nature research

Corresponding author(s): Sanjeev Mariathasan, Shannon J. Turley, Richard Bourgon

☐ Initial submission ☒ Revised version ☐ Final submission

# Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

## Experimental design

1. Sample size

Describe how sample size was determined.

Sample size was constrained by (a) the size of the clinical trial, sample size of which was based on standard power calculations around the primary endpoint of the trial, and (b) the extent to which not all participants had sample suitable for the assays used. That said, we observed many statistically significant effects in the data, indicating that the effective sample size was sufficient for studying the phenomena of interest. Sample size in the mouse studies is based on the number of mice routinely needed to establish statistical significance based on variability within study arms.

2. Data exclusions

Describe any data exclusions.

All data exclusions are described in the methods section. There are only two scenarios of data exclusion (human data): for analyses examining molecular associates with response, patients without defined response were not considered; for analyses of tumor mutation burden (TMB), outlier patients with 0 mutations were excluded (maximum of two patients; this was necessary as TMB was log2 transformed for all analyses). No data points were excluded from the mouse studies.

3. Replication

Describe whether the experimental findings were reliably reproduced.

IMvigor210 trial is a unique data set and no attempt for replication was made; however, relevant papers supporting the clinical findings were cited in our manuscript.
The number of samples for human analyses are specified in the figures. The number of replicates for mouse experiments is stated in the legend of Figure 4, with N specifying the number of biological replicates.
Specifically:
4a: Representative image from multiple control tumors.
4b: Representative image from multiple control tumors.
4c: 6 independent studies - only 2 of them had the anti-TGFb only control, 10 biological replicates for each study.
4d-e: One representative study of 6. 10 biological replicates.
4f-g: Data combined from 3 independent studies, 15 biological replicates in total.
4h-i: Representative images from 1 out of 3 studies.
4j: Data combined from 3 independent studies, 20 biological replicates in total.
4k-m: One experiment, 8 biological replicates (8 mice/tumor model condition).
S13c: One RNAseq dataset, 5 biological replicates (5 mice/tumor model)
S13d: One experiment, 4 biological replicates.
S13e: Data combined from 3 independent studies, 28 biological replicates in total.
S13f: One experiment, 10 biological replicates.
S13g: Data combined from 3 independent studies, 23 biological replicates in total.
S13h: Data combined from 2 independent studies, 10 biological replicates in total.
S13i: Representative images from 1 out of 3 studies.
S13j: One experiment, 10 biological replicates
S13k: One RNAseq dataset, 5 biological replicates (5 mice/tumor model)
S13l: One experiment, 5 biological replicates (5 mice/tumor model).
S13m: Representative image from multiple control tumors

S13n: 2 independent studies - only 1 of them had the anti-TGFb only control, 10 biological replicates for each study.
S13o: One representative study of 2. 10 biological replicates.
S14a-c: Data combined from 3 independent studies, 15 biological replicates in total.
S14d-h: One experiment, 8 biological replicates.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Our analyses were performed on samples from participants of a clinical trial with predetermined in- and exclusion criteria. We show in Extended Data table S9 that the biomarker evaluable population was equivalent in important covariates to the original intent-to-treat population. Grouping of samples in our analyses were based on objective phenotypes, such as response.
For assessing treatment effect in mice, these were grouped using the "Gould-Hanson with Chan modification Grouping Algorithm" as follows: (Step 1) Determine average volume and remove any ungrouped animals based on their deviation from the average; (Step 2) Sort list of animals by alternating descending deviant high and low volumes (H3,L3,H2,L2,H1,L1); (Step 3) Assign to groups based on even or odd number of groups (to create even distribution): (i) Odd # of groups, in straight sequential (1,2,3,4,5,1,2,3,4,5...etc) and (ii) Even # of groups, in snaking-sequential (1,2,3,4,4,3,2,1...etc); (Step 4) All groups are assigned sequentially with the other groups until all animals in that group are assigned, i.e. larger groups will have the least variation in the extra animals)

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Blinding was not relevant as all subjects received the same treatment. All data was analyzed computationally only (no subjective data interpretation).

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

| n/a | Confirmed |
|---|---|
| ☐ | ☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.) |
| ☐ | ☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ A statement indicating how many times each experiment was replicated |
| ☐ | ☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section) |
| ☐ | ☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons |
| ☐ | ☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted |
| ☐ | ☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range) |
| ☐ | ☒ Clearly defined error bars |

*See the web collection on statistics for biologists for further resources and guidance.*

## Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

Analyses were mostly performed using standard software, such as FlowJo for FACS, Matlab/Mathworks for imaging data, GraphPad and the statistical programming language R. Human data were analyzed using functionality of software packages available at Bioconductor. The methods section gives a very detailed description of individual analysis steps for both human and mouse data. In addition, a data and analysis code package will be made available enabling the reproduction of all data presented in the manuscript. This package will also include statistical outputs from GraphPad.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

## Materials and reagents

Policy information about availability of materials

8. Materials availability

| | |
|---|---|
| Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company. | No restrictions on materials. All materials are available for distribution or from commercial sources. |

9. Antibodies

| | |
|---|---|
| Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species). | Anti-human PD-L1 (CD274)SP142 mAb. https://products.springbio.com/products/PD-L1-CD274-SP142-/M442. Also referencend in Powles, T. et al. MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer. Nature 515, 558-562 (2014).<br>anti-human CD8 rabbit monoclonal antibody SP16 (Spring Bioscience; Cat #M3160) Website: http://products.springbio.com//products/CD8-SP16-/M316.<br>The anti-mouse programmed death-ligand 1 (PD-L1 IgG1 monoclonal antibody (Clone 6E11) functional blocking pre-clinical was generated at Genentech by immunization of PD-L1 knockout mice with a PD-L1-Fc fusion protein and cloned onto a murine IgG1 isotype. The anti-TGFb mIgG1 (1D11) functional blocking antibody was generated at Genentech from a stable line using the commercially available sequence. The anti-gp120 mIgG1 antibody used for in vivo studies was also generated by Genentech.<br><br>Antibodies used for flow cytometry: FcR (Clone 2.4G2, 5 µg/ml, BD Biosciences, San Jose, CA); CD45 (BV605, clone 30-F11, BD Biosciences, San Jose, CA); TCRb (PE, clone H57-597, Biolegend, San Diego, CA); CD8 (APC-Cy7, clone 53-6.7, Biolegend, San Diego,CA); CD4 (BV711, clone RM4-5, Biolegend, San Diego, CA), CD25 (BUV395, clone PC61, BD Biosciences, San Jose, CA); GranzymeB (FITC, clone NGZB, eBioscience™, Thermo Fisher Scientific Inc., Waltham, MA); FOXP3 (APC, clone FJK-16s, eBioscience™, Thermo Fisher Scientific Inc., Waltham, MA); pSMAD2/3 (PE, clone O72-670,BD Biosciences, San Jose, CA).<br><br>Antibodies used for IHC: CD3 (clone SP7, Thermo Fisher Scientific, Waltham, MA); pSMAD2 (clone 138D4, Cell Signal Technologies, Danvers, MA).<br><br>Antibodies used for IF: Collagen I (polyclonal ab34710, Abcam, Cambridge, UK); Col III (polyclonal ab7778, Abcam, Cambridge, UK ); Col IV (polyclonal AB756P, Millipore, St. Louis, MO); CD3 (clone 17A2, BD Biosciences, San Jose, CA), PDGRFR-α (polyclonal AF1062, R&D Systems, Minneapolis, MN). |

10. Eukaryotic cell lines

| | |
|---|---|
| a. State the source of each eukaryotic cell line used. | EMT6 and MC38: Original source ATCC and then screened and stored by common cell repository at Genentech. |
| b. Describe the method of cell line authentication used. | EMT6 and MC38 cells were analyzed by RNAseq. |
| c. Report whether the cell lines were tested for mycoplasma contamination. | Cell lines are routinely screened and both EMT6 and MC38 cells used in studies described in this manuscript were negative for mycoplasma. |
| d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use. | N/A |

## Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

As stated in the Methods, 8-10 week old female Balb/c mice from Charles River Laboratories (Hollister, CA) were used. The mice were housed at Genentech in standard rodent micro-isolator cages and were acclimated to study conditions for at least 3 days before tumor cell implantation. Only animals that appeared healthy were used for the studies.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

Full description of the human research participants and the covariate-relevant population characteristics is detailed in the following two publications on the IMvigor210 trial cohorts:
Rosenberg JE, et al. Lancet, 2016;387(10031):1909-1920.
Balar AV, et al. Lancet, 2017;389(10064):67-76.
These papers have been referenced in this manuscript.

nature research

Corresponding author(s): Sanjeev Mariathasan, Shannon J. Turley,Richard Bourgon

☒ Initial submission ☐ Revised version ☐ Final submission

# Flow Cytometry Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

## Data presentation

For all flow cytometry data, confirm that:

☒ 1. The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ 2. The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ 3. All plots are contour plots with outliers or pseudocolor plots.

☒ 4. A numerical value for number of cells or percentage (with statistics) is provided.

## Methodological details

| | |
|---|---|
| 5. Describe the sample preparation. | Tumors were collected, weighted and enzymatically digested using a cocktail of dispase (Life Technologies, Carlsbad, CA), collagenase P and DNaseI (Roche, Penzberg, Germany) for 45 minutes at 37 °C, to obtain a single cell suspension. , cells were first incubated with mouse BD Fc block (clone 2.4G2, 5 µg/ml, BD Biosciences, San Jose, CA) and Live/dead aqua (LIVE/DEAD® Fixable Dead Cell Stain, Invitrogen) for 30 minutes on ice. The cells were then stained with the following antibodies: CD45 (BV605, clone 30-F11, 0.67 µg/ml BD Biosciences, San Jose, CA), TCRb (PE, clone H57-597, 2 µg/ml, Biolegend, San Diego,CA), CD8 (APC-Cy7, clone 53-6.7, 1 µg/ml, Biolegend, San Diego,CA), CD4 (BV711, clone RM4-5, 0.6 µg/ml; Biolegend, San Diego, CA), CD25 (BUV395, clone PC61, 1 µg/ml; BD Biosciences, San Jose, CA) for 30 minutes on ice. Cells were fixed and permeabilized (eBioscience™ Foxp3 / Transcription Factor Staining Buffer Set, Thermo Fisher Scientific Inc., Waltham, MA) to stain for GranzymeB (FITC, clone NGZB, 5 µg/ml, eBioscience™, Thermo Fisher Scientific Inc., Waltham, MA) and FOXP3 (APC, clone FJK-16s, 2 µg/mL; eBioscience™, Thermo Fisher Scientific Inc., Waltham, MA). |
| 6. Identify the instrument used for data collection. | BD Fortessa and BD Symphony |
| 7. Describe the software used to collect and analyze the flow cytometry data. | DIVA (BD) for data collection, FlowJo (LLC) for data analysis |
| 8. Describe the abundance of the relevant cell populations within post-sort fractions. | We did not sort cells |
| 9. Describe the gating strategy used. | First we gated based on physical parameters (SSC-A vs FSC-A), we then excluded doublets (FSC-H vs FSC-A). On the singlets we selected for live cells (LiveDead Aqua negative, Amcyan). We then selected for CD45+ cells and in the CD45+ gate, for TCRb (T cells). In the T cell gate we then selected CD8+ cells. In the CD8 gate we analyzed the Mean fluorescence intensity of Granzyme B the percentage of Granzyme B + cells. |

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information. ☐