinamide can be safely withdrawn from the organoid medium to improve killing without interference with organoid growth (reviewer 2).
A2) Quantify the impact of the culture setting (in matrigel vs. on matrigel vs. in suspension) on the status of the organoids in the absence and presence of CAR NK cells (reviewer 2).
B1) Determine a linear range of the image-based quantification of organoid size, as well as robustness with respect to factors such as organoid morphology, exclusion of dead cells and photo-bleaching (reviewer 2).
B2) Demonstrate that GFP expression in NK-cells can be effectively distinguished by our organoid area quantification algorism using intensity and size filters (reviewer 2).
C1) Perform immunoblot analysis of EGFRvIII expression (reviewer 3) at increased sensitivity using new antibodies and/or immunoprecipitation.
C2) With respect to the detection of Frizzled protein levels (reviewer 3) we have to argue that endogenous protein levels have so far not been successfully studied in gastrointestinal cells. Immunofluorescence detection using the pan-Frizzled antibody (OMP-18R5) has not been successful in our hands, most likely due to the low protein levels. FACS experiments are unrealistic because Frizzled receptors are sensitive to protease treatment required for single cell dissociation. However, there is strong genetic and functional evidence that the conditions studied (Fig. 6) represent distinct surface levels (Koo et a., Nature 2012 and Hao et al., Nature 2012) and we have indirectly visualized Frizzled levels previously in the same organoid genotypes using a knock-in allele of a tagged Wnt-ligand (Farin et al., Nature 2014). Moreover, we have validated the specificity of our CAR cells using an overexpression approach (Fig. S5B).
D1) Investigate the potential toxicity of Frizzled-CAR cells using a panel of human organoids from colon cancer and normal colon (Reviewer 2). We have tumor-organoids to our disposal that contain somatic RNF43 mutations, which justifies Frizzled-CAR NK-cells as a therapeutic rationale. These experiments will increase the translational impact of our study, and further support our mouse organoid models that however are genetically more defined.
D2) Perform dose titration experiments with Frizzled-CAR cells towards WT and RZ-KO organoids to investigate a therapeutic window at low CAR doses to address if our system is "is too sensitive to predict the safety of CAR systems appropriately" (Reviewer 3).

Given that we have so far not been able to detect any therapeutic advantage of Frizzled-CAR NK-cells in vitro, we think that our system can serve as a very informative model for preclinical target validation. Of course, we agree that our model cannot substitute toxicity testing in vivo because CAR-toxicity may include off-cancer activity towards multiple organs and may be strongly modulated by the immune system ("cytokine release syndrome"). The OMP-18R5 antibody (that our

CAR is based on) has shown some efficacy against xenografted cancer cells and has shown toxicity only at high doses. However, we want to stress that these preclinical data were not the motivation to choose Frizzled as a target. Rather, our approach is based on the rational strategy that RNF43 mutant tumors exhibit increased Frizzled surface levels based on the literature (knock-down studies in cell lines, and genetic evidence in mice, see above). Our platform has allowed to test this hypothesis by targeting the endogenous protein in primary cells. In this context the presented data clearly shows toxicity towards normal cells arguing against a therapeutic benefit and potential risks in vivo. In our revised manuscript we will formulate our findings more carefully and mention that Frizzled-CAR cells in some clinical settings may have a therapeutic benefit (e.g. in a liver metastasis model, where gastrointestinal off-tumor toxicity can be avoided by local application).

Our Frizzled-CAR NK-cells would indeed represent an interesting model for future in vivo toxicity studies, due to the cross-species reactivity of the CAR-NK line. However, such experiments will require careful dose escalation, followed by serological and histological analysis, in addition to the establishment of appropriate tumor models. Given the complexity and the time to obtain an official permit, such experiments would certainly justify an independent study.

We think that our additional experiments and our arguments above will hopefully also convince reviewer 3 and we would be very happy if the editor would support our strategy. Given that reviewers 1 and 2 were very supportive and did not explicitly demand for in vivo data argues that our data represent a relevant and complete study by itself. I would be happy to discuss any open questions with you by phone.

1st Editorial Decision 6th Dec 2018

Thank you for following-up on our decision on your manuscript (EMBOJ-2018-100928T) and providing us with a preliminary revision plan. We have now discussed your detailed experiments in the editorial team, and in addition we have also asked referee #3 to consider them. Concluding from those considerations, we encourage you to proceed with your revisional work along the lines of your proposal.

We realise that you would - judging from the information provided in the point-by-point letter - be potentially able to address most issues raised by the referees in a revised version of the manuscript. While we also noted that referee #3 remains unconvinced that his/her critique would be fully addressed by your proposed experiments, we concur with your point that the amount of work needed to address the in vivo relevance of the CAR-Frizzled in new tumorigenesis settings would likely go beyond the scope of the current study. We however ask you to carefully revise your manuscript to relativise your statements and introduce caveats where appropriate.

REFEREE REPORTS:

Referee #3's additional comments

I have read through the authors' revision plan. I understand that the plan includes experiments that would allow them to address most of the reviewers' requests. In response to my request, however, they are only planning to perform dose titration experiments with Frizzled-CAR cells towards WT and Rnf43/Znrf3-DKO organoids to investigate a therapeutic window at low CAR doses in vitro. They state that they would be able to address if the system is too sensitive to predict the safety of CAR systems appropriately. However, it seems to me that they are not responding to my comment directly. What I would like to say is that they should perform in vivo experiments if they want to claim the system to be suitable as a preclinical target validation system. In other words, if they really insist on the utility of the system as a safety/toxicity prediction tool, they should show that the results obtained in in vitro assay systems indeed correspond to what we observe in in vivo experiments. If the authors are not able to provide in vivo data regarding the effect of FRIZZLED CAR cells they should tone down the significance of the system as a preclinical assay.

Point-by-point response (answers in blue)

Referee #1:

This paper describes a method to study the activity of CAR-NK cells against colorectal cancer organoids using engineered NK-92 cells. Having first identified nicotinamide in the medium as inhibitory to cytotoxic activity of NK-92 cells they find that CAR-NK cells cannot penetrate the Matrigel in which organoids are usually grown but they will kill organoids that are place on a thin layer of Matrigel. To measure cytotoxicity they introduced luciferase/GFP into the organoids allowing measurement of cytotoxicity in these multi-cellular organoids.
This model system allowed them to test effector:target cell ratios, the effect of a layer of fibroblasts and to study cell killing in real time, they established a confocal live cell imaging protocol well-illustrated in the accompanying movies with image analysis to measure loss of fluorescence in the organoids. The assay was also able to demonstrate potential off-target activity of one neo-antigen that was a potential target.
Overall this work is well-performed and this model system could have some utility as a pre-clinical screen for engineered NK and T cells.

We thank the reviewer for this appraisal of our work.

1. The movies are impressive and illustrate the main findings of the paper very well. It would help the reader if they were referred to individually at the appropriate point in the text rather than just as S1-6 and S7-9.

The movies are now referred to at the individual text passages.

2. While this paper convincingly shows that multi-cellular structures can be killed by CAR-NK cells and the organoids can be classified as 3D this model does not fully replicate a 3D tumor microenvironment even if the organoids are grown on a fibroblast monolayer. The experiments described here also illustrate that neither the NK-92 cells or the CAR-NK cells can penetrate the Matrigel that normally would enclose the organoids. Presumably would also be unable to penetrate the ECM that will surround malignant cell islands of human colorectal tumors. This finding should addressed in the discussion. Would this be the same for CAR-T cells? Is the absence of NK cells in many tumors at least in part due to their inability to move along ECM matrices.

We are thankful for these thoughts. As shown in Fig. 1G and Fig. EV2D the presence of a 2D layer of ECM is critical for cytotoxicity, most likely because it supports migration of CAR cells or stabilizes the interaction with the organoids, which we now mention in the results. Embedding of target cells in 3D Matrigel efficiently blocks cytotoxicity and here the CAR cells do not efficiently penetrate the dense matrix. In our discussion we mention that insufficient cell infiltration has been recognized as a key problem for CAR therapy in solid cancers. We suggest that: “The limited penetration of NK-cells into 3D Matrigel may serve as a model to study and improve immune cell infiltration into tissues.”

Referee #2:

Off-tumor toxicity of CAR-cells is a significant problem requiring a personalized approach. The authors report an elegant image-based approach for screening efficacy and safety of CAR-cell based therapies. A strength of the platform is that it is based on 3D organoid culture, yet in a format that essentially corresponds to a 2D system. This makes the model potentially widely accessible for labs with basic cell culture equipment and access to confocal or spinning disc microscopy. The approach should be useful to screen other targets, tumor types, or CAR-mediated strategies.
Overall, the study is systematically designed and experimentally sound. The manuscript is well written and of interest to the readership of the EMBO Journal. The following points should be addressed to render the manuscript suitable for publication:

We thank the reviewer for empathizing the applicability of our experimental model and the quality of our data.

• Nicotinamide was identified as a non-essential factor for organoid culture. It seems that this conclusion was made based on bright-field microscopy and a qualitative assessment of growth (based on size and morphological hallmarks)? It would be important to conduct a more in-depth analysis of organoid culture in the presence and absence of nicotinamide (effect on maintenance, cell type composition, etc.) to substantiate this conclusion.

We thank the reviewer for making us aware on this point. We did not intend to state that nicotinamide is dispensable for the long-term culture of organoids. Rather, we have confirmed that withdrawal of nicotinamide only for a short co-culture period (usually 8-10 hours) does not affect cell vitality. We have now quantified that nicotinamide withdrawal for 3 days does not affect ATP levels (Fig. 1D). Concerning an effect on cell-type specific differentiation, we want to mention that culture of human organoids requires a complex medium, which induces a general progenitor-like status (Sato et al., 2011). Unaffected viability after 3 days therefore indicates that the cell status is at least for the time of our killing assays not affected by the withdrawal.

• Figure 1G shows a size reduction of organoids in both Matrigel and suspension. If this is linked to a paracrine effect of the NK, it would be good to show proper controls in which no NK are present, and without NK but conditioned medium from NK-tumor co-cultures. Moreover, how do the authors explain that the specific lysis only occurs on Matrigel, and not in suspension? A proper explanation is required as well as showing related controls with only organoids on Matrigel or in suspension to assess organoid death without NK. Finally, what is the effector to target ratio for the experiments shown in this figure?

The experiment in Fig. 1G was performed to determine the optimal settings for CAR-directed cytotoxicity (E:T ratio was 4:1, we have added this information). Only after seeding on a matrigel layer cytotoxicity is observed as seen by abundant formation of apoptotic bodies. We now show higher magnification images to better illustrate the organoid morphologies and discuss that "Our results suggest that co-culture on an ECM layer increase NK-cell migration and/or stabilize the effector-target cell interaction and was therefore used for all subsequent experiments." In addition, we have quantified the cytotoxicity in the different co-culture conditions by our luciferase assay (shown in Fig. EV2D). In suspension culture, we have found unaffected viability of organoids alone and after co-culture with EPCAM-CAR NK-92 cells. Furthermore, we could show that EPCAM-CAR NK-92 cells do not significantly change the viability after seeding in matrigel. The reduced organoid size after co-culture in this condition could therefore indeed be due to paracrine interactions as suggested. However, because we have not observed size reduction on 'bystander cells' in our competitive killing experiments (Fig. 5D) we are confident that any potential paracrine effect should not strongly impact our image-based read-out (see also two points below).

• Figure 2B: I suggest to show the percentage of organoid death alone, without NK, overtime, or show normalized data based on that. The E:T ratio should also be mentioned.

We thank the reviewer for pointing this out. The E:T ratio was marked in the legend, but we now better explain how 'target cell lysis' is defined. We have now mentioned more clearly in text and methods that our data is always shown as remaining luciferase activity compared to organoids that were cultured in parallel without NK cells.

• In general, image analysis based on mean fluorescence of GFP was used to assess cell death. This read-out may also depend on morphological changes of organoids, the presence of dead or non-fluorescent cells, or photo-bleaching. It would be good to also assess organoid lysis by staining for live and dead cells to validate the GFP-based measurements.

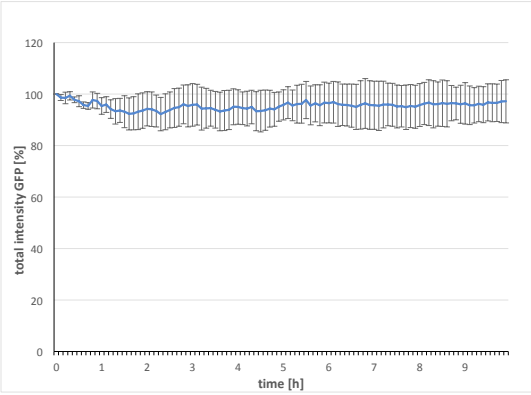
We thank the reviewer for giving us the opportunity to comment on these important points. First, we would like to emphasize that in side-by-side measurements, the results from imaging-based quantification show a very good overlap to the luciferase-based read-out (Fig. 3E), which supports the validity of our approach.

As apparent in the movie S6 the presence of dead cells does not greatly impact on the automatic detection of organoid area (see thin lines). The workflow for detection (see Appendix Fig. 2) contains a size filter (steps 3 and 4), which allows effective exclusion of these small particles. Although some apoptotic bodies may be temporally included into the area if they are very close to the organoid, these structures represent a small percentage of the total area. In addition, movie S6 shows that the cell debris is highly motile and moves away from the organoid and/or disintegrates. This transient nature is actually the key problem to use apoptotic cells as read-out for cytotoxicity (see below), because prolonged monitoring is required to follow killing of the macroscopic structures. The fact that we can readily track the complete loss of organoids (see. Fig. EV3D) further confirms that residual GFP-positive cell debris does not cause false positive signals.

We have now confirmed ubiquitous GFP expression in our organoids by FACS analysis (see Fig. EV2B). Reporter expression can be assured by co-selection of an antibiotic3 marker (Luc2-P2A-GFP-IRES-Puromycin lentivirus). In addition, our workflow quantifies the relative GFP-area compared to the starting area, and is therefore robust against non-homogeneous expression and even insensitive to the presence of fluorescence reporter-negative cell populations (see Fig. 5CD).

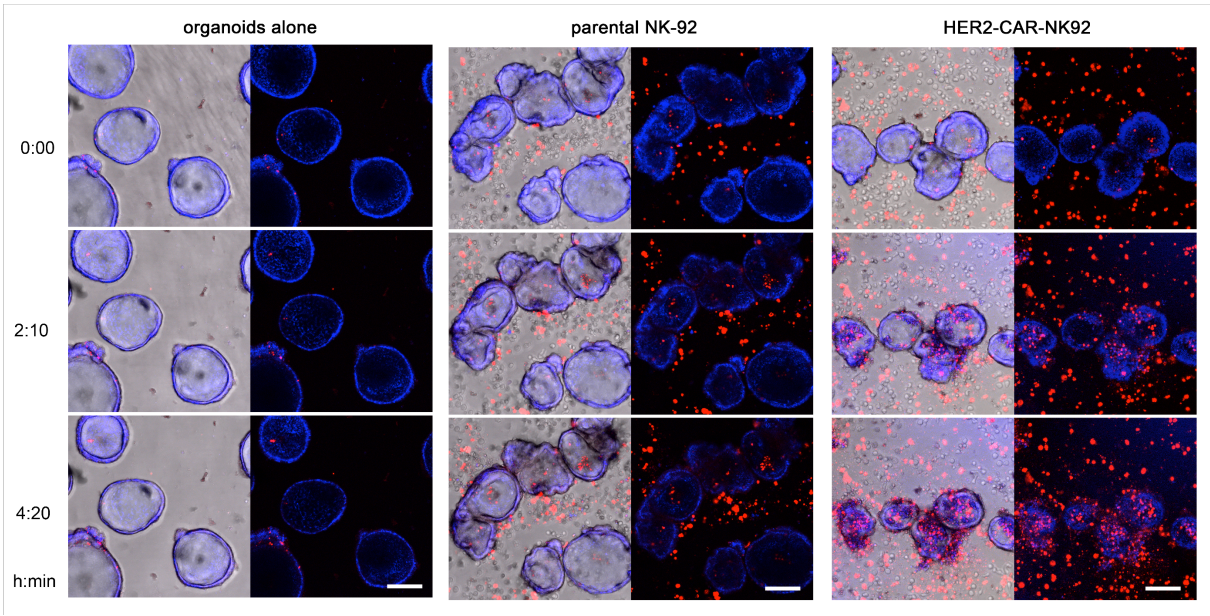
We agree that morphological changes of the organoids could potentially affect the read-out. Organoids are hollow structures and the size may vary after permeabilization of the epithelial layer. While such fluctuation may be prominent on a single organoid level (see Fig. EV3E) we perform analysis of multiple organoids in parallel. Our data shows a progressive and linear reduction of the relative area (see Fig. 3D, Fig. 5C and Fig EV5C), arguing for robustness against fluctuations in individual organoids. In order to further validate the detection we have titrated the organoids and have measured the area before and 8 hours after addition of EPCAM-CAR cells (Fig. EV3B). The data shows that linear quantification of organoid area is possible before and after co-culture.

As shown in movie S1 no photo bleaching is observed in our live imaging (movies show non-processed data). In addition, we have plotted the total GFP intensity during imaging of organoids without NK-cells, which confirmed a stable GFP signal (see Review Fig. 1).



Review Figure 1. Analysis of total GFP intensities during live imaging
Continuous imaging of GFP-expressing normal organoids in the absence of NK-92 cells for 10 hours (data from Fig. 3D). Mean total GFP intensity in n = 4 imaging fields (± SD) is shown relative to t = 0. Intensities were automatically retrieved form the analysis of maximal image projection data. Images were acquired in 6 min intervals. No photo-bleaching is observed during live imaging.

While setting up our image-based assays, we have initially also tested the possibility to quantify cell death by a positive read-out. However, as mentioned above and in our discussion the time for efficient cytotoxicity against organoids exceeds the detection time of apoptotic cells. In addition, apoptotic bodies are highly motile, which complicates the assignment of signal to a specific organoid. Moreover, NK-92 cells also show prominent apoptosis, which further complicates the detection of dead epithelial cells (see Review Fig. 2). We therefore believe that the quantification of remaining live cells is more useful for assessment of cytotoxicity.



Review Figure 2. Detection of dead cells during live imaging of CAR-NK-92 mediated organoid killing
Confocal life imaging (maximal image projections) of human tumor organoids alone or together with parental NK-92 cells or HER2-CAR NK-92 cells (E:T ratio 2:1). Staining for dead cells using propidium iodide that was added to the medium (red channel) and epithelial cells that were pre-stained with calcein violet (blue channel). Overlay with bright field images (left), and fluorescent channels alone (right) are shown. Note the presence of dead NK-cells at t = 0 and the progressive dispersal of apoptotic bodies away from the organoids. Scale bars are 100 µm.

• Figure 3B: What is the E:T ratio?

Thanks, we have added the information to the figure legend (E:T ratio in Fig. 3 was 2:1).

• CAR targeting of tumor neoantigens: What cell type was used for engineering the organoids here? Was this from the same patient than in the first experiments? Tumor derived or healthy tissue?

As shown in Fig. EV4B-D we have introduced the EGFRvIII in both normal and tumor-derived organoids. The normal line was identical to the line used in above (Figs. 2/3) and the tumor organoid line was from an unrelated individual. For killing experiments in Fig. 4AB we have used normal organoids +/- EGFRvIII. For the all other CAR-EGFRvIII experiments normal organoids (N) and tumor organoids +($T_{EGFRvIII}$) or -(T_{ctrl}) EGFRvIII were used as correctly indicated in the legends. We have now added more clear labels to the figures.

• Figure 5B: Why are NK cells now GFP positive? This would affect the image analysis as it is based on green fluorescence and so any NK overlapping the organoids or in the 50um range influences the overall green intensity.

It is correct that the CAR-EGFRvIII cells are GFP positive due to strong CAR expression together with GFP in this NK-92 pool (mentioned in the methods section). We could observe effective CAR-directed lysis by image-based monitoring (Fig. 5C), arguing that the GFP expression does not perturb the detection of organoid area loss. We would like to clarify that we have measured the fluorescent area and not the overall GFP intensity. This is why NK cells that have migrated onto the organoids will have no major effect. Only NK cells that locate on the edge of organoids could influence the detected area. However, as mentioned above our analysis workflow allows effective filtering of small particles to avoid that adjacent NK-cells are added to the GFP+ area. We have now performed control experiments to demonstrate the performance of our algorithm (Fig. EV5). We show that in GFP/DsRED double positive organoids that the workflow allows similar quantification

of organoid area by detection of GFP or DsRED. Here, the GFP-expression in EGFRvIII-CAR NK-92 cells did not interfere with the result, validating our read-out.

• The evaluation of CAR targeting Frizzled receptors was based on mouse organoids. In my view, the impact of the work would be strengthened substantially if the authors could conduct these experiments with patient-derived samples.

We fully agree that experiments in human CRC organoids can increase the translational impact of our study. Our isogenic mouse organoids (shown in Fig 6), however, are genetically more defined. In order to identify an RNF43-deficient human CRC organoid we have now tested a panel of human organoids for IWP-2 sensitivity (which allows to identify RNF43 deficiency), followed by genotypic confirmation. A panel of 4 CRC and one normal organoids was then tested against parental NK-92, Frizzled-CAR and EPCAM-CAR cells, which showed that neither of the 2 CAR lines can be used to selectively target (RNF43-deficient) CRC organoids (Fig. EV6). These results further support our observations made in mouse organoids and illustrate the potential of our platform for personalized testing of CAR-efficiency and specificity.

• Please indicate if the patient samples are fresh or frozen.

All organoid lines were derived from fresh tissue following the standard culture protocol (Sato et al., Gastroenterology) as mention in the methods.

Referee #3:

The authors have developed a novel experimental system that allows to assess the cytotoxic effect of chimeric antigen receptor (CAR)-engineered NK-92 cells on colon epithelial organoids. They first identified a co-culture condition where CAR cells target organoid cells on a matrigel-coated culture plate. They then established a quantitative assay for CAR-mediated cytotoxicity toward colon organoids and also a live imaging system for visualization of dynamic behaviors of CAR cells. The well-designed methodology demonstrated in this study may have potential to become a novel platform to assess the interaction between colonic organoids and CAR-engineered NK-92 cells under particular circumstances. However, the data presentation, especially those on the FRIZZLED CAR cells, is insufficient and this makes it hard to understand the significance of the system as a tool to evaluate efficacy and safety of CAR therapy against colorectal carcinoma in vivo.

We thank the reviewer for this appraisal of our work.

Major

The authors have generated a novel line of CAR NK-92 cells so that they bind to and react with frizzled proteins, based on the sequence data of the previously described pan-FRIZZLED monoclonal antibody OMP-18R5 (Gurney et al. PNAS 2012). The FRIZZLED CAR cells showed cytotoxic effects on FZD5-expressing MDA-MB453 cells but not on parental cells, demonstrating that the engineered NK-92 cells are capable of killing organoids involving CAR-dependent mechanisms. By using this line of CAR cells, the authors showed that there was no significant difference in the CAR-mediated cell lysis between wild-type, APC-KO and Rnf43/Znrf3-DKO organoids.

We thank the reviewer for carefully delineating the rationale of our approach and summarizing our data.

Based on this observation, the authors claimed that the strategy to target Frizzled proteins by CAR systems may have potential risk of cytotoxicity against normal colonic tissue. Furthermore, it appears to me that, on the grounds of these data, the authors claim that the co-culture system presented in this study may serve as a suitable preclinical assay to foresee the risk of therapy. However, this is not a persuasive argument unless the data on in vivo effects of the FRIZZLED CAR cells are presented. The pan-FRIZZLED antibody OMP-18R5 was previously shown to inhibit the growth of a variety of tumor cells including colon tumors in xenograft models (Gurney et al. PNAS 2012). Gurney et al. showed in this previous paper that administration of OMP-18R5 did not

show any apparent side effects but resulted in colitis, presumably due to suppression of the Wnt pathway, when it was used at very large doses. The present study should test the in vivo effect of FRIZZLED CAR cells. Do these CAR cells have anti-tumor activity together without any lytic effects on Frizzled-expressing normal tissues? Or do they induce severe colitis as in the previously reported xenograft models treated with high-dose OMP-18R5? If we see a safe therapeutic window in which colon tumor xenografts can be targeted without any other tissue damage, this may suggest that the co-culture model presented in this study is too sensitive to predict the safety of CAR systems appropriately.

We are thankful for this comment, and we agree that our model cannot substitute for toxicity testing in vivo. CAR-toxicity may include off-cancer activity towards multiple organs and may be strongly modulated by the immune system. Therefore, we have now avoided all speculative claims on safety assessment of CAR therapy in vivo. Yet, our data show that the newly created FZD-CAR cells exhibit no therapeutic window towards isogenic mouse organoids (Fig. 6A and new Fig. 6D) and human CRC organoids of relevant genotypes (new Fig EV6). Therefore, we think that our system can serve as an informative model for early preclinical target validation.

The OMP-18R5 antibody (that our CAR is based on) has shown some efficacy against xenografted cancer cells and toxicity only at high doses. We want to stress that these preclinical data were not our motivation to choose Frizzled as a target. Rather, our approach was based on the rational strategy that RNF43 mutant tumor cells exhibit increased Frizzled surface levels based on knock-down studies in cell lines and genetic evidence in mice (Koo et a., Nature 2012; Hao et al., Nature 2012). Our platform has allowed testing this hypothesis by targeting the endogenous protein in primary cells. In this context the presented data clearly shows toxicity towards normal cells and no therapeutic benefit.

Lethal adverse effects have been described in CAR therapy and a case report of ERBB2 CAR-T cells describes fatal cytokine release syndrome (Morgan et al., Molecular Therapy 2010). This was not predicable from the toxicology of the therapeutic antibody that the CAR was based on (Herceptin), arguing for the importance of sensitive in vitro models. We also agree that our Frizzled-CAR NK-cells would represent an interesting model for future *in vivo* toxicity studies, given the cross-species reactivity of the CAR-NK line. However, such *in vivo* experiments would require careful dose escalation, followed by serological and histological analysis, and establishment of appropriate tumor models, which would exceed the scope of our study und certainly justify an independent project.

The authors should also perform flow cytometry or some quantitative assays to show the expression levels of Frizzled proteins on wild-type, APC-KO and Rnf43/Znrf3-DKO organoids cultured in the presence and absence of Rspo1. This will help readers understand how DKO of Rnf43/Znrf3 or Rspo1 treatment induces the change of Frizzled expression relative to the basal expression levels.

We are thankful for this comment. However, it has not been possible to measure endogenous surface levels of Frizzled in gastrointestinal cells. Immunofluorescence detection using the pan-Frizzled antibody (OMP-18R5) has not been successful in our hands, most likely due to the low protein levels. FACS experiments are challenging because Frizzled receptors are sensitive to protease treatment, which is required for single cell dissociation. However, there is strong molecular evidence that the conditions and genotypes studied here (summarized in Fig. 6A), represent distinct surface levels (Koo et a., Nature 2012 and Hao et al., Nature 2012) and we have indirectly visualized Frizzled levels previously in the same organoids by surface detection of the Wnt-ligand (Farin et al., Nature 2014). We now show functional dependence on Frizzled-signaling by IWP-2 sensitivity of human RNF43 deficient CRC organoids. Furthermore, our comparison of cytotoxicity of Frizzled-CAR cells and EPCAM-CAR cells that target a uniformly expressed antigen (Fig. EV6BC) argues for a non-tumor-specific activity.

Minor

The authors described that EGFRviii, which was introduced by using the lentiviral system, was undetectable by immunoblot assays due to low expression levels (Fig. 4C). Are the organoids analyzed in Fig. 4A-E the same as those assayed in Fig. 4F-I, which co-expressed detectable levels

of EGFP or RFP? If so, what do the authors think is the reason for different levels of expression between EGFRviii and fluorescent proteins? If not, the immunoblot data of the EGFRviii expression in organoids doubly positive for EGFRviii and EGFP should be presented. The authors will be able to try immunohistochemistry to detect truncated EGFRviii on sections of fixed organoids.

Yes, the same lines are shown here and were also used for the WB analysis. But, the expression level of small and cytoplasmic reporters (EGF/RFP) cannot be compared to a large transmembrane receptor. In our revision we have not been successful to increase the sensitivity our WB detection of EGFRvIII but we have now confirmed transgene expression using qRT-PCR analysis. In fact a similar mRNA expression level as for HPRT was found (Fig. EV4D). Although protein translation was selected by continuous culture in the presence of antibiotics (EGFRvIII-P2A-blasticidin lentivirus), it appears that only very low EGFRvIII protein levels can be achieved by lentiviral overexpression. This could be due to a limiting protein trafficking and/or posttranscriptional regulation. Because we could detect a strong and specific cytotoxic response against EGFRvIII transduced organoids, our data indicates that these low levels are sufficient for CAR-mediated killing.

References

Farin HF, Jordens I, Mosa MH, Basak O, Korving J, Tauriello DVF, de Punder K, Angers S, Peters PJ, Maurice MM, et al. (2016) Visualization of a short-range Wnt gradient in the intestinal stem-cell niche. *Nature* **530**: 340–343.

Hao H-X, Xie Y, Zhang Y, Charlat O, Oster E, Avello M, Lei H, Mickanin C, Liu D, Ruffner H, et al. (2012) ZNRF3 promotes Wnt receptor turnover in an R-spondin-sensitive manner. *Nature* **485**: 195–200.

Morgan RA, Yang JC, Kitano M, Dudley ME, Laurencot CM, Rosenberg SA (2010) Case Report of a Serious Adverse Event Following the Administration of T Cells Transduced With a Chimeric Antigen Receptor Recognizing ERBB2. *Mol Ther* **18**: 843–851.

Koo B-K, Spit M, Jordens I, Low TY, Stange DE, van de Wetering M, van Es JH, Mohammed S, Heck AJR, Maurice MM, et al. (2012) Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors. *Nature* **488**: 665–669.

2nd Editorial Decision 12th Mar 2019

Thank you for submitting the revised version of your manuscript. My apologies again for the delay in processing your revised manuscript which was due to protracted referee input. Your revised study has now been re-evaluated by the three original referees, please find their comments enclosed below. As you will see the referees find that their concerns have been sufficiently addressed and they are now broadly favour of publication.

Thus, we are pleased to inform you that your manuscript has been accepted in principle for publication in The EMBO Journal, pending some minor issues regarding formatting and data representation, as outlined below, which need to be adjusted at re-submission.

REFEREE REPORTS:

Referee #1:

I have looked at the authors' responses and I am happy with this. Looking at the rebuttal, it seems that the authors have answered reviewers' comments well.

Referee #2:

The authors did an excellent job in addressing all my concerns. I suggest acceptance of this exciting manuscript.

Referee #3:

The authors have fully addressed my concerns in the revised manuscript. The changes made in response to other reviewers' comments have also strengthened the manuscript.

2nd Revision - authors' response 24th Mar 2019

The authors performed the requested editorial changes.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Henner Farin
Journal Submitted to: EMBO Journal
Manuscript Number: EMBOJ-2018-100928T

Reporting Checklist For Life Sciences Articles (Rev. June 2017)

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

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

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

2. Captions

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

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple $\chi 2$ tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

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

In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

B- Statistics and general methods

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

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	Sample sizes were generally determined based on previous studies involving similar experimental setup.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	as described in the methods: Killing experiments were excluded if the CAR-specific lysis in positive controls was less than 30% and the experiment was repeated using a freshly recovered cryovial of CAR NK-92 cells. If organoid/matrigel layer accidently detached during the washing steps the respective wells were marked and excluded from the analysis.
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	No group allocation was required for the experiments, no randomization procedure was performed. Automatic imaging-based analysis was performed on randomly chosen imaging fields.
For animal studies, include a statement about randomization even if no randomization was used.	NA
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	No group allocation was required for the experiments, no blinding procedure was performed. Automatic imaging-based analysis was performed using similar analysis settings for all conditions.
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	Yes statistical tests were chosen, depending on kind of comparision and single or multiple comparisons. Appropriate statistical tests were performed. As detailed in Materials and Methods: All statistical analysis was performed using GraphPad Prism 7.0d software. Data are presented as
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	the data is normally distributed
Is there an estimate of variation within each group of data?	only standard deviation is shown
Is the variance similar between the groups that are being statistically compared?	unequal variance was assumed

USEFUL LINKS FOR COMPLETING THIS FORM

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

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

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

http://datadryad.org

http://figshare.com

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

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

http://biomodels.net/

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

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	Antibodies were anti Myc-tag (clone 9E10) 647-conjugated from Santa Cruz Biotechnology (cat: sc-40 AF647) anti-EGFR (clone D38B1) Cell Signaling Technology (cat: #4267) anti EPCAM (clone HEA125) APC-coupled; Miltenyi Biotec (cat: 130-113-822) anti-EPCAM (clone REA764) VioBlue-coupled; Miltenyi Biotec (cat: 130-111-004) anti-CD45 (clone 2D1) APC-conjugated; Invitrogen (cat: # 17-9459-42)
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	HEK-293T (ICLC HTL04001). Routinely tests for mycoplasma contamination was performed

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

D- Animal Models

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

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	The establishment and use of human organoids was approved by the institutional Review Boards of the University Cancer Center Frankfurt (UCT) and the Ethics Committee at the University Hospital Frankfurt (project-number: SGI-06-2015).
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	Informed written consent was optained for all patient samples. Patient identity for pathological specimens remained anonymous in the context of this study.
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	NA
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	Human samples can only be distributed after permission by the ethics commitee.
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	NA
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	NA
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	NA

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	NA
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	excact image analysis pipeline is reported in the Appendix figures 2 and 3
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	NA

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	NA