nature portfolio

Peer Review File

Complementary dual-virus strategy drives synthetic target and cognate T-cell engager expression for endogenous-antigen agnostic immunotherapy

Open Access This file is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. In the cases where the authors are anonymous, such as is the case for the reports of anonymous peer reviewers, author attribution should be to 'Anonymous Referee' followed by a clear attribution to the source work. The images or other third party material in this file are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author): with expertise in oncolytic virotherapy, cancer immunotherapy

In this work, the authors pursue an approach to overcome the limitations of targeted cancer immunotherapies by engineering an oncolytic VSVMΔ51 to express a truncated target antigen (HER2) to allow for tumor antigen targeting with an ADC. Moreover, they combine the truncated HER2-expressing VSVMΔ51 with an oncolytic VV expressing a T cell engager targeting HER2 for enhanced therapeutic effects. This work represents a novel use of oncolytic virotherapy to modify the tumor in order to enable off-the-shelf targeted therapies, as a more robust alternative to personalized therapies.

This work comprises an impressive multitude of tumor models and methods, as well as a rationally designed approach. The manuscript is well-written and easy to follow. A general limitation of the strategy is that the targeted therapy will only address the infected tumor cells, which in principle, may already be killed by the direct oncolytic effect of the virus.

Additional specific comments:

1. What is the mechanism of the proposed enhanced virus replication in the presence of T-DM1 (trastuzumab)? As the titers are only expressed as a fold-increase, the differences appear more substantial than they might really be. If expressed as absolute titer, it would seem that the increase in the presence of trastuzumab would only be around a 1/2-log.

2. In the tumor models where bilateral subcutaneous tumors are implanted, are IT injected treatments administered to both tumors, or only 1?

3. It is unclear why ex vivo treatments were performed from in vivo tumors of implanted mouse cell lines. While this is obviously the only option for treatment of patient samples, the mouse studies could have easily been performed from in vitro cultured cell lines, and the rationale for implanting them and then investigating the treatment ex vivo is not clear.

4. OV infections in the presence and absence of T-DM1 in patient specimens are shown in Figures 1e/f and Extended Figure 2c/d. Viral titers were only increased in the samples shown in Figure 1. What is the difference between the patient samples in the two figures, and why were titers only enhanced in Figure 1? Also, in addition to the virus titers as a readout, it would be helpful to show cell viability/cytotoxicity.

5. Although not explicitly stated in the Methods section or Figure Legends, it appears from the treatment schemes (Figures 1g and 5a) that treatment already began in the lung tumor model on day 2 post-implantation. Were the tumors already detectable at this time-point? Why was treatment started so early?

6. In Figure 3c and d, the VV-antiHER2 TCE monotherapy seems to have some effect (is this statistically significant?) compared to VV-Ctrl TCE in MC38 cells, despite the fact that MC38 do not express HER2. How can this be explained?

7. In general, it seems that the OV monotherapies are not effective in the in vivo models. Why is this? Even without the additional targeting effect, one would expect that a direct oncolytic effect from the virus should cause delayed tumor growth and prolonged survival, but this does not appear to be the case in the tumor models used.

8. For Figure 5c, it is stated in the text that VSVΔ51+VV led to a 60% reduction in tumor nodules compared to VSVΔ51-HER2T alone; however, according to the p-value shown on the graph (0.57), this is not statistically significant.

9. In Figure 5f, p-values are shown for 2 of the treatments, but it is not indicated what the treatments are being compared to. For VSVΔ51-HER2T + VV-antiHER2-TCE, no p-value is indicated. Is the survival significant compared to the respective control therepies?

10. It is stated in line 288 that VSVΔ51-HER2T + VV-antiHER2-TCE treatment enhanced infiltration of CD45+ leucocytes; however, this does not appear to be the case according to the data shown in Extended Data Figure 10f. Please explain.

11. The majority of the data are presented in the Extended Data section. This makes it quite cumbersome to follow the story, because one needs to constantly flip back and forth from the main body to the Extended Data section. Could some of this data be shifted to the main figures?

Minor comment:

1. In Extended Figure 1d, for MOI 1, what is time-point “w”?

Reviewer #2 (Remarks to the Author): with expertise in oncolytic virotherapy

The article proposes an interesting new approach to treating tumors using oncolytic viruses and anti-HER2 therapy. Despite the potential difficulty of implementing this approach into clinical practice, the concept itself is new and promising. In addition to a tumor-selective strain of vesicular stomatitis virus expressing a truncated form of HER2, a strain of vaccinia virus expressing a HER2-targeted T-cell activator (TCE) was used. The synergistic effect of the two approaches ensures the formation of a systemic antitumor immune response. All experiments were carried out with a substantial and controls were thoughtfully executed. To my opinion, the validity and reliability of the results is beyond doubt. The work was carried out at the excellent level. I am sure that the work would have important impact to the entire fields of oncolytic viruses and cancer immunotherapy.

Although I found the factual presentation as impeccable, I would expect a little more detailed presentation of the results and discussion, if the article’s format permits.

Minor comments:

Lines 111 -114 Figure 1 g Colocalization is not entirely obvious; in the single section shown, the areas stained with antibodies for VSV and HER2 only partially overlap

The article could be supplemented with a detailed histological analysis of tumors after injection to characterize the type of cell death of tumor cells. This is especially true for the most significant in vivo experiments about the combination of two viral strains

The only wish is to obtain a strain of vaccinia virus that, without any special problems for replication efficiency, could express both therapeutic proteins - HER2T and TCE- at different loci, or even in one. The selectivity of the vaccinia virus for tumor cells is very high, there will

be complete colocalization, and in addition, with demonstrated effectiveness, there will be no problems with the approval of such an approach for use in clinical practice.

Reviewer #3 (Remarks to the Author): with expertise in oncolytic virotherapy, cancer immunotherapy

In this manuscript the authors seek to deliver a non-signaling HER2T (T for truncated) protein to tumors and a VV to deliver a T Cell Engager (TCE) that will pull in T cells to recognize the HER 2T expressed on the tumor cells and activate a T cell response against the tumors. This is an attractive strategy that extends the concept of OV beyond whether the 'pure oncolysis' or the 'oncolysis leading to T cell priming' strategies. Overall, the experiments are plentiful and do suggest that the combination of novel antigen display along with T cell recruitment can potentially generate therapy which allows for shaping the tumor to the therapy. However, there are some points that I feel could greatly strengthen the story, especially for this Journal.

Conceptual:

Whilst I understand completely that these groups are a world leader in the use of both VSV and VV as oncolytics, if the goal is to express a novel tumor antigen (TA) de novo on tumor cells and then to use that TA to engage incoming T cells, is VSV really the most valuable vector system to do this? Presumably the cells infected by the VSV-HER-2T will also be predominantly killed by the virus (the 'pure oncolysis' dogma) and so de novo display of the TA will be transient. The window then for (non viral) T cells to come into the tumor, see the TA and engage new T cells will be limited – compared to for example display of the TA by a non lytic replicating virus. Some of my comments below reflect this concern which may be why some of the models are both artificial (treatment rapidly after tumor seeding for iv delivery) and some of the in vivo effects are rather limited in efficacy. Perhaps the authors could address this slight disconnect with virus type for delivery of the TA and therapy – if they agree.

A major hurdle with TCE or vaccination or OV strategies is achieving high level virus

expression in all the tumor cells and/or antigen heterogeneity in tumors being targeted with TCE or T cells. Thus, tumors escape these strategies very efficiently due to heterogeneity of TA expression and/or antigen loss. With the strategy described, there is I believe, very little chance of infecting 100% of the tumor cells and generating T cells responses against all of these cells. So in a sense, this inefficient delivery of the TA to a fraction of the tumor cells does not improve upon the issues of tumor heterogeneity and antigen loss – as perhaps some of the moderate efficacy expts suggest. This raises the issue of whether there is really endogenous anti-tumor immunity generated – and this is a concern I raise below.

The authors raise the issue of the dependence of this strategy upon co-delivery of two different viruses. This may be less of an issue for s.c. or even ip delivery in a mouse but clinically this seems a very difficult concept to take on board given the hurdles experienced so far in achieving high efficiency in vivo delivery of systemic OV clinically. The TCE presumably does not have to be expressed in the same tumor cells as the VSV-HER-2T but they will have to be co-expressed relatively close. Once again the very artificial nature of the iv delivery expts used in the manuscript suggest that this may be difficult as a translational strategy (please see below).

Specific Comments:

Figure 1

Do the authors have some view on what percentage of the tumor cells that are infected by the VSV-HER-2T that are not killed by oncolysis will then be killed by the T-DM1? This speaks to the nature of any bystander effect – ie is there added value to the T cell engagement of killing cells infected by the OV?

In Figure 1g-i., the authors try to show that iv delivery of the dual virus strategy is effective. They use a model in which tumor is seeded i.v. and then viruses are given 1 day later. They count resultant lung nodules at day 10. This is not a model of treatment of established tumors by this strategy. To convince the reader that here is true iv delivery of the two virus strategy to treat tumours the treatment should be given once we are certain that there are

seeded (established) tumors in the lungs. This expt also lacks the control of VSV-GFP or VSV-luciferase which is very important – if, for example, the tumor cells are simply infected by the virus before being established as tumors (making iv delivery much easier than having to penetrate an established tumor) the oncolytic virus alone needs to be tested. The authors show a significant difference between the virus+TCE in F. compared to the virus alone suggesting a real difference. However, is it not likely/possible that the VSV-HER-2T displays the HER-2T TA and that the T-DM1 treatment clears these tumor cells way before the tumor is established?

Why are there only 5 mice per treatment group? Please provide a full explanation of how this sample size provides enough power to make the statistical conclusions drawn.

In my opinion, this expt should be repeated to show iv delivery is efficacious against clearly established tumor, that the VSV-HER-2T virus is better than the control virus, and with large sample sizes to make valid conclusions. This seems important to me to support the authors claim that VSV-HER-2T can be used in combination with an off the shelf therapy already in the clinic.

Figure 2:

'The αHER2-TCE was further shown to induce target-dependent cytolytic activity by reducing the viability of HER2+ or HER2T+ target cells in the presence of naïve syngeneic splenocytes (Fig. 2h).' These are splenocytes used in these experiments. Could the authors not have used CD3 T cells to prove the activity of the TCE? These mouse tumor cells are expressing a human HER-2 which could stimulate various innate and or xenogeneic T cell responses. Why was the Control TCE (with an anti-human CD3) not used in these experiments?

Figure 3:

The quality of Fig.3A is very poor at least in my version. Why is GFP examined?

J69 cells are human T cells and the MC38 targets are mouse T cells. If the mouse T cells are engineered to express an anti-human CD3 TCE is it clear that this TCE is not enhancing a xenogeneic reactivity against the mouse cells? Would it be good to have the control VSV-GFP in this mix to guard against that possibility?

Figure 4:

4C: 'P values indicated next to each treatment group relative to VSVΔ51-HER2T + VV-αHER2-

TCE'. My apologies for being unclear. There is a p value of 0.0017 shown against the VSVΔ51-HER2T + VV-αHER2-TCE line. How can this be significant against itself?
Why are there only 5 mice per treatment group? Please provide a full explanation of how this sample size provides enough power to make the statistical conclusions drawn. How many times was this expt repeated?

D,E: Am I correct in the reading of this expts that 2 mice survived long term the treatment with VSVΔ51-HER2T + VV-αHER2-TCE; these two mice were then re-challenged on the left and right flank with MC438 parental (left flank) and MC38-HER-2T (right flank). The MC38 left flank tumors grew but the right flank MC38-HER-2T tumors were rejected.
If I am correct I think the authors could make this a little clearer in the Figure legend. More importantly this shows that the effect of the therapy was to raise endogenous immunity against the human HER-2T and not against the tumor cells themselves. This is contrary to the authors' overall claims that this strategy raises endogenous anti-tumor T cell responses and is rather ignored in the discussion.

I&J. Here the same experimental set up is in place and the rejection of the parental CT26 tumors on the left flank is more convincing of a genuine anti-tumor T cell response. However, this is not a good control to show the generation of anti-tumor immunity. When naïve mice are re-challenged with the CT26 cells they do not reject the tumor. When the 2 cured mice are re-challenged they do. However, the cured mice have already seen the CT26 tumors – so it may be that any mouse vaccinated with CT26 (perhaps irradiated vaccine etc) would reject the re-challenge (ie CT26 is an immunogenic tumor). Therefore, rejection of the CT26 by the cured mice does not show that treatment itself generates the anti-tumor immunity; it could be simply that the CT26 is inherently immunogenic and any or no treatment generates anti-tumor immunity.
These data using only 2 surviving/re-challenged mice are not statistically enough to make the conclusions the authors draw.
I think it would be much more persuasive if the authors could show that treated and cured mice generate real anti-tumor CD8+ T cell responses enhanced by the VSVΔ51-HER2T + VV-αHER2-TCE combination treatment. This could be done much more elegantly by ELISPOT, ICS or even ELISAs against parental tumors or even against known peptide epitopes from

these tumor models. In addition, real anti-tumor immunity could be shown by T cell depletion expts. On balance, I do not think the authors have shown what they claim that immunological memory against the tumor is generated. This is also in the title: '…. or endogenous-antigen agnostic immunotherapy' so I think that precision, statistical relevance and full interpretation of the data is important.

Figure 5:

A-C: The same critique of the experimental set up holds here. Virus delivery 1 day after IV injection of tumor cells does not show efficacy against established tumors. I think this should be extended to convince the reader of the efficacy and feasibility of this approach of delivery of two viruses to established tumors.

D-E. Is it a surprise that the VSV-Her-2T virus has no therapy by itself? The therapy in this ID8 model is very modest – is it T cell regulated in the way that the authors are claiming? There are no immune correlative studies to show this.

Statistics need to be explained and validated throughout the manuscript.
For example, in Extended Data Fig.3F. '(f) overall survival was monitored. P-values indicated next to each treatment group relative to VSVΔ51-HER2+T-DM1.' There is a p value of 0.0090 next to the VSVΔ51-HER2+T-DM1.This makes no sense unless I am mis-understanding this.

Extended Data Fig. 9C: I find it very difficult to believe that there is any statistical difference between the VSV-HER2-T + VV-Cntrl group and the VSV-HER2-T+VV-anti-HER2 TCE (co) group. And yet the text reads:

To evaluate the extent of tumour control our dual-virus combination exerts, we implemented our approach in disseminated disease models. First, we optimized a co-

formulated dose containing both OVs and validated this co-formulation yields similar efficacy to our original sequential treatment regimen (Extended Data Fig. 9a-c). This co-formulation enables more rapid administration within a shorter timeframe, allowing the treatment to fit within the narrow therapeutic window in these disseminated disease models. We moreover demonstrated that switching the order of the OVs in the sequential regimen does not significantly impact efficacy (Extended Data Fig. 9c).

I do not believe that these data show therapy of the combination. Therefore they do not show that switching the timing does not impact efficacy. There is no statistical difference between the different timings of the therapy but if the therapy is not significant itself relative to controls then this is not a valid conclusion. I think that throughout the manuscript the authors need to be considerably more rigorous with their statistics, given their sample sizes and their interpretations of the data.

In summary, I apologize for such a long review which actually results from my excitement about the overall concept. This is an attractive concept from an excellent group which is why the review is so long. Multiple models are tested and some interesting data suggestive of efficacy provided. There are some conceptual issues which could, perhaps, be addressed in more rational detail (how an oncolytic virus will display the TA long enough to interact with the TCE and how systemic delivery of 2 vectors will be possible clinically). The immune studies here are rather crude and not, in my opinion, persuasive of the mechanisms that are proposed due to low sample sizes ((2 survivors re-challenged, a lack of effective assays (ELISPOTS, ICS, depletions etc), immunity against a human HER-2T may be a confounding factor, and inconsistencies about whether real immune T cells reactivity against endogenous TA are really raised). The statistical power of many experiments seems very low and interpretations of the data are not always persuasive or clear. Finally, some of the model systems used are not realistic; whilst we all put our best foot forward with the optimal models, using IV delivery to 1 day old tumors is not convincing for a systemic delivery model. Overall these issues detract for me from the impact of the work.

POINT-BY-POINT RESPONSE TO REVIEWER COMMENTS:

Reviewer #1

In this work, the authors pursue an approach to overcome the limitations of targeted cancer immunotherapies by engineering an oncolytic VSVMΔ51 to express a truncated target antigen (HER2) to allow for tumor antigen targeting with an ADC. Moreover, they combine the truncated HER2-expressing VSVMΔ51 with an oncolytic VV expressing a T cell engager targeting HER2 for enhanced therapeutic effects. This work represents a novel use of oncolytic virotherapy to modify the tumor in order to enable off-the-shelf targeted therapies, as a more robust alternative to personalized therapies.

This work comprises an impressive multitude of tumor models and methods, as well as a rationally designed approach. The manuscript is well-written and easy to follow. A general limitation of the strategy is that the targeted therapy will only address the infected tumor cells, which in principle, may already be killed by the direct oncolytic effect of the virus.

We thank the reviewer for the positive comments about our manuscript. We have expanded on this point regarding HER2T-expression within infected cells in the discussion. We have also performed a series of experiments that address this point (Extended Data Fig. 11). Briefly, we injected the virotherapies into CT26 tumours implanted in BALB/c mice. We harvested these tumours at 24 and 48 hours post-injection and assessed them by flow cytometry or by IHC to determine the proportion of infected cells. Indeed we observed a detectable level of VSVΔ51-HER2T infected cells, expressing HER2T, until at least 48 hours post-injection. These data suggest that infected cells can be successfully labelled with HER2T and persist following infection, without undergoing lysis.

Specific comments to address:

1. What is the mechanism of the proposed enhanced virus replication in the presence of T-DM1 (trastuzumab)?

We thank the reviewer for this question, and we have briefly expanded on this in the text: "Our group previously published on the mechanistic synergy between microtubule destabilizing agents (MDAs) and VSVΔ51. We found that MDAs enhance VSVΔ51 spread and bystander killing within tumours, as a result of disrupting type I interferon secretion. We moreover demonstrated that ADCs bearing an MDA payload lead to the same combinatorial synergy with VSVΔ51 although in a very specific targeted fashion."

2. As the titers are only expressed as a fold-increase, the differences appear more substantial than they might really be. If expressed as absolute titer, it would seem that the increase in the presence of trastuzumab would only be around a 1/2-log.

Thank you for this question. In Fig. 1e the absolute viral titers are shown, and Fig. 1f the fold-change in viral titers are shown. We agree that in many of these specimens the increase in viral titer is modest, but the general trend is positive and often statistically significant, with 3 samples showing a log-fold increase (or greater) (Ovarian 3, Breast 1, Parotid 2) when VSVΔ51-HER2T is used. Importantly, we report no impact of the combination when the control virus VSVΔ51-GFP is used. In sum, we approach these results with excitement, but we recognize that there is sample variability due to patient tumour heterogeneity and sample quality.

3. In the tumor models where bilateral subcutaneous tumors are implanted, are IT injected treatments administered to both tumors, or only 1?

We apologize for any confusion. All bilateral tumour models used were for tumour rechallenge experiments done using cured mice, or control naïve mice. In this case, tumours implanted in the left flank are parental tumours, while tumours implanted in the right flank are HER2-expressing derivatives of the parental tumours. No tumours were injected in these bilateral models, but the tumour progression of both tumours were monitored. We have clarified this in the materials and methods as follows:

"All rechallenge studies were performed D90 post-implantation, by implanting bilateral tumours. For these bilateral implantations, cells were seeded at double the original seeding density, distal or contralateral to the initial site of tumour implantation. Parental tumours were implanted in the contralateral side, while HER2-expressing derivatives of the parental tumours were implanted in the right flank at a distal site from the initial tumour. Tumour progression and overall survival were monitored over time."

4. It is unclear why ex vivo treatments were performed from in vivo tumors of implanted mouse cell lines. While this is obviously the only option for treatment of patient samples, the mouse studies could have easily been performed from in vitro cultured cell lines, and the rationale for implanting them and then investigating the treatment ex vivo is not clear.

Thank you for this question. Indeed, *ex vivo* tumour treatment was the only option for patient tumours, and these patient samples were scarce. For this reason, we wanted to ensure we had an optimized treatment pipeline for these patient samples to avoid wasting precious specimens. As such, we used mouse tumour cores to refine and optimize our tumour core treatment pipeline.
In addition, the *ex vivo* treatment of mouse tumours was a method for us to validate our combination strategies in a 3-dimensional model, which is a key stepping stone prior to moving to *in vivo* mouse experiments. We have included this rationale in the results, as follows:
"This was first done in these CT26 tumour cores ex vivo so as to refine the treatment conditions before implementing more challenging patient post-surgical specimens."

5. OV infections in the presence and absence of T-DM1 in patient specimens are shown in Figures 1e/f and Extended Figure 2c/d. Viral titers were only increased in the samples shown in Figure 1. What is the difference between the patient samples in the two figures, and why were titers only enhanced in Figure 1? Also, in addition to the virus titers as a readout, it would be helpful to show cell viability/cytotoxicity.

Thank you for this question. Figure 1e shows the absolute viral titers following treatment of patient tumour cores with VSVΔ51-**HER2T** + T-DM1. Extended Figure 2c shows absolute viral titers following treatment of patient tumour cores with the *control* virus VSVΔ51-**GFP** + T-DM1.
Viral titers were only enhanced in Figure 1e because the presence of VSVΔ51-induced HER2T enabled binding of T-DM1, which internalized and enhanced VSVΔ51 output. In the control condition using VSVΔ51-GFP, there was no target for T-DM1 to bind to, and therefore there was no enhancement of VSVΔ51 output. This was clarified in the text as follows:
"VSVΔ51-HER2T + T-DM1 (Fig. 1 e,f), or VSVΔ51-GFP + T-DM1 as control (Extended Data Fig. 2c,d). Viral titers were only enhanced in the presence of VSVΔ51-induced HER2T, which enabled binding of T-DM1 and its subsequent internalization, leading to enhanced VSVΔ51 output. In the control condition using VSVΔ51-GFP, there was no target for T-DM1 to bind to, and therefore there was no enhancement of VSVΔ51 output."

Unfortunately we no longer have these specimens, as they had to be frozen and thawed to enable quantification of viral titer. As such we are unable to perform cytotoxicity/viability assays on prior samples, but we have previously demonstrated that VSVΔ51 +T-DM1 combination leads to reduced viability in HER2+ cells. (*Arulanandam, R., Taha, Z., Garcia, V. et al. The strategic combination of trastuzumab emtansine with oncolytic rhabdoviruses leads to therapeutic synergy. Commun Biol 3, 254 (2020). https://doi.org/10.1038/s42003-020-0972-7).*

6. Although not explicitly stated in the Methods section or Figure Legends, it appears from the treatment schemes (Figures 1g and 5a) that treatment already began in the lung tumor model on day 2 post-implantation. Were the tumors already detectable at this time-point? Why was treatment started so early?

We thank the reviewer for this question. We and others have previously utilized this lung metastasis model, whereby we inject cells i.v. by tail vein injection, and assess the impact of subsequent systemic therapy upon the number of lung nodules. We have also evaluated the success of treatments starting at multiple timepoints following i.v. cell injection, and concluded that treatment at day 1+2 is the only effective therapeutic window, explaining why we begin treatment very early.

To directly address this comment, we performed a set of experiments that validate our model of lung metastasis. Briefly, we injected 4T1.2 cells stained with an infrared dye, by tail vein injection into mice. We imaged these mice by fluorescence IVIS imaging at 1 hour and 24 hours post-injection, and detected the injected cells within the lungs. We also harvested these lungs at 24 hours, fixed, paraffin embedded, and H&E stained the lung sections; a pathologist evaluated these sections and confirmed presence of the 4T1.2 cells within the lung sections at the 24 hour timepoint post-injection. These data have been incorporated into Extended Data Figure 3a-e, the text, and also into the materials and methods section. We thank you again for your question, as it has helped us strengthen the validation of our model.

To add, other groups have also assessed the phenotypic and genetic characteristics of this model, and it has been described as a metastasis model that bypasses extravasation. It was found that there are no genetic or phenotypic differences between lung metastases that form as a result of gradual metastases disseminating from a primary tumour *vs* those that form as a result of i.v. tail vein injection. For these reasons, we have implemented this model as it enables us to demonstrate therapeutic efficacy of our strategy following systemic administration. Relevant references are indicated below and have been added to our text.

- Lücke J, Zhang T, Zazara DE, Seeger P, Izbicki JR, Hackert T, Huber S, Giannou AD. Protocol for generating lung and liver metastasis in mice using models that bypass intravasation. STAR Protoc. 2024 Mar 15;5(1):102696. doi: 10.1016/j.xpro.2023.102696. Epub 2024 Jan 18. PMID: 38244200; PMCID: PMC10831314.

- Rashid OM, Nagahashi M, Ramachandran S, Dumur CI, Schaum JC, Yamada A, Aoyagi T, Milstien S, Spiegel S, Takabe K. Is tail vein injection a relevant breast cancer lung metastasis model? J Thorac Dis. 2013 Aug;5(4):385-92. doi: 10.3978/j.issn.2072-1439.2013.06.17. PMID: 23991292; PMCID: PMC3755653.

- Taha Z, Crupi MJF, Alluqmani N, Fareez F, Ng K, Sobh J, Lee E, Chen A, Thomson M, Spinelli MM, Ilkow CS, Bell JC, Arulanandam R, Diallo JS. Syngeneic mouse model of human HER2+ metastatic breast cancer for the evaluation of trastuzumab emtansine combined with oncolytic rhabdovirus. Front Immunol. 2023 Apr 19;14:1181014. doi: 10.3389/fimmu.2023.1181014. PMID: 37153626; PMCID: PMC10154558.

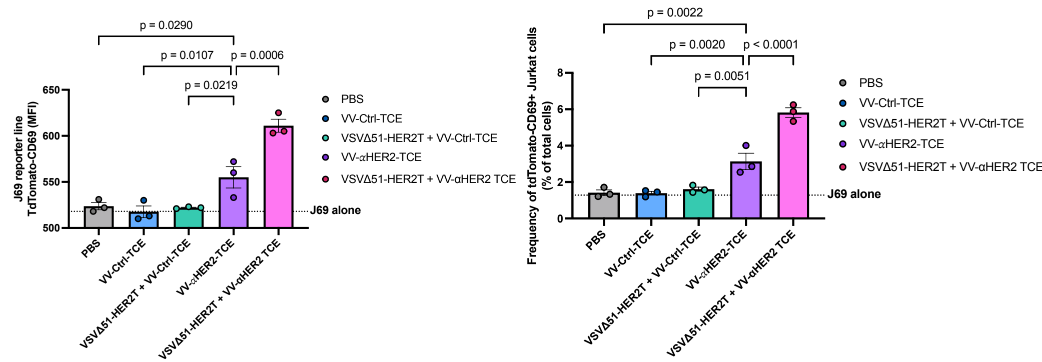
7. In Figure 3c and d, the VV-antiHER2 TCE monotherapy seems to have some effect (is this statistically significant?) compared to VV-Ctrl TCE in MC38 cells, despite the fact that MC38 do not express HER2. How can this be explained?

Thank you the reviewers for this question. This activation of J69 T-cells by the VV-αHER2-TCE alone is a result of non-specific T-cell activation through the functionally intact anti-mouse CD3 scFv of the TCE. These results are statistically significant, as confirmed by a one-way ANOVA with Dunnett's correction for multiple comparisons (indicated below). These statistics will be omitted from the manuscript figure to avoid clutter, as we aim to highlight the effect of the combination (pink bar). However, we observe this result only in the MC38 tumour cores, but not the CT26 or the human HT29 tumour cores, and we have moved these figures from the supplemental data into the main figure to better complement the data set.

In the context of therapeutic efficacy *in vivo*, VV-αHER2-TCE monotherapy had no impact on survival as demonstrated in the survival studies in Fig. 4 and 6, where neither the VV-αHER2-TCE monotherapy, nor the control combination VV-αHER2-TCE + VSVΔ51-Fluc, led to any therapeutic efficacy. This confirms that the

non-specific J69 activation may not be functionally anti-tumour when there is no targeting moiety (HER2T) present.

We clarified this point in the text, and we also added data from CT26 and human HT29 tumour cores where we do not observe the same extent of background J69 activation following infection with *VV-αHER2-TCE* alone.



8. In general, it seems that the OV monotherapies are not effective in the in vivo models. Why is this? Even without the additional targeting effect, one would expect that a direct oncolytic effect from the virus should cause delayed tumor growth and prolonged survival, but this does not appear to be the case in the tumor models used.

To simply some of our survival plots and reduce clutter (Fig 4, Extended Data Fig. 9), we initially removed our untreated PBS control groups all of which endpoint uniformly early. Our statistical comparisons in these figures are against the VV-αHER2-TCE monotherapy, which we believe is a more robust and relevant control group than the PBS control. However, we have added in the PBS control groups to Fig.4 and Extended Data Fig. 9. You can now appreciate the modest impact of virotherapy on survival compared to PBS alone in these figures, none of which are statistically significant. Other data (Fig 6a-c) demonstrate modest efficacy of OV monotherapy compared to PBS (VSVΔ51 leads to reduced lung metastases compared to PBS). The mouse tumour models we have implemented in our experiments are known to be refractory to OV monotherapy, and even combinations of OVs, making it critical to encode synergizing and rationally-selected payloads to enhance efficacy.

9. For Figure 5c, it is stated in the text that VSVΔ51+VV led to a 60% reduction in tumor nodules compared to VSVΔ51-HER2T alone; however, according to the p-value shown on the graph (0.57), this is not statistically significant.

We apologize for the confusion. The comparison mentioned in the text was between VV-αHER2-TCE+ VSVΔ51-HER2T (pink bar) *vs* VSVΔ51-HER2T alone, where the p-value is 0.0262 (one-way ANOVA). The Light green and yellow bars represent the control combination viruses, and these did not lead to a statistically significant reduction compared to VSVΔ51-HER2T alone, although we did observe a trend towards reduced nodule count. We have made the intended comparison more clear on the graph (**bolded p-value**).

10. In Figure 5f, p-values are shown for 2 of the treatments, but it is not indicated what the treatments are being compared to. For VSVΔ51-HER2T + VV-antiHER2-TCE, no p-value is indicated. Is the survival significant compared to the respective control therapies?

We apologize for the lack of clarity with this annotation style. We have changed this and all similar p-value annotation styles in Fig. 4, Fig. 6, Extended Data Fig. 4, and Extended Data Fig. 7.

11. It is stated in line 288 that VSVΔ51-HER2T + VV-antiHER2-TCE treatment enhanced infiltration of CD45+ leucocytes; however, this does not appear to be the case according to the data shown in Extended Data Figure 10f. Please explain.

We thank the reviewer for pointing this out. We modified the text to correct our interpretation of our results and take caution not to over-state our results. We have omitted that specific CD45 data piece as it was misinterpreted.

12. The majority of the data are presented in the Extended Data section. This makes it quite cumbersome to follow the story, because one needs to constantly flip back and forth from the main body to the Extended Data section. Could some of this data be shifted to the main figures?

That you for this suggestion; we have moved some of the extended data into the main figures (Fig. 4d,e, and Fig. 5).

Minor comment:

1. In Extended Figure 1d, for MOI 1, what is time-point “w”?

We clarified this notation in the figure and caption. “W” is the wash, used to rinse off excess virus from cells infected at this high MOI of 1.0.

Reviewer #2

The article proposes an interesting new approach to treating tumors using oncolytic viruses and anti-HER2 therapy. Despite the potential difficulty of implementing this approach into clinical practice, the concept itself is new and promising. In addition to a tumor-selective strain of vesicular stomatitis virus expressing a truncated form of HER2, a strain of vaccinia virus expressing a HER2-targeted T-cell activator (TCE) was used. The synergistic effect of the two approaches ensures the formation of a systemic antitumor immune response. All experiments were carried out with a substantial and controls were thoughtfully executed.
To my opinion, the validity and reliability of the results is beyond doubt. The work was carried out at the excellent level. I am sure that the work would have important impact to the entire fields of oncolytic viruses and cancer immunotherapy.
Although I found the factual presentation as impeccable, I would expect a little more detailed presentation of the results and discussion, if the article's format permits.

We thank the reviewer very much for their kind words and positive comments. We have added additional data to the manuscript, additional details to the discussion, and clarified some of the presentation of different data pieces.

Minor comments:

1. Lines 111 -114 Figure 1 g Colocalization is not entirely obvious; in the single section shown, the areas stained with antibodies for VSV and HER2 only partially overlap.

We thank the reviewer for this comment. We apologize for the confusion. No colocalization testing was done here. Each vertical panel is a separate tumour section individually stained with either an anti-VSV antibody or anti-HER2 trastuzumab, to assess HER2T expression in tumours infected with VSVΔ51-HER2T. We did use consecutive sections to ensure spatial adjacency, but we did not multiplex the stains for VSV and HER2. We have modified the "Merge" panels to clarify that these are merges of the DAPI/Trast or DAPI/VSV channels, and not merges of VSV/Trast.

2. The article could be supplemented with a detailed histological analysis of tumors after injection to characterize the type of cell death of tumor cells. This is especially true for the most significant in vivo experiments about the combination of two viral strains.

We thank the reviewer for this suggestion. We have performed histological analyses of CT26 tumours post-injection with our virus combination and monotherapy controls. To evaluate cell death, we assessed levels of cleaved caspase 3 as well as cleaved caspase 1. This data has been added as Extended Data Fig. 12.

3. To obtain a strain of vaccinia virus that, without any special problems for replication efficiency, could express both therapeutic proteins - HER2T and TCE- at different loci, or even in one. The selectivity of the vaccinia virus for tumor cells is very high, there will be complete colocalization, and in addition, with demonstrated effectiveness, there will be no problems with the approval of such an approach for use in clinical practice.

We thank the reviewer for this excellent suggestion. We are currently working on this concept as part of a separate manuscript with novel targets, and for this reason we have decided to keep this work separate.

Reviewer #3

Reviewer #3 (Remarks to the Author): with expertise in oncolytic virotherapy, cancer immunotherapy

In this manuscript the authors seek to deliver a non-signaling HER2T (T for truncated) protein to tumors and a VV to deliver a T Cell Engager (TCE) that will pull in T cells to recognize the HER 2T expressed on the tumor cells and activate a T cell response against the tumors. This is an attractive strategy that extends the concept of OV beyond whether the 'pure oncolysis' or the 'oncolysis leading to T cell priming' strategies. Overall, the experiments are plentiful and do suggest that the combination of novel antigen display along with T cell recruitment can potentially generate therapy which allows for shaping the tumor to the therapy. However, there are some points that I feel could greatly strengthen the story, especially for this Journal.

Conceptual comments to address:

1. Whilst I understand completely that these groups are a world leader in the use of both VSV and VV as oncolytics, if the goal is to express a novel tumor antigen (TA) de novo on tumor cells and then to use that TA to engage incoming T cells, is VSV really the most valuable vector system to do this? Presumably the cells infected by the VSV-HER-2T will also be predominantly killed by the virus (the 'pure oncolysis' dogma) and so de novo display of the TA will be transient. The window then for (non viral) T cells to come into the tumor, see the TA and engage new T cells will be limited – compared to for example display of the TA by a non lytic replicating virus. Some of my comments below reflect this concern which may be why some of the models are both artificial (treatment rapidly after tumor seeding for iv delivery) and some of the in vivo effects are rather limited in efficacy. Perhaps the authors could address this slight disconnect with virus type for delivery of the TA and therapy – if they agree.

We thank the reviewer for this question and comment. We do indeed agree that there is the possibility of some VSV-infected cells to undergo oncolysis and become eliminated, before the T-cell mediated effects of the counterpart VV can become active. Fortunately, this outcome was not enough to prevent the therapeutic efficacy of the dual-virus combination strategy.

We initially selected the combination of VSV and VV because of the synergy between the two viruses. We and others from our group have demonstrated that VV can significantly enhance the ability of VSV to replicate and spread in cancer cells, and more importantly, in tumour tissue (Extended Data Fig. 11). In addition, VSV can spread rapidly within tumour tissue and persist for at least 48 hours post-injection, while VV spreads much slower. For this reason, we utilized VSV to spread the HER2T throughout the tumour tissue as efficiently as possible, because VV is not capable of spreading rapidly in the same timeframe.

Conveniently, T-cell engagers are biologically and functionally effective at picomolar concentrations, meaning the less well-spread VV will be able to secrete enough TCE to elicit a localized and therapeutically relevant effect. If we were to switch the transgenes in the viruses to VV-HER2T + VSV-TCE, we would likely have significantly less tumour tagged with HER2T, and significantly higher levels of secreted TCE with no target to bind to. We have expanded upon this point in our discussion and believe it strengthens our story.

Furthermore, although we have not explored non-lytic viruses for this purpose, we are actively exploring different permutations of oncolytic viruses, but that work is within the scope of a separate manuscript.

2. A major hurdle with TCE or vaccination or OV strategies is achieving high level virus expression in all the tumor cells and/or antigen heterogeneity in tumors being targeted with TCE or T cells. Thus, tumors escape these strategies very efficiently due to heterogeneity of TA expression and/or antigen loss. With the strategy described, there is I believe, very little chance of infecting 100% of the tumor cells and generating T cells responses against all of these cells. So in a sense, this inefficient delivery of the TA to a fraction of the tumor cells may not improve upon the issues of tumor heterogeneity and antigen loss – as perhaps some of the moderate efficacy experiments suggest. This raises the issue of whether there is really endogenous anti-tumor immunity generated.

We thank the reviewer for this comment. We agree that not 100% of the entire tumour can be infected. In fact, we tested this out directly (Extended Data Fig. 15) and report that only ~2% of tumour cells are infected by VSV, increased to 4% when co-injected with VV, by 48 h post-injection of CT26 tumours. In the same experiment, we observed that when tumours are infected with VV, only 0.1% of tumour cells are infected, increased to 0.5-2% when co-injected with VSV. Indeed, this extent of viral infection can not lead to direct oncolysis of the entire tumour. We have performed follow up experiments that reveal this modest infection is sufficient to induce anti-tumour immune responses, as well as anti-HER2T (antigen, exogenous) and anti-viral immune responses (Fig. 5, Extended Data Fig. 13-16).

3. The authors raise the issue of the dependence of this strategy upon co-delivery of two different viruses. This may be less of an issue for s.c. or even ip delivery in a mouse but clinically this seems a very difficult concept to take on board given the hurdles experienced so far in achieving high efficiency in vivo delivery of systemic OV clinically. The TCE presumably does not have to be expressed in the same tumor cells as the VSV-HER-2T but they will have to be co-expressed relatively close. Once again the nature of the iv delivery experiments used in the manuscript suggest that this may be difficult as a translational strategy.

We thank the reviewer for this comment. The two viruses must ideally infect tumour cells in relatively close proximity. TCEs have a short half life, so distal expression and secretion will likely not reach HER2T at a different anatomical site. However we have found an effective method of co-formulating the two viruses to reduce the probability of non-overlapping infection, and we observe overlapping infection by IHC in serial tumour sections (Extended Data Fig. 12).

Specific Comments:

1. Figure 1 - Do the authors have some view on what percentage of the tumor cells that are infected by the VSV-HER-2T that are not killed by oncolysis will then be killed by the T-DM1? This speaks to the nature of any bystander effect – ie is there added value to the T cell engagement of killing cells infected by the OV?

We thank the reviewer for this question. Our group has previously published on the bystander killing mediated by combining VSVΔ51 with microtubule destabilizing agents, including T-DM1, hypothesized to be mediated through TNF-alpha secretion. For this reason, we did not explore this mechanism in this manuscript. We have included this in the discussion. Relevant references below:

- Arulanandam R, Taha Z, Garcia V, Selman M, Chen A, Varette O, Jirovec A, Sutherland K, Macdonald E, Tzelepis F, Birdi H, Alluqmani N, Landry A, Bergeron A, Vanderhyden B, Diallo JS. The strategic combination of trastuzumab emtansine with oncolytic rhabdoviruses leads to therapeutic synergy. Commun Biol. 2020 May 22;3(1):254. doi: 10.1038/s42003-020-0972-7. PMID: 32444806; PMCID: PMC7244474.

- Arulanandam R, Batenchuk C, Varette O, Zakaria C, Garcia V, Forbes NE, Davis C, Krishnan R, Karmacharya R, Cox J, Sinha A, Babawy A, Waite K, Weinstein E, Falls T, Chen A, Hamill J, De Silva N, Conrad DP, Atkins H, Garson K, Ilkow C, Kærn M, Vanderhyden B, Sonenberg N, Alain T, Le Boeuf F, Bell JC, Diallo JS. Microtubule disruption synergizes with oncolytic virotherapy by inhibiting interferon translation and potentiating bystander killing. Nat Commun. 2015 Mar 30;6:6410. doi: 10.1038/ncomms7410. PMID: 25817275.

We have added in Extended Data Fig. 11 to address this point. Briefly, we injected CT26 tumours with VSVΔ51-HER2T (or other treatments as indicated). We harvested tumours at 24 and 48 hours post-infection and performed IHC on fixed tumour sections as well as flow cytometry on dissociated tumours. For flow cytometry, we quantified the percentage of VSVΔ51-HER2T infected tumour cells by using trastuzumab as a primary antibody, which would detect cell-surface HER2+ cells. We report 2-4% of all tumour cells infected by VSVΔ51-HER2T at 24 and 48 hpi, corroborated by IHC data. As trastuzumab was used to detect these cells by flow cytometry, we hypothesize that theoretically all of these cells could bind to and be killed by T-DM1, but the proportion of cells eliminated through bystander killing are expected to be greater.

We have also demonstrated that combining VSVΔ51-HER2T + T-DM1 + VV-αHER2-TCE leads to no added therapeutic benefit compared to VSVΔ51-HER2T + T-DM1 or VSVΔ51-HER2T + VV-αHER2-TCE. This outcome is likely the result of the αHER2-TCE and the T-DM1 sharing an identical epitope and leading to competition at that site. We have chosen not to include that data in this manuscript.

2. In Figure 1g-i., the authors try to show that iv delivery of the dual virus strategy is effective. They use a model in which tumor is seeded i.v. and then viruses are given 1 day later. They count resultant lung nodules at day 10. This is not a model of treatment of established tumors by this strategy. To convince the reader that here is true iv delivery of the two virus strategy to treat tumours the treatment should be given once we are certain that there are seeded (established) tumors in the lungs. This experiment also lacks the control of VSV-GFP or VSV-luciferase which is very important – if, for example, the tumor cells are simply infected by the virus before being established as tumors (making iv delivery much easier than having to penetrate an established tumor) the oncolytic virus alone needs to be tested. The authors show a significant difference between the virus+TCE in F. compared to the virus alone suggesting a real difference. However, is it not likely/possible that the VSV-HER-2T displays the HER-2T TA and that the T-DM1 treatment clears these tumor cells way before the tumor is established?

We apologize for any confusion. To clarify, all 4T1.2 lung metastasis model experiments had a VSVΔ51-HER2T monotherapy condition. This group was present in Fig. 1h-i, and it was also present in Fig. 6c. In fact, in Fig. 6c all statistical tests were relevant to the VSVΔ51-HER2T monotherapy.

To comment on the control virus, we used the VSVΔ51-HER2T virus as a control monotherapy instead of VSVΔ51-Fluc or VSVΔ51-GFP because those two reporter transgenes are not as relevant as the HER2T transgene. We aimed to identify any therapeutic impacts of the VSVΔ51-HER2T, attributable to the virus itself or the potential immunogenicity of the HER2T transgene.

In addition, we performed a set of experiments that validate our 4T1.2 lung metastasis model (Extended Data Fig. 3). Briefly, we injected 4T1.2 cells stained with an infrared dye, by tail vein injection into mice. We imaged these mice by fluorescence IVIS imaging at 1 hour and 24 hours post-injection, and detected the injected cells within the lungs. We also harvested these lungs at 24 hours, fixed, paraffin embedded, and H&E stained the lung sections; a pathologist evaluated these sections and confirmed presence of microscopy 4T1.2 lesions within the lung sections at 24 hours post-injection. This timepoint was selected to coincide with the timing of when we normally begin systemic treatments. These data have been incorporated into Extended Data Figure 3a-e, the results, and also into the materials and methods section. We thank the reviewer again for this comment, as it has helped us strengthen the validation of our models.

3. In my opinion, this expt should be repeated to show iv delivery is efficacious against clearly established tumor, that the VSV-HER-2T virus is better than the control virus, and with large sample sizes to make valid conclusions. This seems important to me to support the authors claim that VSV-HER-2T can be used in combination with an off the shelf therapy already in the clinic.

We thank the reviewer for this suggestion. We have performed a new experiment to address this point. We recognize that the 4T1.2 lung metastases model may not necessarily be representative of natural dissemination of metastatic cells from a primary tumour. To address this issue, we added a new model of metastasis (Extended Data Fig. 17). Here we implanted B16-F10 metastatic melanoma cells in C57BL/6 mice. We surgically excised the tumours 10 days post-implantation (~250 mm^3), giving primary tumours enough time to establish and to disseminate cells. We followed up with two intravenous injections of our treatments at 48 and 72 hours post-surgery (adjuvant setting). We euthanized animals at day 35 post-surgery and assessed lungs and spleens for metastases. We observed no primary tumour recurrence in our combination treated group, as well as no lung or spleen metastases, compared with control groups.

To add, other groups have also assessed the phenotypic and genetic characteristics of this model, and it has been described as a metastasis model that bypasses extravasation. It was found that there are no genetic or phenotypic differences between lung metastases that form as a result of gradual metastases disseminating from a primary tumour *vs* those that form as a result of i.v. tail vein injection. For these reasons, we have implemented this model as it enables us to demonstrate therapeutic efficacy of our strategy following systemic administration. Relevant references are indicated below and have been added to our text.

- Lücke J, Zhang T, Zazara DE, Seeger P, Izbicki JR, Hackert T, Huber S, Giannou AD. Protocol for generating lung and liver metastasis in mice using models that bypass intravasation. STAR Protoc. 2024 Mar 15;5(1):102696. doi: 10.1016/j.xpro.2023.102696. Epub 2024 Jan 18. PMID: 38244200; PMCID: PMC10831314.

- Rashid OM, Nagahashi M, Ramachandran S, Dumur CI, Schaum JC, Yamada A, Aoyagi T, Milstien S, Spiegel S, Takabe K. Is tail vein injection a relevant breast cancer lung metastasis model? J Thorac Dis. 2013 Aug;5(4):385-92. doi: 10.3978/j.issn.2072-1439.2013.06.17. PMID: 23991292; PMCID: PMC3755653.

- Taha Z, Crupi MJF, Alluqmani N, Fareez F, Ng K, Sobh J, Lee E, Chen A, Thomson M, Spinelli MM, Ilkow CS, Bell JC, Arulanandam R, Diallo JS. Syngeneic mouse model of human HER2+ metastatic breast cancer for the evaluation of trastuzumab emtansine combined with oncolytic rhabdovirus. Front Immunol. 2023 Apr 19;14:1181014. doi: 10.3389/fimmu.2023.1181014. PMID: 37153626; PMCID: PMC10154558.

4. Why are there only 5 mice per treatment group? Please provide a full explanation of how this sample size provides enough power to make the statistical conclusions drawn.

We thank the reviewer for this question. All of our experiments involving animals are based on experimental design with a power of 0.8 and an alpha level of 0.05. We use the NIH sample size calculator to estimate and decide on sample size. Statistical tests done are indicated in each figure or in the methods section, but all survival studies were assessed using the log-rank test.

We typically run our animal experiments in increments of 5 mice per group because the animal facility housing limitations are 5 animals per cage. In experiments where the effect size is variable and modest we aim to have a minimum of 10 mice per group. For experiments where we expect to observe a more robust and consistent response we typically use 5 mice per group. We also ensure that all control groups are included in the experiment, which also adds a significant number of animals to the experiment and adds to the overall statistical confidence.

For the 4T1.2 lung metastasis experiments, we started with n = 5 mice per treatment group. The impact of the combination treatment we observed was a ~90%+ reduction in the number of tumour nodules relative to PBS controls (Fig. 1i, Fig. 6c). At this effect size, if we power the experiment to 0.8 then the prescribed sample size

computed would be n ≥ 4 mice per group. These results were also consistent with a tight standard deviation and as such we did not increase the sample size.

For the survival studies performed, rather than increasing the sample size in a single tumour model we instead repeated the experiment in six different tumour models. Although the n-value of each treatment group within a single tumour model is n = 5, we believe that observing consistent outcomes across 6 different tumour models, in 2 different mouse genetic backgrounds (BALB/c and C57BL/6), is convincing in that we observe consistent outcomes of our treatment applicable to multiple tumour models. We also note that these survival experiments were not performed in parallel, rather they were performed sequentially, where each cohort of mice was from a different lot or colony. We therefore believe that these experiments then serve as independent experimental and biological replicates.

5. Figure 2 - 'The αHER2-TCE was further shown to induce target-dependent cytolytic activity by reducing the viability of HER2+ or HER2T+ target cells in the presence of naïve syngeneic splenocytes (Fig. 2h).' These are splenocytes used in these experiments. Could the authors not have used CD3 T cells to prove the activity of the TCE? These mouse tumor cells are expressing a human HER-2 which could stimulate various innate and or xenogeneic T cell responses. Why was the Control TCE (with an anti-human CD3) not used in these experiments?

We have revised the sentence to better represent the data, as follows: "The αHER2-TCE was further shown to induce target-dependent cytolytic activity by reducing the viability of HER2+ or HER2T+ target cells in the presence of naïve syngeneic splenocytes (Fig. 2h), with no impact on HER2- cells."
We could have also used CD3+ T-cells, which would likely have shown more robust results, but we aimed to use an immune population more representative of an *in vivo* setting. Although we did not use the Ctrl-TCE in this specific experiment, we did have 4T1.2 cells that are a HER2-negative control; the aim of this experiment was to demonstrate HER2-dependent cell killing. In subsequent experiments (Fig. 3-6) we used the Ctrl-TCE.

6. Figure 3 - The quality of Fig.3A is very poor at least in my version. Why is GFP examined? J69 cells are human T cells and the MC38 targets are mouse T cells.

The GFP channel was to ensure that the TdTomato signal is a real signal being detected and not autofluorescence. We have removed the GFP panels as we agree they may be confusing. We also used higher resolution images for higher quality.

If the mouse T cells are engineered to express an anti-human CD3 TCE is it clear that this TCE is not enhancing a xenogeneic reactivity against the mouse cells? Would it be good to have the control VSV-GFP in this mix to guard against that possibility?

We thank the reviewer for pointing this out. MC38 colorectal cancer tumour cores were co-cultured with J69 cells *ex vivo*. The images are focused on the J69 cells, and tumour cores are not shown. We have modified the figure legends (Fig. 3a, Extended Data Fig. 8c) for better clarity.

7. Figure 4 C - 'P values indicated next to each treatment group relative to VSVΔ51-HER2T + VV-αHER2-TCE'. My apologies for being unclear. There is a p value of 0.0017 shown against the VSVΔ51-HER2T + VV-αHER2-TCE line. How can this be significant against itself?

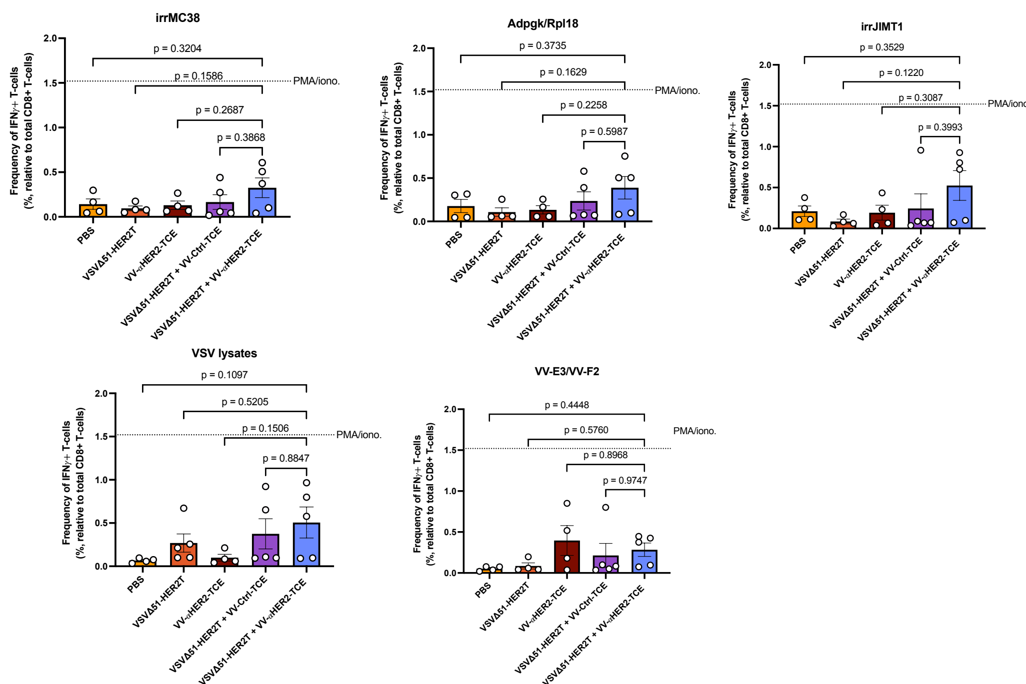
We thank the reviewer for pointing this out. We have changed all p-value annotations in all survival curves to have better clarity.

8. Why are there only 5 mice per treatment group? Please provide a full explanation of how this sample size provides enough power to make the statistical conclusions drawn. How many times was this experiment repeated?

We thank the reviewer for this question. We addressed this point in a prior comment. Briefly, n=5 mice per group is a standard sample size given housing limitations at our animal facility. Additionally, we increase sample size if experimental outcomes are more variable. In these series of experiments we found our results consistent, and instead of increasing the n-value in a single tumour model, we repeated the experiment in 6 different tumour models. These experiments were not conducted in parallel, but sequentially. Each experiment was thus conducted independently, at a different time, using syngeneic mice sourced from different colonies (Jackson Laboratory or Charles River). Each of these experiments involved 5-7 treatment groups when accounting for all control groups, making it challenging to increase sample size for all experiments.

9. Fig.4D,E: Am I correct in the reading of this experiments that 2 mice survived long term the treatment with VSVΔ51-HER2T + VV-αHER2-TCE; these two mice were then re-challenged on the left and right flank with MC38 parental (left flank) and MC38-HER-2T (right flank). The MC38 left flank tumors grew but the right flank MC38-HER-2T tumors were rejected. If I am correct I think the authors could make this a little clearer in the Figure legend. More importantly this shows that the effect of the therapy was to raise endogenous immunity against the human HER-2T and not against the tumor cells themselves. This is contrary to the authors' overall claims that this strategy raises endogenous anti-tumor T cell responses and is rather ignored in the discussion.

We thank the reviewer for their suggestion. We have clarified the rechallenge design in the figure legends. In the rechallenge experiments with the MC38 model we did observe graft rejection of the HER2T tumours only, and not of the parental MC38 tumours. These data might suggest that our treatment strategy triggered potent anti-HER2T immunity, with comparably less potent or less effective anti-tumour immune responses. We speculate this outcome may partly be a result of waning anti-tumour immune responses following the 90-day waiting period before tumour rechallenge. To evaluate this hypothesis, we isolated spleens 14 days post-treatment from mice bearing MC38 tumours and performed ICS following stimulation (see below). This experiment was performed in parallel to the data in Fig. 5h-m. These results were not statistically significant and for this reason we omitted these data from the manuscript, but we did indeed detect responses against both anti-tumour (irrMC38 and MC38 peptides Adpgk/Rpl18) and anti-HER2 (irrJIMT1), as well as viral antigens, at day 14 post-treatment.



10. Fig.4 I&J. Here the same experimental set up is in place and the rejection of the parental CT26 tumors on the left flank is more convincing of a genuine anti-tumor T cell response. However, this is not a good control to show the generation of anti-tumor immunity. When naïve mice are re-challenged with the CT26 cells they do not reject the tumor. When the 2 cured mice are re-challenged they do. However, the cured mice have already seen the CT26 tumors – so it may be that any mouse vaccinated with CT26 (perhaps irradiated vaccine etc) would reject the re-challenge (ie CT26 is an immunogenic tumor). Therefore, rejection of the CT26 by the cured mice does not show that treatment itself generates the anti-tumor immunity; it could be simply that the CT26 is inherently immunogenic and any or no treatment generates anti-tumor immunity. These data using only 2 surviving/re-challenged mice are not statistically enough to make the conclusions the authors draw. I think it would be much more persuasive if the authors could show that treated and cured mice generate real anti-tumor CD8+ T cell responses enhanced by the VSVΔ51-HER2T + VV-αHER2-TCE combination treatment. This could be done much more elegantly by ELISPOT, ICS or even ELISAs against parental tumors or even against known peptide epitopes from these tumor models. In addition, real anti-tumor immunity could be shown by T cell depletion experiments. On balance, I do not think the authors have shown what they claim that immunological memory against the tumor is generated. This is also in the title: '…. or endogenous-antigen agnostic immunotherapy' so I think that precision, statistical relevance and full interpretation of the data is important.

We thank the reviewer for this comment. Our group uses the rechallenge model as standard practice to evaluate immunologic memory, but we recognize the flawed nature of this experiment. To address your comment, we have performed a series of experiments that directly evaluate anti-tumour immune responses (Fig 5h-m, Extended Data Fig. 15-16). Briefly, CT26 tumour-bearing mice were administered i.t. treatment injections as indicated. Serum was collected at day 7 and 14 post treatment, and at day 14 post-treatment splenocytes were harvested and cultured overnight in the presence of different stimuli for both an ELISpot or ICS. Treatment with our dual-virus strategy elicited potent CD8+ T-cell responses against whole tumour cells (irrCT26) and tumour antigens (gp70). We detected responses against the HER2T antigen and viral antigens as well. We moreover detected tumour-reactive (CT26) IgG at both D7 and D14 post-treatment, as well as HER2T-reactive IgG. Combined with the re-challenge data, we believe these new findings more strongly support our conclusion that treatment with our dual-virus combination elicits anti-tumour immune responses.

11. Figure (6) A-C: The same critique of the experimental set up holds here. Virus delivery 1 day after IV injection of tumor cells does not show efficacy against established tumors. I think this should be extended to convince the reader of the efficacy and feasibility of this approach of delivery of two viruses to established tumors.

We have addressed this point in a prior comment. Briefly, we validated that within 24 hours of IV cell injection we detect microscopic lesions within lungs by histological examination, and by IVIS imaging. We moreover utilized another metastasis model, using B16-F10 post-surgical metastases (Extended Data Fig. 17).

12. Fig. (6)D-E. Is it a surprise that the VSV-Her-2T virus has no therapy by itself? The therapy in this ID8 model is very modest – is it T cell regulated in the way that the authors are claiming? There are no immune correlative studies to show this.

We thank the reviewer for their question. This ID8 *Pten (-/-) Tp53(-/-)* tumour model is very aggressive and very challenging to treat using oncolytic virus monotherapy given its refractory nature against VSVΔ51. We observe very similar outcomes with VSVΔ51 monotherapy in a separate experiment (Extended Data Fig. 4a-c) and in other projects using the same ID8 model.

Indeed the therapeutic response in the ID8 model is often less robust than in other models, owing to the disseminated, aggressive, and virus-refractory nature of this tumour model. Nonetheless the therapeutic benefit of the dual-virus treatment led to a statistically significant enhancement in median survival.

Although we have not performed any immune correlative experiments in this ID8 model to validate the mechanism of action of our dual virus combination, we are confident that T-cell responses are responsible, at least in part, for the observed efficacy. Fig. 5e includes 6 controls groups, including 3 virus monotherapy control groups and 2 virus combination control groups. Neither of the control virus combinations (VSVΔ51-

Fluc + VV-αHER2-TCE or VSVΔ51-HER2T + VV-Ctrl-TCE) led to therapeutic efficacy despite having received the same overall viral dose as the efficacious VSVΔ51-HER2T + VV-αHER2-TCE group. It was only the combination with the compatible payloads that led to therapeutic efficacy, where the payloads operate through direct engagement of CD3+ T-cells.

13. Statistics need to be explained and validated throughout the manuscript. For example, in Extended Data Fig.3F. '(f) overall survival was monitored. P-values indicated next to each treatment group relative to VSVΔ51-HER2+T-DM1.' There is a p value of 0.0090 next to the VSVΔ51-HER2+T-DM1.This makes no sense unless I am mis-understanding this.

Thank you for pointing this out. We have changed all p-value annotations in all survival curves to have better clarity.

14. Extended Data Fig. (10)C: I find it very difficult to believe that there is any statistical difference between the VSV-HER2-T + VV-Ctrl group and the VSV-HER2-T+VV-anti-HER2 TCE (co) group. And yet the text reads: To evaluate the extent of tumour control our dual-virus combination exerts, we implemented our approach in disseminated disease models. First, we optimized a co-formulated dose containing both OVs and validated this co-formulation yields similar efficacy to our original sequential treatment regimen (Extended Data Fig. 9a-c). This co-formulation enables more rapid administration within a shorter timeframe, allowing the treatment to fit within the narrow therapeutic window in these disseminated disease models. We moreover demonstrated that switching the order of the OVs in the sequential regimen does not significantly impact efficacy (Extended Data Fig. 9c). I do not believe that these data show therapy of the combination. Therefore they do not show that switching the timing does not impact efficacy. There is no statistical difference between the different timings of the therapy but if the therapy is not significant itself relative to controls then this is not a valid conclusion. I think that throughout the manuscript the authors need to be considerably more rigorous with their statistics, given their sample sizes and their interpretations of the data.

We thank the reviewer for this comment, and we apologize for the lack of detail. We have included the p-values for the survival curve in Extended Data Fig. 10c, which were computed by the Log-rank test.

We have also included the statistical analysis for the tumour volume curve (Extended Data Fig. 10b, d); here the p-values indicate significant stunting of tumour progression by both the seq. or the co. injection strategies (purple or teal curves) at days 21-28. These p-values were computed by two-way ANOVA, and all comparisons are relative to the co. injection group (teal). The colour of the asterisks match the colour of the treatment group being compared. We also included a table that contains the p-values for these comparisons to avoid cluttering the tumour volume curve.

Based on these tumour volume and survival datasets, we conclude that the therapeutic benefit conferred by the dual-virus strategy is not significantly affected if it is administered sequentially or as a single co-injection.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

The authors have made considerable efforts to address the comments of the reviewers. As a result, the manuscript is much clearer and stronger. There do remain some weaknesses - for example, as the other reviewer agreed, the treatment of lung lesions one day after i.v. injection of tumor cells is not a treatment model of established lesions. Even though tumor cells could be detected in the lungs by luminescent imaging, these are likely individual tumor cells, as this time period is far too short for the establishment of macroscopic lesions. However, since the manuscript is generally convincing and already contains a huge amount of data, I would be satisfied if the authors would simply acknowledge the limitation of the model within the text. Otherwise, I am satisfied and recommend the manuscript for publication.

Reviewer #2 (Remarks to the Author):

The article proposes an interesting new approach to treating tumors using oncolytic viruses and anti-HER2 therapy. Despite the potential difficulty of implementing this approach into clinical practice, the concept itself is new and promising. In addition to a tumor-selective strain of vesicular stomatitis virus expressing a truncated form of HER2, a strain of vaccinia virus expressing a HER2-targeted T-cell activator (TCE) was used. The synergistic effect of the two approaches ensures the formation of a systemic antitumor immune response. All experiments were carried out with a substantial and controls were thoughtfully executed.

In my opinion, the authors significantly improved the article and supplemented it with all the necessary materials, responding to all comments.

I look forward to the continuation of research and subsequent implementation of the proposed approach of combined biotherapy into clinical practice

Reviewer #3 (Remarks to the Author):

The authors have responded very fully to al of the Reviewers' comments with both added Discussion and in some cases with new data. I think this is an interesting new pathway for the use/combination of Oncolytic virotherapy and I recommend the manuscript for publication.

POINT-BY-POINT RESPONSE TO REVIEWER COMMENTS:

Reviewer #1

The authors have made considerable efforts to address the comments of the reviewers. As a result, the manuscript is much clearer and stronger. There do remain some weaknesses - for example, as the other reviewer agreed, the treatment of lung lesions one day after i.v. injection of tumor cells is not a treatment model of established lesions. Even though tumor cells could be detected in the lungs by luminescent imaging, these are likely individual tumor cells, as this time period is far too short for the establishment of macroscopic lesions. However, since the manuscript is generally convincing and already contains a huge amount of data, I would be satisfied if the authors would simply acknowledge the limitation of the model within the text. Otherwise, I am satisfied and recommend the manuscript for publication.

We thank the reviewer for their positive outlook towards our revisions. We have also included an additional line in the results section to address the limitation of the i.v. metastasis model as follows:
"The lung metastasis model demonstrated efficacy of our strategy in early-stage disseminated disease, but it does not capture the true progression of metastasis over time. We therefore further evaluated the therapeutic benefit of the VSVΔ51-HER2T+VV-*α*HER2-TCE combination strategy in a post-surgical B16-F10 metastasis model…"

Reviewer #2 (Remarks to the Author):

The article proposes an interesting new approach to treating tumors using oncolytic viruses and anti-HER2 therapy. Despite the potential difficulty of implementing this approach into clinical practice, the concept itself is new and promising. In addition to a tumor-selective strain of vesicular stomatitis virus expressing a truncated form of HER2, a strain of vaccinia virus expressing a HER2-targeted T-cell activator (TCE) was used. The synergistic effect of the two approaches ensures the formation of a systemic antitumor immune response. All experiments were carried out with a substantial and controls were thoughtfully executed.

In my opinion, the authors significantly improved the article and supplemented it with all the necessary materials, responding to all comments.

I look forward to the continuation of research and subsequent implementation of the proposed approach of combined biotherapy into clinical practice

We thank the reviewer for their positive reception of our revisions and we greatly appreciate all the feedback received, which we agree has considerably strengthened our work.

Reviewer #3 (Remarks to the Author):

The authors have responded very fully to all of the Reviewers' comments with both added Discussion and in some cases with new data. I think this is an interesting new pathway for the use/combination of Oncolytic virotherapy and I recommend the manuscript for publication.

We thank the reviewer for their positive reception of our revisions and we greatly appreciate all the feedback received, which we agree has considerably strengthened our work.