

TGF β attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells

Sanjeev Mariathasan^{1*}, Shannon J. Turley^{1*}, Dorothee Nickles^{1*}, Alessandra Castiglioni¹, Kobe Yuen¹, Yulei Wang¹, Edward E. Kadel III¹, Hartmut Koeppen¹, Jillian L. Astarita¹, Rafael Cubas¹, Suchit Jhunjhunwala¹, Romain Banchereau¹, Yagai Yang¹, Yinghui Guan¹, Cecile Chalouni¹, James Zhai¹, Yasin Şenbabaoğlu¹, Stephen Santoro¹, Daniel Sheinson¹, Jeffrey Hung¹, Jennifer M. Giltmire¹, Andrew A. Pierce¹, Kathryn Mesh¹, Steve Lianoglou¹, Johannes Rieger¹, Richard A. D. Carano¹, Pontus Eriksson², Mattias Höglund², Loan Somarriba³, Daniel L. Halligan³, Michiel S. van der Heijden⁴, Yohann Loriot⁵, Jonathan E. Rosenberg⁶, Lawrence Fong⁷, Ira Mellman¹, Daniel S. Chen¹, Marjorie Green¹, Christina Derleth¹, Gregg D. Fine¹, Priti S. Hegde¹, Richard Bourgon¹ & Thomas Powles⁸

Therapeutic antibodies that block the programmed death-1 (PD-1)–programmed death-ligand 1 (PD-L1) pathway can induce robust and durable responses in patients with various cancers, including metastatic urothelial cancer^{1–5}. However, these responses only occur in a subset of patients. Elucidating the determinants of response and resistance is key to improving outcomes and developing new treatment strategies. Here we examined tumours from a large cohort of patients with metastatic urothelial cancer who were treated with an anti-PD-L1 agent (atezolizumab) and identified major determinants of clinical outcome. Response to treatment was associated with CD8⁺ T-effector cell phenotype and, to an even greater extent, high neoantigen or tumour mutation burden. Lack of response was associated with a signature of transforming growth factor β (TGF β) signalling in fibroblasts. This occurred particularly in patients with tumours, which showed exclusion of CD8⁺ T cells from the tumour parenchyma that were instead found in the fibroblast- and collagen-rich peritumoural stroma; a common phenotype among patients with metastatic urothelial cancer. Using a mouse model that recapitulates this immune-excluded phenotype, we found that therapeutic co-administration of TGF β -blocking and anti-PD-L1 antibodies reduced TGF β signalling in stromal cells, facilitated T-cell penetration into the centre of tumours, and provoked vigorous anti-tumour immunity and tumour regression. Integration of these three independent biological features provides the best basis for understanding patient outcome in this setting and suggests that TGF β shapes the tumour microenvironment to restrain anti-tumour immunity by restricting T-cell infiltration.

Pre-treatment tumour samples from a large phase 2 trial (IMvigor210) investigating the clinical activity of PD-L1 blockade with atezolizumab in metastatic urothelial cancer (mUC) were used for an integrated evaluation of biomarkers (Extended Data Fig. 1a; Supplementary Discussion). We categorized patients who responded to treatment with a complete or partial response as responders and compared them with non-responders, who displayed stable or progressive disease. As previously reported^{2,4}, PD-L1 expression on immune cells (more than 5% of immune cells in mUC are detected by the PD-L1 antibody SP142) was significantly associated with response (Fig. 1a). By contrast, PD-L1 expression on tumour cells was not associated with response (Extended Data Fig. 1b). We next performed transcriptome RNA sequencing (RNA-seq) in 298 tissue samples and assessed correlations with PD-L1 expression on immune cells and response. A gene

set associated with CD8⁺ T-effector (T_{eff}) cells^{4,5} was highly correlated with PD-L1 expression on immune cells (Extended Data Fig. 1c, d; Supplementary Discussion). This gene set was also significantly associated with response, particularly with complete response, and with overall survival (Fig. 1b, c).

mUC is characterized by one of the highest somatic mutation rates^{6,7}. In mUC, tumour mutation burden (TMB) correlates with response to immune checkpoint inhibitors^{4,5}. We confirmed these findings (Fig. 1d, e) and showed that computationally predicted tumour neoantigen burden behaved similarly (Extended Data Fig. 1e, f), suggesting that the TMB reflects an increased potential for immunogenicity^{8–11}. We next assessed the transcriptional and mutational correlates of TMB in mUC. The pathways that were most significantly associated with TMB were those involved in the cell cycle, DNA replication and DNA damage response (DDR; Extended Data Fig. 1g; Supplementary Table 1). Signatures for these pathways were correlated with *MKI67* and thus with proliferation (Extended Data Fig. 1h). Expression levels of *APOBEC3A* and *APOBEC3B*, which encode two cytidine deaminases that are upregulated in urothelial and other cancers^{12,13}, were also correlated with TMB and response (Extended Data Fig. 1i, j; Supplementary Tables 2, 3). Finally, tumours with one or more mutations in DDR or cell cycle regulator gene sets had significantly higher TMB and response rates (Fig. 1f, g; Extended Data Fig. 1k, l; Supplementary Tables 4, 5).

Next, we investigated features beyond CD8⁺ T-cell immunity and TMB that we associated with response. Gene set enrichment analysis identified the cytokine–cytokine-receptor gene set as the only feature associated with non-response (Extended Data Fig. 2a; Supplementary Table 6). The most significantly associated genes within this pathway were *IFNGR1* and genes involved in the TGF β signalling pathway: *TGFB1*, *ACVR1* and *TGFB2* (Supplementary Table 3). Whereas *IFN γ* is known to have favourable effects on anti-tumour immunity, persistent signalling by this cytokine has been associated with adaptive resistance to checkpoint therapy^{14–17}. We observed significantly enhanced expression of *IFNG* in responders, whereas *IFNGR1* expression was significantly higher in non-responders (Extended Data Fig. 2b, c). TGF β is a pleiotropic cytokine that is associated with poor prognosis in multiple tumour types^{18–20} and is thought to have a protumorigenic role in advanced cancers by promoting immunosuppression, angiogenesis, metastasis, tumour cell epithelial-to-mesenchymal transition (EMT), fibroblast activation and desmoplasia^{19,21–23}. The two top-scoring TGF β pathway genes in our data represent a TGF β

¹Genentech, South San Francisco, California 94080, USA. ²Division of Oncology and Pathology, Department of Clinical Sciences Lund, Lund University, Lund, Skåne 223 81, Sweden. ³Fios Genomics, Edinburgh EH16 4UX, UK. ⁴Netherlands Cancer Institute, 1066 CX Amsterdam, The Netherlands. ⁵Department of Cancer Medicine, Institut Gustave Roussy, University of Paris Sud, 94800 Villejuif, France. ⁶Genitourinary Oncology Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA. ⁷University of California San Francisco, Helen Diller Family Comprehensive Cancer Center, San Francisco, California 94158, USA. ⁸Barts Experimental Cancer Medicine Centre, Barts Cancer Institute, Queen Mary University of London, London EC1M 6BQ, UK.

*These authors contributed equally to this work.

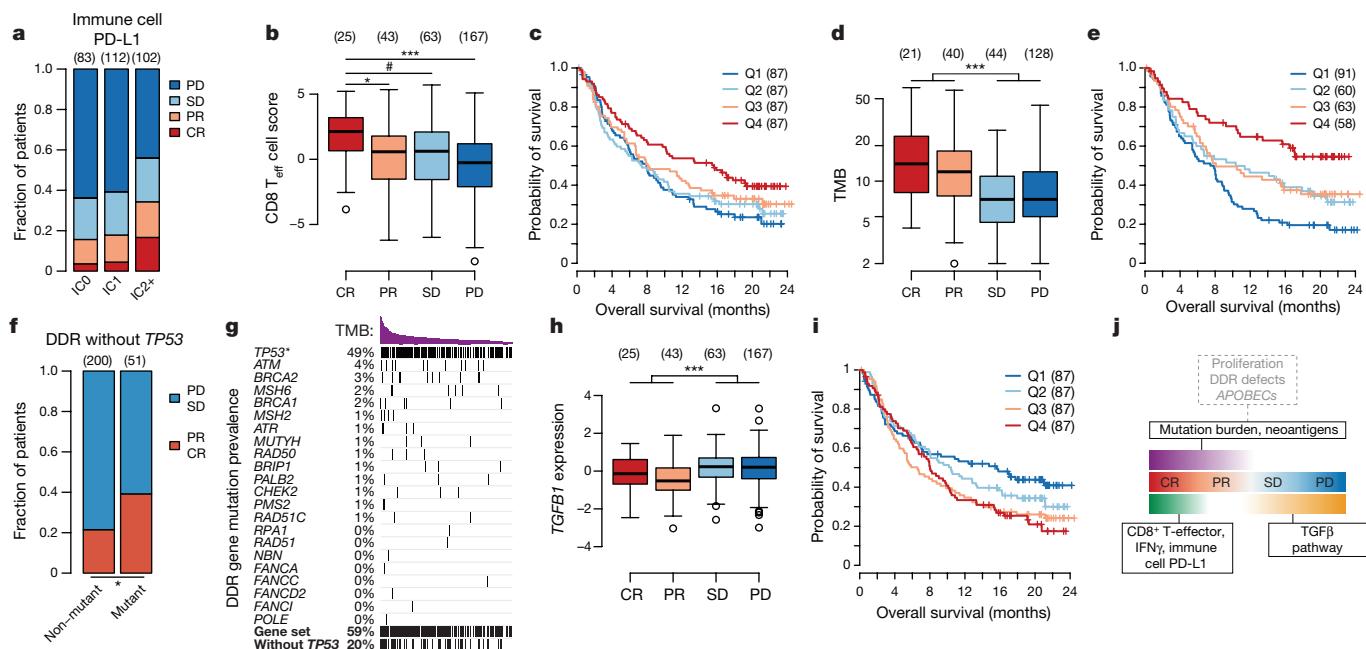


Figure 1 | Three core biological pathways are associated with response to atezolizumab. **a**, PD-L1 on immune cells is associated with response (two-sided Fisher's exact test $P = 0.0038$). IC2/3 tumours had a significantly higher complete response (CR) rate ($P = 0.0006$). PD, progressive disease; SD, stable disease; PR, partial response. PD-L1 expression on immune cells (IC) is assessed by SP142 immunohistochemistry assay and scored as IC0 (<1%), IC1 (≥1% and <5%), or IC2/3 (≥5%). **b**, CD8⁺ T_{eff}-signature score is positively associated with response (two-tailed t -test, $P = 0.0087$), with association driven by the complete response group (CR versus PR, $P = 0.0388$; CR versus SD, $P = 0.0668$; CR versus PD, $P = 0.0003$). **c**, CD8⁺ T_{eff}-signature quartiles are significantly associated with overall survival (likelihood ratio test, $P = 0.0092$). **d**, **e**, TMB is positively associated with

response to atezolizumab (two-tailed t -test, $P = 6.9 \times 10^{-7}$) and overall survival (likelihood ratio test, $P = 2.0 \times 10^{-5}$). A similar plot for tumour neoantigen burden is shown in Extended Data Fig. 1e, f. **f**, **g**, There is a significant association between DDR mutation status and response (**f**; two-sided Fisher's exact test, $P = 0.0117$ excluding *TP53*) and TMB (**g**), both with (two-sided Fisher's exact test, $P = 6.01 \times 10^{-8}$) and without inclusion of *TP53* ($P = 1.95 \times 10^{-5}$). **h**, **i**, *TGFB1* gene expression is significantly associated with non-response (two-tailed t -test, $P = 0.00011$) and reduced overall survival (likelihood ratio test, $P = 0.0096$). **j**, The relationship between response and three core biological pathways. # $P < 0.10$, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Sample sizes are given in parentheses. Q1–Q4, quartiles, lowest to highest. Box plots show median, whiskers represent minimum and maximum excluding outliers, and circles are outliers.

ligand, *TGFB1*, and a TGF β receptor, *TGFB2R2*. Both showed increased mean expression in non-responders and association with reduced overall survival (Fig. 1h, i; Extended Data Fig. 2d, e). Taken together, these results suggest that the impact of checkpoint inhibition on patient outcome in mUC is dictated by three core biological pathways: (i) pre-existing T-cell immunity and (ii) TMB are positively associated with outcome, whereas (iii) TGF β is associated with lack of response and reduced survival (Fig. 1j).

Most human solid tumours exhibit one of three distinct immunological phenotypes: immune inflamed, immune excluded, or immune desert^{1,24}. Previous evidence, mostly from melanomas, has suggested that inflamed tumours are the most responsive to checkpoint blockade^{24,25}, but the relevance of the tumour-immune phenotype to the response of mUCs was previously unknown. In our mUC cohort, a large proportion of tumours (47%) exhibited the excluded phenotype; by contrast, only 26% and 27% of tumours exhibited the inflamed and desert phenotypes, respectively (Fig. 2a, b; Extended Data Fig. 2f). The mean signal for the CD8⁺ T_{eff} signature was lowest in the desert phenotype, intermediate in the excluded phenotype and highest in inflamed tumours (Fig. 2c), and was significantly associated with response only in inflamed tumours, consistent with a model in which the absence of CD8⁺ T cells, or their spatial separation from tumour cells, makes the signature irrelevant.

The observed proximity of CD8⁺ T cells to desmoplastic stroma in immune-excluded tumours (Fig. 2a; Extended Data Fig. 2f) suggests that the relevance of TGF β in this phenotype may be driven by its impact on stromal cells. To measure TGF β pathway activity specifically in fibroblasts, we generated a pan-fibroblast TGF β response signature (F-TBRS). Average expression of this signature was low in immune deserts but significantly higher in inflamed and excluded tumours.

Consistent with a role for TGF β pathway activation in fibroblasts in the tumour microenvironment (TME), the F-TBRS was significantly associated with non-response only in excluded tumours (Fig. 2d). TMB was significantly associated with response in both excluded and inflamed tumours (Fig. 2e).

To better understand how the three core pathways relate to one another and to measure the relative strengths of their association with outcome, we performed a statistical analysis of competing models. Logistic regression pseudo- R^2 was used as a measure of 'explained variance' in patient response²⁶.

In immune desert tumours, the pathways showed negligible explanatory power (Fig. 2f). For excluded tumours, both the F-TBRS signature and TMB were significantly associated with response, and combining the two provided a significant improvement over either term in isolation. By contrast, in inflamed tumours, the CD8⁺ T_{eff} signature combined with TMB exhibited the strongest correlation with response. In an analysis in which all the samples were combined, a model incorporating the three core pathways significantly improved on all single- and two-pathway models (Fig. 2f; Extended Data Fig. 2g), demonstrating that the information provided by each pathway is at least partially independent.

The Cancer Genome Atlas (TCGA) molecular-subtype taxonomy¹² has been widely investigated with respect to mUC, but with inconsistent results. Here we contrast the TCGA taxonomy with the Lund taxonomy, which includes a genetically unstable subtype^{27,28} (Extended Data Fig. 3). Consistent with the importance of TMB, we observed that the genetically unstable subtype was significantly enriched for response (Fig. 3). This effect could not, however, be attributed to TMB alone, as the TCGA luminal II subtype was similarly enriched for high-TMB tumours (Extended Data Fig. 4a, b). Instead, we identified the source

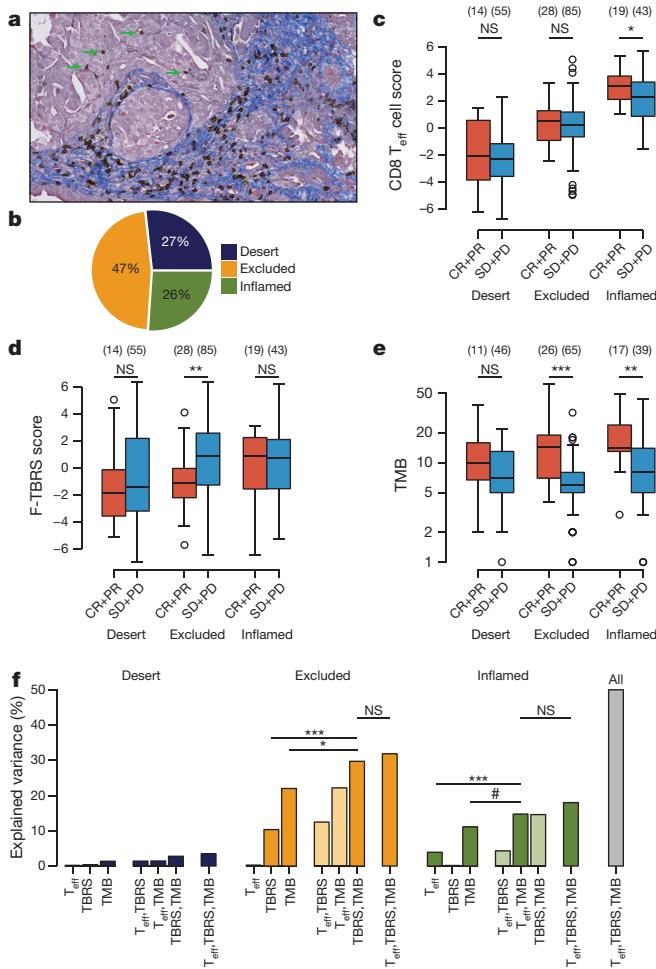


Figure 2 | TGFβ is associated with lack of response in the excluded tumour-immune phenotype. **a**, Combined CD8 immunohistochemistry-trichrome stain. CD8⁺ T cells (stained with 3,3'-diaminobenzidine (DAB), brown) are primarily within collagenous stroma (blue). Rare CD8⁺ T cells are interspersed between tumour cells (green arrows). **b**, Distribution of tumour-immune phenotypes in IMvigor210. **c**, CD8⁺ Teff score is significantly associated with response only in the inflamed phenotype (two-tailed *t*-test, $P = 0.042$). **d**, F-TBRS is significantly associated with response only in the excluded phenotype (two-tailed *t*-test, $P = 0.0066$). **e**, TMB is significantly associated with response in the excluded and inflamed phenotypes (two-tailed *t*-test, $P = 0.00009$ and 0.00258 , respectively). **f**, Explained variance in patient response for generalized linear models fit using single core pathways or pathway combinations. In immune desert tumours ($n = 57$), core pathways showed negligible explanatory power. For excluded tumours ($n = 91$), F-TBRS and TMB were associated with response as single terms; combining the two provided statistically significant improvement over single terms (likelihood ratio, $P = 6.23 \times 10^{-5}$ and 0.02125 , respectively). For inflamed tumours ($n = 56$), CD8⁺ Teff and TMB were associated with response as single terms; combining the two provided statistically significant improvement over single terms (likelihood ratio, $P = 0.00099$ and 0.0557 , respectively). Standard forward selection applied to all patients (grey) identified a three-pathway model as a significantly better fit than all single- or two-pathway models (Extended Data Fig. 2g). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. NS, nonsignificant. For box plots, centre mark is median, whiskers are minimum/maximum excluding outliers, and circles are outliers. Sample sizes are given in parentheses.

of the difference by separately evaluating patients classified as TCGA luminal II only, Lund genetically unstable only, or both (Extended Data Fig. 4c). Among these groups, genetically unstable-only tumours had the lowest CD8⁺ Teff signature scores and lowest TMB, but responded favourably; by contrast, luminal II-only tumours showed sharply increased F-TBRS scores and responded poorly (Extended Data

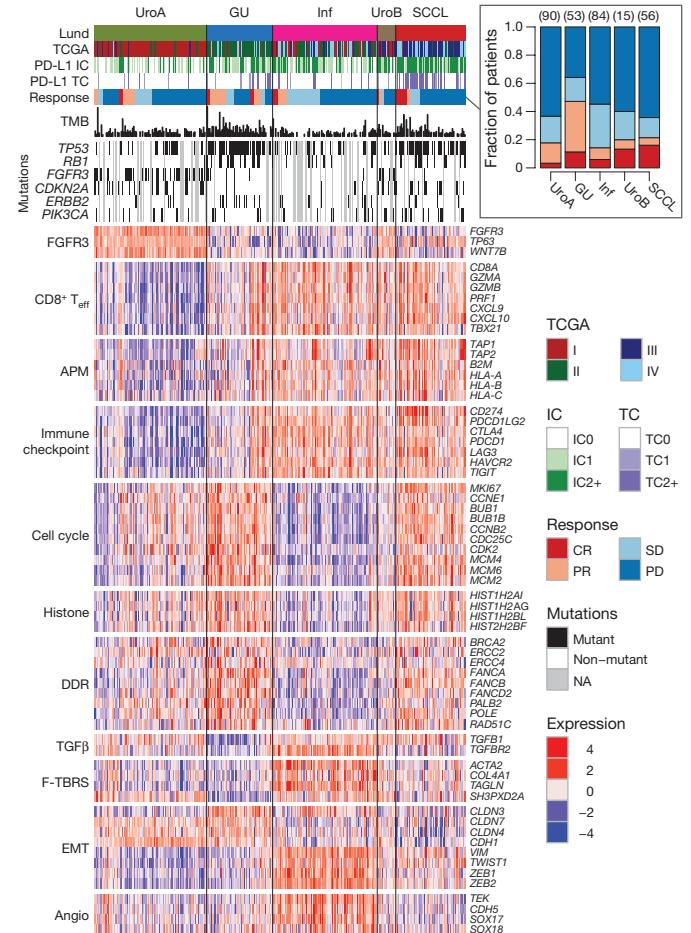


Figure 3 | Relationship between molecular subtypes (Lund scheme) and core biological pathways. Rows of the heat map show gene expression (Z scores) grouped by pathway. Top right, response versus Lund molecular subtype, showing that the genetically unstable (GU) subtype had a significantly higher response rate (two-sided Fisher's exact test, $P = 1.6 \times 10^{-5}$). Sample sizes are given in parentheses. Inf, infiltrated; SCCL, basal/SCC-like; TC, tumour cells; UroA, urothelial-like A; UroB, urothelial-like B.

Fig. 4d, e). These results demonstrate the importance of interactions among the three core pathways in determining response. The Lund subtypes are discussed further in the Supplementary Discussion.

To test our hypothesis that physical exclusion of T cells by the stromal barrier limits response to atezolizumab in immune-excluded tumours, we studied the EMT6 mouse mammary carcinoma model to determine whether there is a role for TGFβ-activated stroma in this context. The EMT6 model exhibits the immune-excluded phenotype (Extended Data Fig. 5a-d) and expresses all TGFβ isoforms as well as PD-L1 (Extended Data Fig. 5e, f). Although therapeutic blockade of PD-L1 or TGFβ alone had little or no effect, mice treated with antibodies against both PD-L1 and TGFβ exhibited a significant reduction in tumour burden (Fig. 4a, b; Extended Data Fig. 5g, h). Regression of EMT6 tumours in these studies was completely dependent on CD8⁺ T cells (Extended Data Fig. 5i). These findings were reproduced in a second relevant mouse tumour model, MC38 (Extended Data Fig. 5j-n).

Combined antibody blockade also led to a significant increase in tumour-infiltrating T cells (Extended Data Fig. 6a), particularly CD8⁺ Teff cells (Fig. 4c, d; Extended Data Fig. 6b). Blockade of TGFβ, alone or in combination with anti-PD-L1, had no effect on the number of CD4⁺ T-regulatory (T_{reg}) cells in the tumour (Extended Data Fig. 6c-e). RNA-seq data revealed that the CD8⁺ Teff signature was increased in mouse tumours treated with a combination of anti-PD-L1 and anti-TGFβ (Fig. 4e). Quantitative histopathology demonstrated that T cell

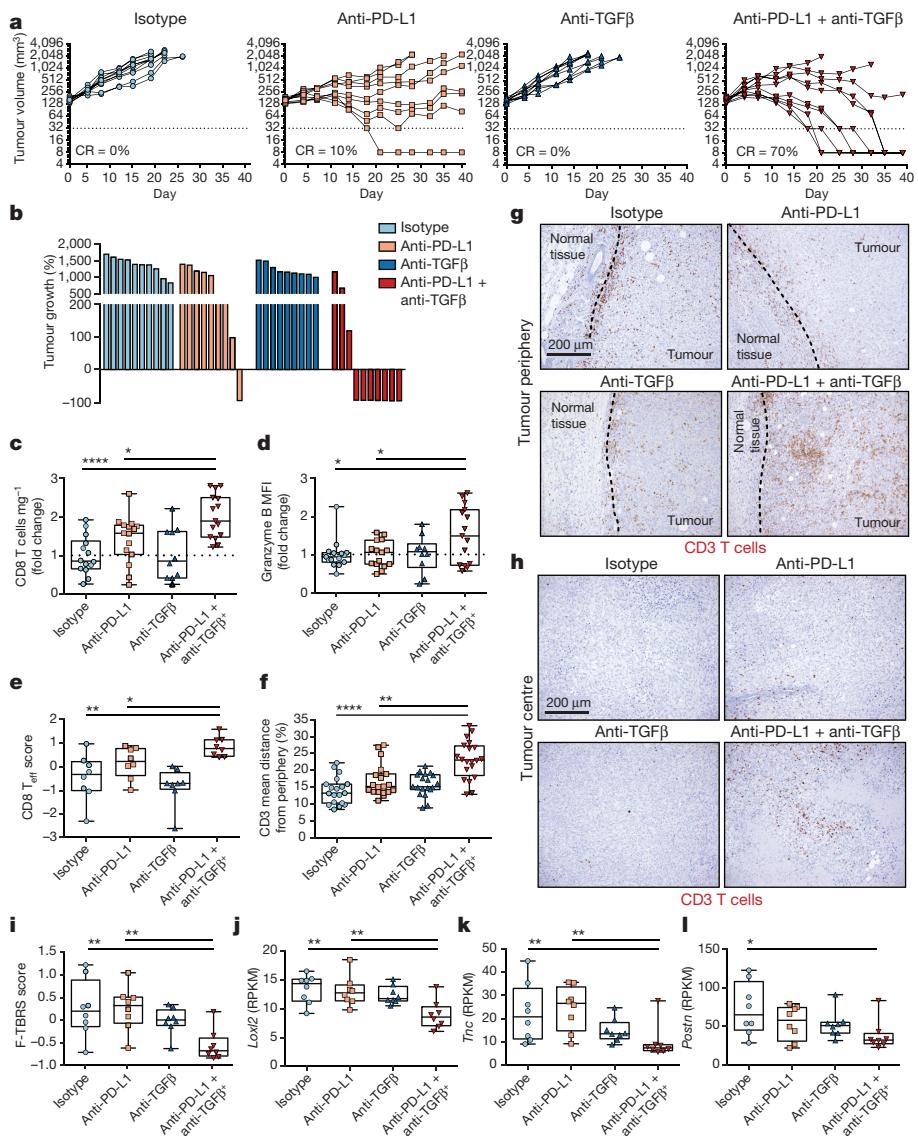


Figure 4 | Tumour regression and changes in TME following therapeutic anti-TGF β and anti-PD-L1 treatment in EMT6 tumours.

a, Tumour growth curves. **b**, Change in tumour volume compared to baseline. Six independent experiments, 10 mice per group. **c, d**, Fold change relative to the isotype mean in total CD8 T-cell abundance (cells mg^{-1} , **c**) and CD8 $^{+}$ T-cell GRZB mean fluorescence intensity (MFI, **d**) by flow cytometry. Three independent experiments, $n=15$ mice for all groups except anti-TGF β , where $n=10$ mice. **e**, CD8 $^{+}$ T_{eff} signature. One experiment, $n=8$ mice per group. **f**, Quantification of tumour-infiltrating lymphocyte (TIL) localization by immunohistochemistry. Three independent experiments, $n=19$ mice for all groups except anti-PD-L1/

distribution significantly changed following combination therapy, with mean distance from the stromal border increasing, and mean distance from the tumour centre decreasing. However, T-cell localization did not change with either single antibody treatment (Fig. 4f–h; Extended Data Fig. 6f). Together, these results suggest that TGF β inhibition potentiated the ability of anti-PD-L1 to enhance anti-tumour immunity, resulting in optimal T cell positioning and ensuing tumour regression.

Anti-TGF β treatment significantly reduced TGF β receptor signalling (as shown by the reduction of phosphorylated (p)SMAD2/3) in EMT6 tumours, particularly in non-immune cells (Extended Data Fig. 6g, h). Given that TGF β is associated with fibroblast differentiation and EMT²², we investigated whether the benefits of dual-antibody blockade

anti-TGF β , where $n=20$ mice; Tukey's honest significant difference multiple-comparison adjustment. **g, h**, Representative CD3 staining of tumour periphery (**g**) and centre (**h**). Scale bar, 200 μm . **i–l**, F-TBRS signature (**i**) and expression of fibroblast genes (**j–l**). RPKM, reads per kilobase of transcript per million mapped reads. One experiment, $n=8$ mice per group. All data shown in **c–l** are from tumours collected at day seven after treatment initiation. For box plots, centre mark is median, and whiskers are minimum/maximum. All P values are based on two-sided Mann–Whitney U -tests unless otherwise indicated. * $P<0.05$, ** $P<0.01$, *** $P<0.0001$.

could be attributed to direct effects on tumour cells or effects on stromal compartments. While single-agent inhibition of TGF β reduced one of the three EMT signatures that we considered, dual-antibody treatment had no significant impact (Extended Data Fig. 6i). By contrast, the F-TBRS signature and expression of canonical fibroblast genes associated with matrix remodelling were significantly reduced in the combined treatment (Fig. 4i–l). Consistent with the phospho-flow analysis that showed no change in pSMAD2/3 in haematopoietic cells, no reduction was observed in two alternate TBRS signatures associated with T cells or macrophages²⁰ (Extended Data Fig. 6j, k). Blockade of TGF β can thus synergize with anti-PD-L1 in the EMT6 model to reprogram peritumoural stromal fibroblasts and increase CD8 $^{+}$ T_{eff} cell counts in the tumour bed, leading to robust anti-tumour immunity. Interestingly,

although TGF β inhibition might also be expected to reduce the T_{reg} cell population, we did not observe lower T_{reg} numbers as a result of TGF β blockade, which therefore did not appear to contribute to the efficacy of the combined treatment.

The comprehensive evaluation of molecular, cellular and genetic factors associated with response and resistance to checkpoint blockade (atezolizumab) in this large cohort of mUC patients has yielded several important conclusions. Three non-redundant factors were found to contribute: (i) pre-existing immunity, as represented by PD-L1 gene expression on immune cells, IFN γ expression and histological correlates of CD8 $^+$ T_{eff} activity; (ii) TMB, measured directly (Extended Data Fig. 7), but also reflected in signatures of proliferation and DNA damage response; and (iii) TGF β -pathway signalling, reflected by a distinct gene expression signature and by pSMAD2/3. These tightly linked factors and their relationships with each other may partly explain why predicting outcomes from PD-L1 expression alone is challenging. The enrichment of the fibroblast TGF β -response signature in non-responding immune-excluded tumours, combined with results from Lund molecular subtyping and preclinical models showing that co-inhibition of TGF β and PD-L1 converted tumours from an excluded to an inflamed phenotype, support a model in which TGF β signalling may counteract anti-tumour immunity by restricting the movement of T-cells in the TME. The observed multifactorial basis of response to immunotherapy may be applicable to other tumour types beyond mUC. Work in this area, across multiple tumour types and therapies, is still in its infancy, but these results open new avenues for disease-agnostic exploration of the mechanisms underlying responses to, and primary immune escape from, cancer immunotherapy.

Code Availability The source code and processed data used for all analyses presented here have been made available in IMvigor210CoreBiologics, a fully documented software and data package for the R statistical computing environment²⁹. This package is freely available under the Creative Commons 3.0 license and can be downloaded from <http://research-pub.gene.com/IMvigor210CoreBiologics>.

Data Availability All raw sequencing data required for RNA-seq analyses have been deposited to the European Genome-Phenome Archive under accession number EGAS00001002556.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 30 June 2017; accepted 8 January 2018.

Published online 14 February 2018.

- Herbst, R. S. et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature* **515**, 563–567 (2014).
- Powles, T. et al. MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer. *Nature* **515**, 558–562 (2014).
- Bellmunt, J. et al. Pembrolizumab as second-line therapy for advanced urothelial carcinoma. *N. Engl. J. Med.* **376**, 1015–1026 (2017).
- Rosenberg, J. E. et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. *Lancet* **387**, 1909–1920 (2016).
- Balar, A. V. et al. Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: a single-arm, multicentre, phase 2 trial. *Lancet* **389**, 67–76 (2017).
- Alexandrov, L. B. et al. Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
- Chalmers, Z. R. et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med.* **9**, 34 (2017).
- Snyder, A. et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* **371**, 2189–2199 (2014).
- Rizvi, N. A. et al. Cancer immunology: Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348**, 124–128 (2015).
- Van Allen, E. M. et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* **350**, 207–211 (2015).
- Chen, D. S. & Mellman, I. Oncology meets immunology: the cancer-immunity cycle. *Immunity* **39**, 1–10 (2013).

- Cancer Genome Atlas Research Network. Comprehensive molecular characterization of urothelial bladder carcinoma. *Nature* **507**, 315–322 (2014).
- Burns, M. B., Temiz, N. A. & Harris, R. S. Evidence for APOBEC3B mutagenesis in multiple human cancers. *Nat. Genet.* **45**, 977–983 (2013).
- Gao, J. et al. Loss of IFN- γ pathway genes in tumor cells as a mechanism of resistance to anti-CTLA-4 therapy. *Cell* **167**, 397–404 (2016).
- Zaretsky, J. M. et al. Mutations associated with acquired resistance to PD-1 blockade in melanoma. *N. Engl. J. Med.* **375**, 819–829 (2016).
- Ribas, A. Adaptive immune resistance: how cancer protects from immune attack. *Cancer Discov.* **5**, 915–919 (2015).
- Bencic, J. L. et al. Tumor interferon signaling regulates a multigenic resistance program to immune checkpoint blockade. *Cell* **167**, 1540–1554 (2016).
- Lin, R.-L. & Zhao, L.-J. Mechanistic basis and clinical relevance of the role of transforming growth factor- β in cancer. *Cancer Biol. Med.* **12**, 385–393 (2015).
- Massagué, J. TGF β in cancer. *Cell* **134**, 215–230 (2008).
- Calon, A. et al. Stromal gene expression defines poor-prognosis subtypes in colorectal cancer. *Nat. Genet.* **47**, 320–329 (2015).
- Deryck, R. & Zhang, Y. E. SMAD-dependent and SMAD-independent pathways in TGF- β family signalling. *Nature* **425**, 577–584 (2003).
- Morikawa, M., Deryck, R. & Miyazono, K. TGF- β and the TGF- β family: context-dependent roles in cell and tissue physiology. *Cold Spring Harb. Perspect. Biol.* **8**, a021873 (2016).
- Flavell, R. A., Sanjabi, S., Wrzesinski, S. H. & Licona-Limón, P. The polarization of immune cells in the tumour environment by TGF β . *Nat. Rev. Immunol.* **10**, 554–567 (2010).
- Hegde, P. S., Karanikas, V. & Evers, S. The where, the when, and the how of immune monitoring for cancer immunotherapies in the era of checkpoint inhibition. *Clin. Cancer Res.* **22**, 1865–1874 (2016).
- Tume, P. C. et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* **515**, 568–571 (2014).
- Dobson, A. J. & Barnett, A. G. An Introduction to Generalized Linear Models 3rd edn. (Chapman & Hall/CRC, 2008).
- Sjödahl, G. et al. A molecular taxonomy for urothelial carcinoma. *Clin. Cancer Res.* **18**, 3377–3386 (2012).
- Sjödahl, G. et al. Toward a molecular pathologic classification of urothelial carcinoma. *Am. J. Pathol.* **183**, 681–691 (2013).
- R Core Team. *R: A Language and Environment for Statistical Computing* <http://www.r-project.org> (R Foundation for Statistical Computing, Vienna, Austria, 2016).
- Rusinova, I. et al. Interferome v2.0: an updated database of annotated interferon-regulated genes. *Nucleic Acids Res.* **41**, D1040–D1046 (2013).
- Spranger, S., Bao, R. & Gajewski, T. F. Melanoma-intrinsic β -catenin signalling prevents anti-tumour immunity. *Nature* **523**, 231–235 (2015).
- Damrauer, J. S. et al. Intrinsic subtypes of high-grade bladder cancer reflect the hallmarks of breast cancer biology. *Proc. Natl. Acad. Sci. USA* **111**, 3110–3115 (2014).
- Hugo, W. et al. Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. *Cell* **165**, 35–44 (2016).
- Hedegaard, J. et al. Comprehensive Transcriptional Analysis of Early-Stage Urothelial Carcinoma. *Cancer Cell* **30**, 27–42 (2016).

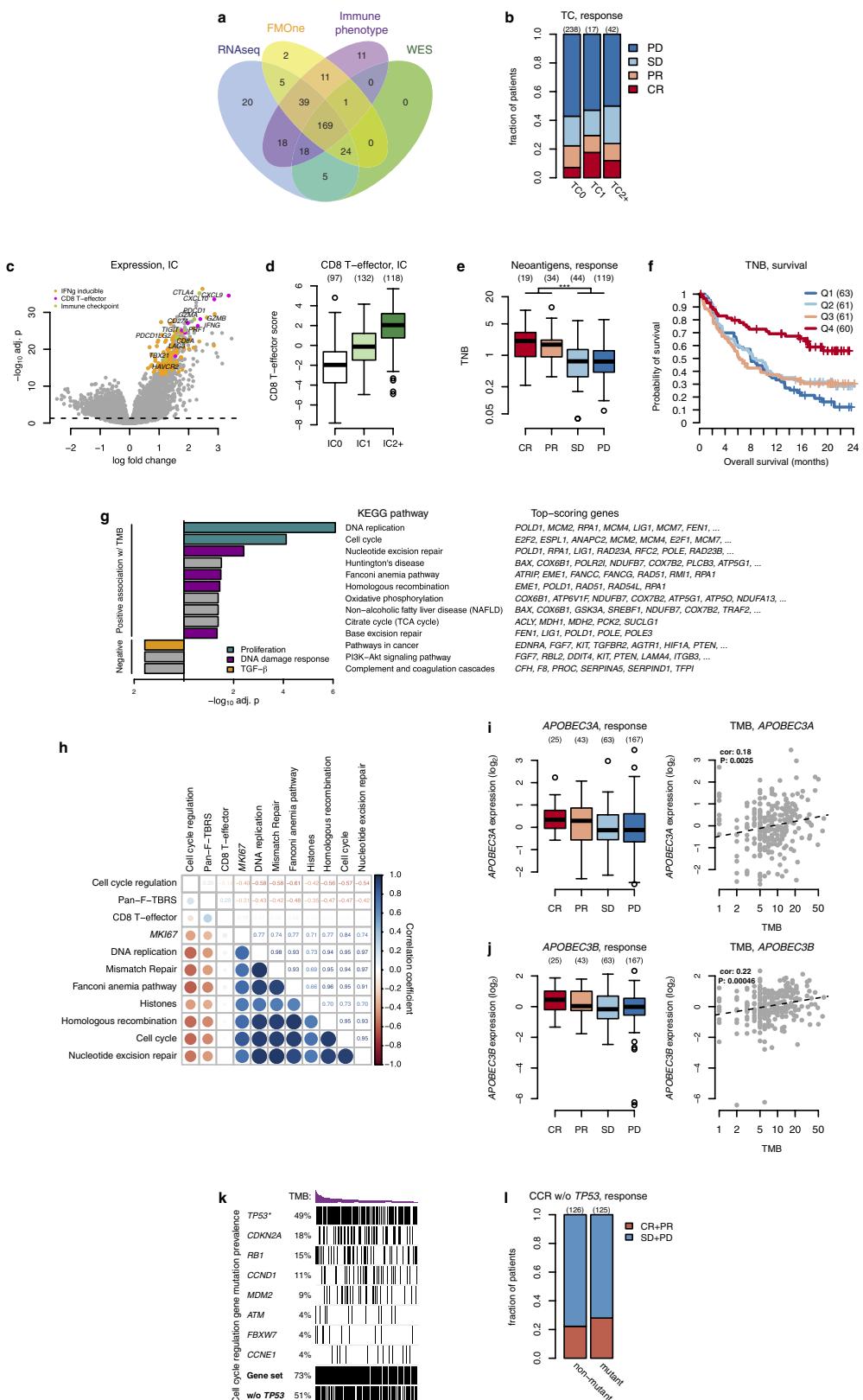
Supplementary Information is available in the online version of the paper.

Acknowledgements We thank the patients and their families; all of the investigators and their staff involved in IMvigor210 study; C. Ahearn, S. Lau, C. Havnar, Z. Boyd, S. Sampath, D. Wilson, J. Doss and medical writers at Health Interactions. J.E.R. acknowledges support from P30 CA008748. L.F. acknowledges support from NCI 1R01CA194511.

Author Contributions S.M., S.J.T., D.N., L.F., I.M., D.S.C., J.E.R., Y.L., M.S.H., M.G., C.D., G.D.F., P.S.H., R.Bo and T.P. contributed to the overall study design. S.M., S.J.T., D.N., Y.W., E.E.K., K.Y., R.Ba., Y.G., Y.S., S.J., S.L., P.E., M.H., L.S., D.L.H., P.S.H. and R.Bo performed the biomarker and statistical analyses. H.I.K., C.C., J.Z., S.S., D.S., J.H., J.M.G., A.A.P. and K.M. conducted microscopy studies. S.J.T., A.C., J.L.A. and R.C. designed all the preclinical experiments and S.J.T., A.C., J.L.A., R.C., Y.Y., C.C., J.Z., Y.S., S.S., D.S., J.H., J.M.G., A.A.P., K.M., J.R. and R.A.D.C. analysed the corresponding preclinical data. S.M., S.J.T., D.N., I.M., P.S.H., R.Bo and T.P. wrote the paper. All authors contributed to data interpretation, discussion of results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to S.M. (mariathasan.sanjeev@gene.com); S.T. (turley.shannon@gene.com) or R.Bo. (bourgon.richard@gene.com).

Reviewer Information *Nature* thanks D. McConkey and the other anonymous reviewer(s) for their contribution to the peer review of this work.

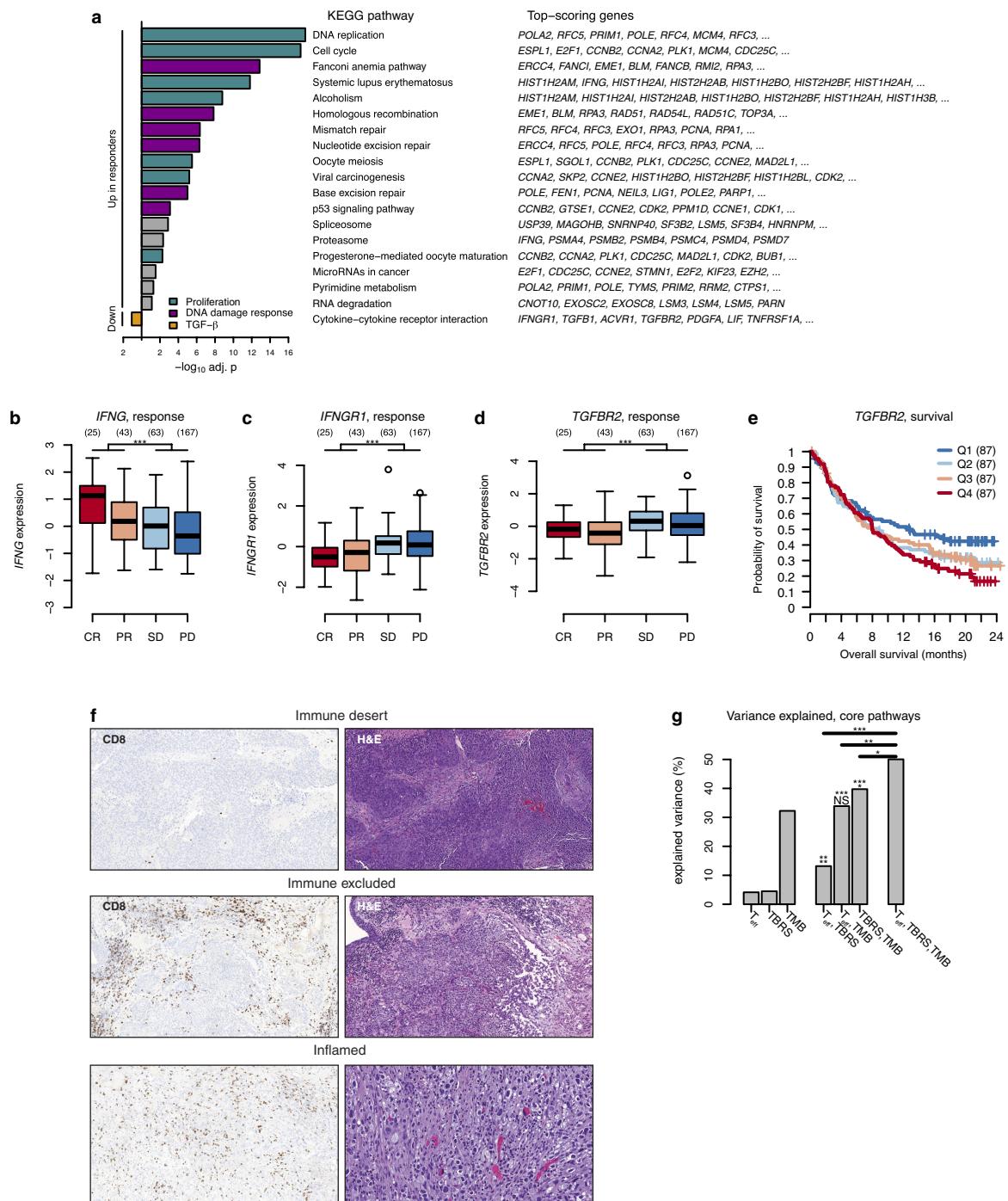


Extended Data Figure 1 | See next page for caption.

Extended Data Figure 1 | Molecular correlates of outcome and TMB.

a, Overlap of the efficacy-evaluable patient populations with assays used in this study ($n = 326$ for one or more of these assays). For gene expression analyses with respect to response, the complete RNA-seq dataset was used ($n = 298$). For gene expression analyses in the context of TMB or immune phenotype, the intersect between RNA-seq and FoundationOne (FMO, $n = 237$) or immune phenotype ($n = 244$) were used, respectively. For mutation analysis around immune phenotypes, the intersect between FoundationOne and immune phenotype was used ($n = 220$). For associations between response or genes mutation status with TMB, the complete FoundationOne dataset was used ($n = 251$). **b**, PD-L1 protein expression on tumour cells (TC), in contrast to expression on immune cells (Fig. 1a), was not associated with response to atezolizumab (two-sided Fisher's exact test, $P = 0.72$). **c, d**, Transcriptional correlates of PD-L1 protein expression on immune cells. **c**, Genes associated with PD-L1 immunohistochemistry positivity on immune cells. Normalized log₂-transformed gene expression was compared with PD-L1 protein expression on immune cells. Interferon-stimulated genes³⁰ and previously reported CD8 T_{eff} and immune checkpoint-molecule gene sets^{4,5} were among the most upregulated (complete list of associated genes is given in Supplementary Table 10). **d**, Association between CD8 T_{eff}-signature score and PD-L1-staining on tumour-infiltrating immune cells. There is a significant positive relationship between the signature score and PD-L1 staining on immune cells (likelihood ratio test, $P = 4.2 \times 10^{-35}$). **e, f**, Tumour neoantigen burden (TNB) is associated with outcome. **e**, Box plots showing the relationship between response status and TNB. Shown are the number of mutations based on whole-exome sequencing, filtering for those mutations that are predicted to be expressed neoantigens. TNB is positively associated with response to atezolizumab (two-tailed t -test, $P = 2.7 \times 10^{-9}$). **f**, TNB, split into quartiles, is also associated with overall survival (likelihood ratio test, $P = 0.00015$). **g**, KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways enriched in genes whose expression is correlated with TMB. Shown are adjusted $-\log_{10} P$ values for enrichment of KEGG gene sets that are significantly (false discovery

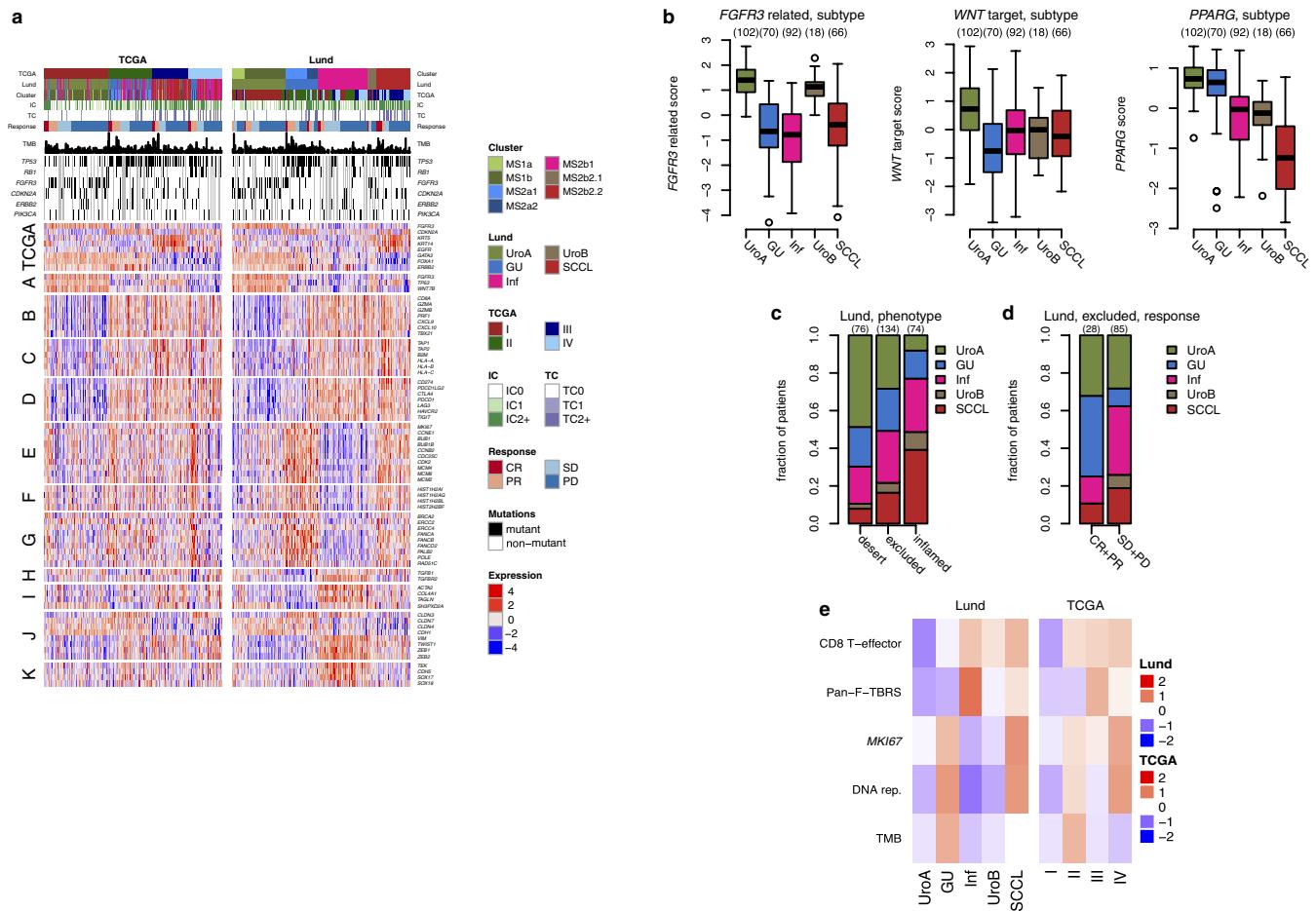
rate < 0.05) enriched in genes that are correlated with TMB (272 samples analysed). Gene sets that are inferred to reflect key underlying biological processes are coloured. Only the top seven genes per set (ranked by single-gene P value) are shown. *ATP5G1* is also known as *ATP5MC1*, *ATP5O* is also known as *ATP5PO*. **h**, Relationship between different gene-expression signatures as well as the single-gene expression values for *MKI67*, a marker for proliferation. Dot size and colour show correlation between signature scores and gene expression (348 samples analysed). Numbers are the Pearson correlation coefficients. Gene-set membership is shown in Supplementary Table 8. **i, j**, *APOBEC3A* and *APOBEC3B* gene expression and their association with response and TMB. Both *APOBEC3A* (two-tailed t -test, $P = 0.015$; **i**) and *APOBEC3B* (two-tailed t -test, $P = 0.0025$; **j**) exhibit higher mean expression in responders. For TMB, Pearson correlation coefficients and P values are given. In **j**, the two extreme expression outliers were excluded when calculating correlation between gene expression and TMB. **k**, Mutations in cell-cycle-regulator genes are associated with TMB. Rows represent genes and columns represent patients ($n = 293$); patients with a mutation are indicated by a black rectangle. The top bar plot depicts TMB in each patient. The final rows represent the mutation status of the pathway with or without *TP53*. Percentages on the left of the plot indicate prevalence. Genes with significant single-gene associations with TMB are marked by an asterisk. Mutations in cell-cycle-regulator genes are associated with TMB with inclusion of *TP53* (two-tailed t -test, $P = 4.01 \times 10^{-8}$), but not without inclusion of *TP53* (two-tailed t -test, $P = 0.0652156$; Supplementary Table 4). **l**, Mutation status in the cell-cycle-regulation (CCR) pathway by response. For each patient, it is determined whether they have any mutations in genes belonging to the CCR pathway, excluding *TP53*. Excluding *TP53*, there is no association between mutation status in the CCR pathway and response (two-sided Fisher's test, $P = 0.31104$; Supplementary Table 4). Sample sizes are given in parentheses. For box plots, centre mark is median, whiskers are minimum/maximum excluding outliers, and circles are outliers.



Extended Data Figure 2 | See next page for caption.

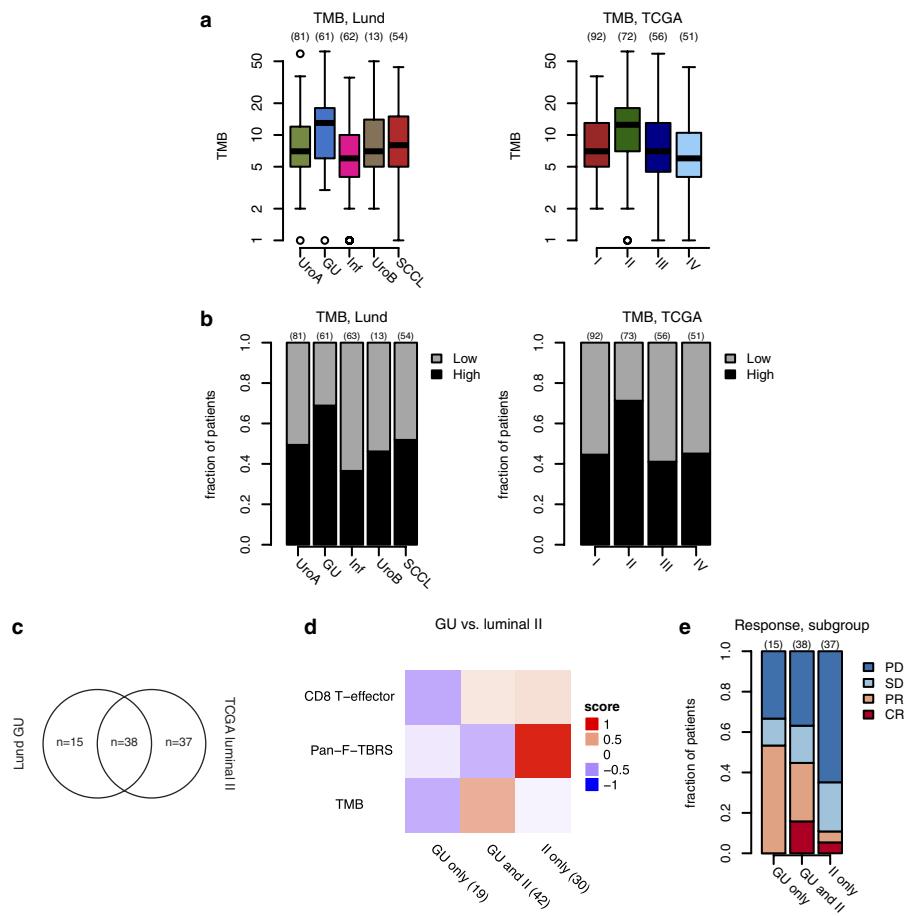
Extended Data Figure 2 | Pathways associated with response and cancer-immune phenotypes. **a**, KEGG pathways are significantly associated with response to atezolizumab (adjusted $P < 0.10$; comparing 68 responders to 230 non-responders). The top seven genes per set are shown; complete lists are given in Supplementary Table 6. *SGOL1* is also known as *SGO1*. **b, c**, *IFNG* (**b**) and *IFNGR1* (**c**) gene expression (*y* axis) are significantly associated with response (two-tailed *t*-test, $P = 9.1 \times 10^{-5}$) and non-response (two-tailed *t*-test, $P = 0.00012$), respectively. **d, e**, *TGFBR2* gene expression (*y* axis) is significantly associated with non-response (two-tailed *t*-test, $P = 0.00019$, **d**) and, when split by quartiles, with reduced overall survival (likelihood ratio test, $P = 0.022$, **e**). In **b–d**, the numbers above the graphs specify sample numbers in each bin. 87 samples per quartile (**e**). **f**, Histology of tumour-immune phenotypes: desert, excluded, and inflamed. **g**, Explained variance in patient response. Generalized linear models were fitted to all efficacy-evaluable, immune-phenotyped samples ($n = 204$) using binary response (complete response/partial response versus stable disease/progressive disease) as the dependent variable and scores from single input or input combinations (*x* axis) as independent variables. Percentage of explained variance of response is plotted on the *y* axis. Comparisons between different

models were made using a likelihood ratio test. Horizontal bars describe likelihood ratio test results for two-biology versus three-biology model comparisons. Stacked significance symbols for two-biology models show results of likelihood ratio test comparison to the first single-biology model and separately to the second single-biology model, in the same order as the *x*-axis bar labels; for example, the T_{eff} , TMB model achieves three asterisks when compared to the T_{eff} single model, but an NS when compared to the TMB single model—due to dilution of its inflamed-specific signal in this all-sample analysis. A model that includes both DNA (TMB) and RNA markers ($CD8^+$ T_{eff} signature and F-TBRS) as well as interactions between the F-TBRS and both TMB and cancer-immune phenotype explains 50% of the variance observed in response, and it significantly improves on all single- and two-biology models. This final bar is also given on the far right in Fig. 2f. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. NS, $P > 0.1$. Exact likelihood ratio test *P* values: T_{eff} , TBRS versus T_{eff} , 0.0026; T_{eff} , TBRS versus TBRS, 0.0032; T_{eff} , TMB versus T_{eff} , 4.9×10^{-8} ; T_{eff} , TMB versus TMB, 0.2; TBRS, TMB versus TBRS, 6.6×10^{-8} ; TBRS, TMB versus TMB, 0.014; T_{eff} , TBRS, TMB versus T_{eff} , TBRS, 1.9×10^{-7} ; T_{eff} , TBRS, TMB versus T_{eff} , TMB, 0.0028; T_{eff} , TBRS, TMB versus TBRS, TMB, 0.016. T_{eff} , CD8 T-effector gene signature.



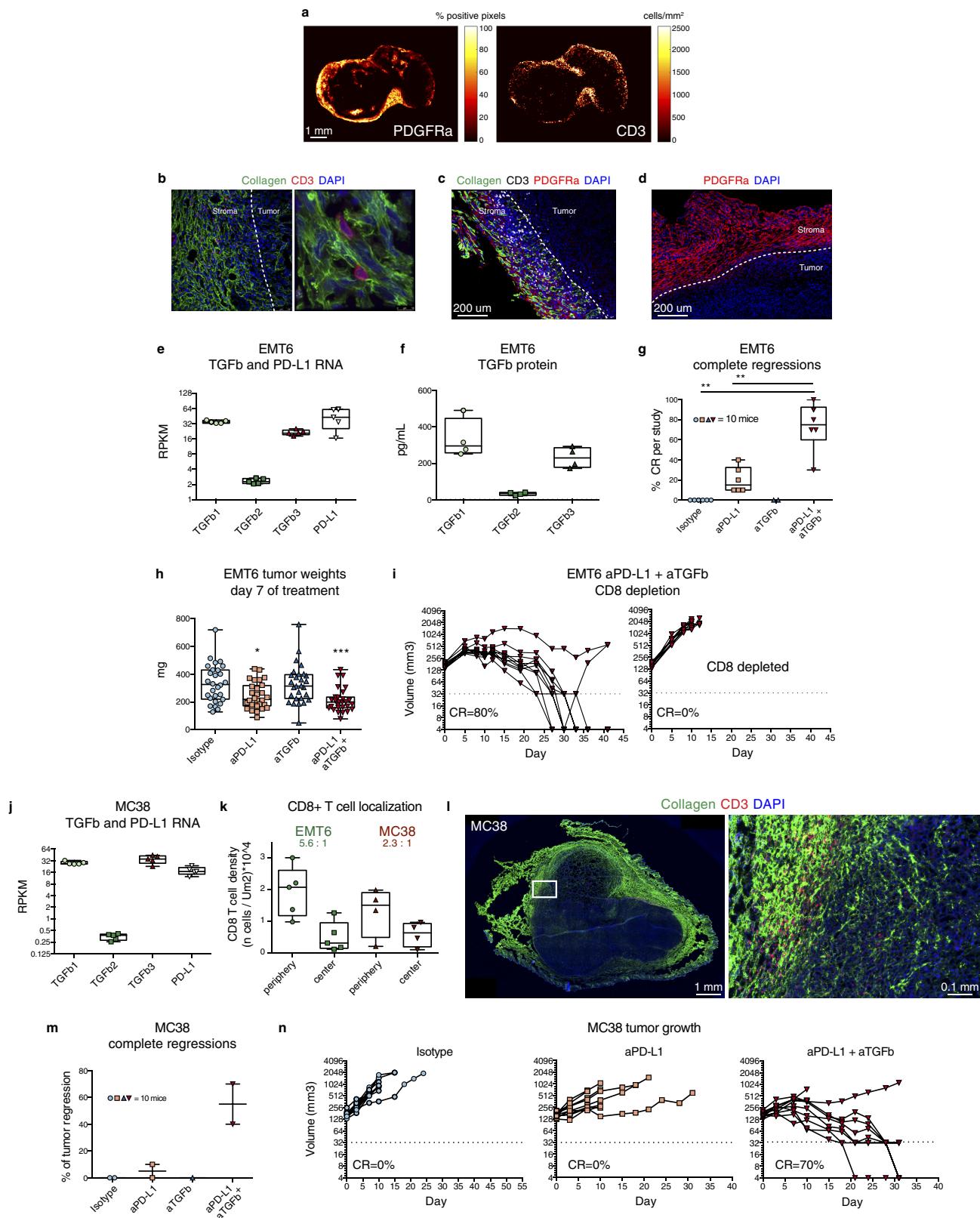
Extended Data Figure 3 | Comparison between Lund and TCGA subtyping schemes. **a**, Heat map representing all evaluated patients, except for patients without defined response, arranged in columns and sorted first by molecular subtype, then by response to atezolizumab. Left, patients were sorted based on a TCGA-based subtyping scheme. Right, patients were sorted by a Lund-based subtyping scheme (as in Fig. 3). Immune cell and tumour cell PD-L1 status are given. In addition, TMB and mutation status (black, mutated; grey, patients without mutation data) for a few genes of interest are shown. The rows of the heat map show expression (Z scores) of genes of interest, grouped into the following biologies and/or pathways: TCGA, TCGA subtyping genes; A, *FGFR3* gene signature; B, CD8 T_{eff} signature; C, antigen-processing machinery; D, immune checkpoint signature; E, *MKI67* and cell cycle genes; F, DNA replication-dependent histones; G, DNA damage-repair genes; H, TGF β receptor and ligand; I, F-TBRS genes; J, angiogenesis signature; K, EMT markers (for details on these signatures see Supplementary Methods). **b**, *FGFR3*-related and WNT target genes³¹ as well as *PPARG* are significantly differentially expressed by Lund subtype (Wald test;

P-values: *FGFR3*-related, 2.7×10^{-43} ; WNT target, 1.3×10^{-15} ; *PPARG*, 1.2×10^{-53}). Gene set membership is given in Supplementary Table 8. **c, d**, Distribution of Lund subtypes by cancer-immune phenotypes and response status. **c**, The fraction of patients within the different Lund subtypes (y axis) is plotted by tumour-immune phenotype. There is a significant difference in Lund-subtype composition between cancer-immune phenotypes (χ^2 test, $P = 1.2 \times 10^{-7}$). **d**, For excluded tumours, the fraction of patients within the different Lund subtypes (y axis) is plotted by response status. Responders, complete and partial response (CR + PR); non-responders, stable and progressive disease (SD + PD). There is a significant difference in Lund-subtype composition between response groups (χ^2 test, $P = 0.00061$). The numbers above the graphs specify sample numbers in each bin. **e**, Assessment of *MKI67* expression and signatures of interest as well as TMB relative to molecular subtypes. Biologies of interest were scaled before medians were calculated across the Lund (left) and TCGA (right) molecular subtypes (columns). Red, high; blue, low. DNA rep., DNA replication.



Extended Data Figure 4 | Contrasting Lund and TCGA molecular subtyping. **a**, TMB (y axis) is plotted against Lund and TCGA subtypes (x axis). The Lund genetically unstable (two-tailed *t*-test, $P = 0.00018$) and TCGA luminal II subtypes ($P = 0.00024$) have a higher median TMB. **b**, Patients are split into TMB low (grey) and high (black), on the basis of median TMB, and the fraction of patients in these two groups is shown for the Lund and TCGA molecular subtypes. **c–e**, TGF β as a probable driver

of differential response in the Lund genetically unstable subtype. **c**, Three patient subgroups: Lund genetically unstable but not TCGA luminal II, both genetically unstable and luminal II, or luminal II but not genetically unstable. **d**, CD8 $^{+}$ T_{eff}, F-TBRS and TMB by subgroup. **e**, Response differs significantly by subgroups (two-tailed Fisher's exact test, $P = 0.00062$). The numbers above the graphs or in parentheses specify sample numbers in each bin.



Extended Data Figure 5 | See next page for caption.

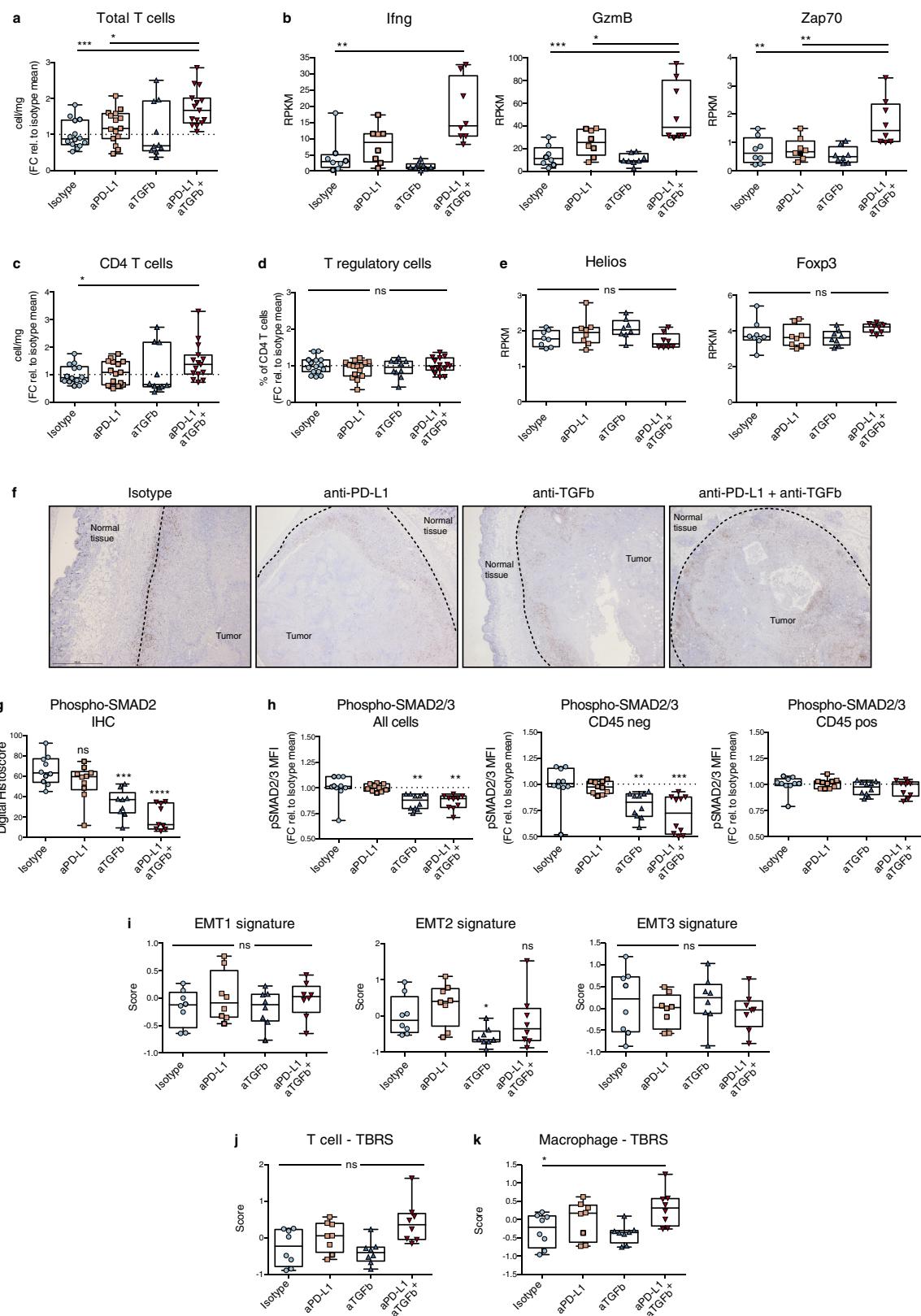
Extended Data Figure 5 | Efficacy data of anti-TGF β and anti-PD-L1 treatment in EMT6 and MC38 immune-excluded tumour models.

a, Fibroblast (PDGFR α , left) and T cell (CD3, right) parametric maps. Left image shows PDGFR α density (per cent positive pixels) and right shows T-cell density (cells mm $^{-2}$). Scale bar, 1 mm. Representative images of eight biological replicates. **b**, Collagen (green) and T cells (CD3, red) stained by immunofluorescence. Representative images of five biological replicates. **c**, Collagen (green), T cells (CD3, white) and PDGFR α (red) in EMT6 tumours stained by immunofluorescence. Scale bar, 200 μ m. **d**, PDGFR α (red) in EMT6 tumours stained by immunofluorescence. Scale bar: 200 μ m. Representative images of four biological replicates. **e**, Quantification of TGF β and PD-L1 RNA in whole EMT6 tumours by RNA-seq. The tumours were inoculated orthotopically and collected when volume reached 300 mm 3 ($n = 5$ mice; data from one experiment). **f**, Quantification of TGF β protein within whole EMT6 tumours by ELISA. Tumours were collected 14 days after inoculation, flash-frozen and lysed for protein quantification ($n = 4$ mice; data from one experiment). **g**, BALB/cJ mice were inoculated orthotopically with EMT6 tumour cells. When tumour volumes reached around 160 mm 3 approximately nine days after inoculation, mice were treated with isotype control, anti-PD-L1, anti-TGF β , or a combination of anti-PD-L1 and anti-TGF β . Tumours were measured two times per week for approximately eight weeks by calliper. When tumour volumes fell below 32 mm 3 (lowest limit of detection), they were considered complete response. Percentage of complete regressions across 2–6-independent studies (10 mice per group per study). **h**, Tumour weights at day seven after initiation of treatment ($n = 28$ mice per treatment group; data from three independent experiments).

i, CD8 depletion experiment. CD8 T cells were depleted before initiation of treatment ($n = 10$ mice per group, data from one experiment).

j, Quantification of TGF β and PD-L1 RNA in whole MC38 tumours by RNA-seq. The tumours were inoculated subcutaneously and collected when volume reached 300 mm 3 ($n = 5$ mice; data from one experiment).

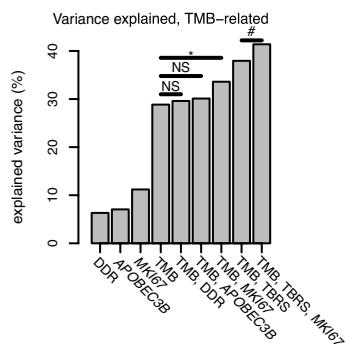
k, Quantification of CD8 T cells in the centre and in the periphery of EMT6 and MC38 tumours from immunohistochemistry stains. Data expressed as number of cells per tissue area (periphery is defined as 400–600 μ m from the tumour edge, centre is the remaining distance to the centre point). EMT6, $n = 5$ mice; MC38, $n = 4$ mice. **l**, Collagen (green) and T cells (CD3, red) in MC38 tumours stained by immunofluorescence. Scale bar, 1 mm (left); 0.1 mm (right, representative images of 5 biological replicates). **m**, C57BL/6 mice were inoculated subcutaneously with MC38 tumour cells. When tumour volumes reached around 180 mm 3 approximately eight days after inoculation, mice were treated with isotype control, anti-PD-L1, anti-TGF β , or a combination of anti-PD-L1 and anti-TGF β . Tumours were measured two times per week for approximately eight weeks by calliper. When tumour volumes fell below 32 mm 3 (lowest limit of detection), they were considered complete response. Percentage of complete response across two independent studies (one for anti-TGF β alone) shown with 10 mice per treatment group for each independent study. **n**, Tumour growth curves for each individual mouse are shown. The data are representative of two independent experiments with 10 mice per treatment group. For box plots, centre mark is median, and whiskers are minimum and maximum. All statistics are two-sided Mann–Whitney *U*-tests compared to isotype group. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.



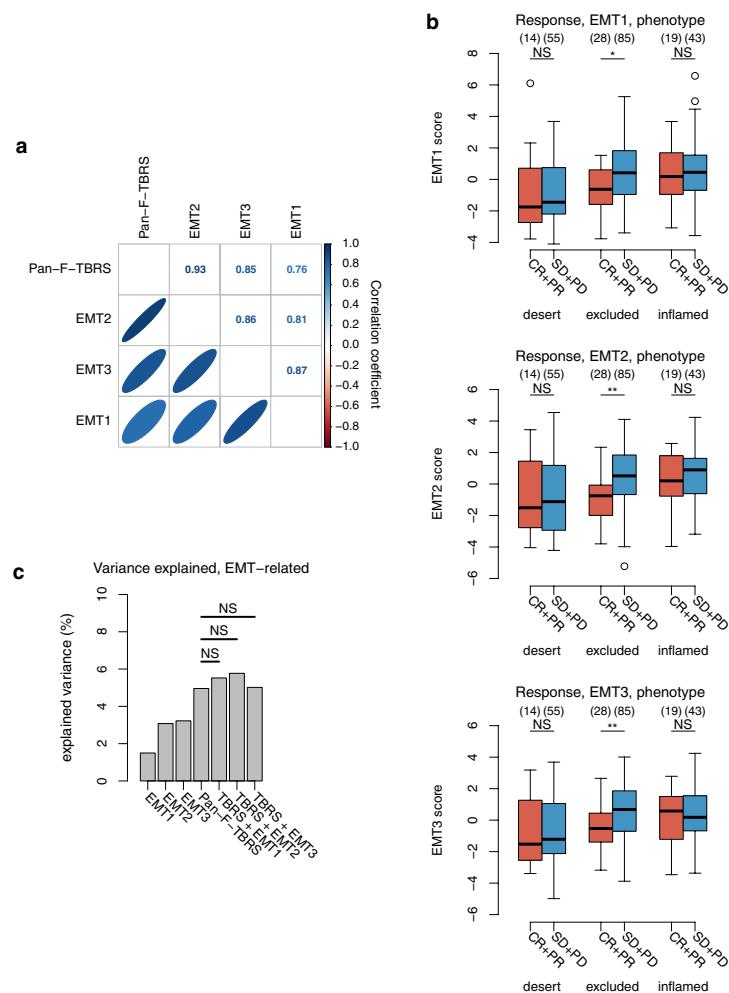
Extended Data Figure 6 | See next page for caption.

Extended Data Figure 6 | Changes in TME following anti-TGF β + anti-PD-L1 treatment in EMT6 tumours. **a, c, d,** Cytofluorimetric analysis of T cells seven days after initiation of the treatment. The abundance of total T cells (**a**), total CD4 $^{+}$ cells (**c**) and the percentage of T-regulatory cells (CD25 $^{+}$ FOXP3 $^{+}$) in the CD4 $^{+}$ population (**d**) are shown. $n = 15$ mice for all treatment groups except for anti-TGF β alone in which $n = 10$; data combined from three independent experiments expressed as fold change relative to the isotype mean. **b, e,** RNA-seq analysis on whole tumours collected seven days after the initiation of treatment. Single-gene expression for *Ifng*, *Gzmb* and *Zap70* (**b**) *Ikzf2* (also known as Helios) and *Foxp3* (**e**) are shown ($n = 8$ mice per treatment group; data from one experiment). **f,** Distribution of tumour-infiltrating lymphocytes in tumours as assessed by immunohistochemistry and digital imaging seven days after the initiation of treatment as above. Representative CD3 staining (brown). Dashed line indicates tumour boundaries. ($n = 19$ for all

groups except anti-PD-L1/anti-TGF β , in which $n = 20$; three independent experiments). Scale bar, 500 μ m. **g,** Quantification of pSMAD2 by immunohistochemistry at day seven after initiation of treatment. ($n = 9$ or 10 mice per treatment group; data from one experiment). **h,** Phosphoflow analysis of SMAD2/3 in tumours seven days after the initiation of treatment as above. MFI of pSMAD2/3 among total cells, CD45 $^{-}$ or CD45 $^{+}$ cells are shown. Data are expressed as fold change (FC) relative to the isotype MFI average. Ten mice per treatment group from two independent experiments. **i–k,** RNA-seq analysis on whole tumours collected seven days after the initiation of treatment. Three EMT signatures (**i**), and TGF β -response signatures for T cells (**j**) and macrophages (**k**) are also shown. ($n = 8$ mice per treatment group; data from one experiment). All statistics in the figure use two-sided Mann–Whitney *U*-test. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$. NS, not significant compared to isotype group.


Extended Data Figure 7 | Explained variance in patient response.

Generalized linear models were fitted using binary response (complete or partial response versus stable or progressive disease) as the dependent variable and scores from single input or input combinations (*x* axis) as independent variables (236 samples). Percentage of explained variance of response is plotted on the *y* axis. Comparisons between different models were made using the likelihood ratio test; a significant *P* value means that the additional variable contributed some independent information to the model. The association of TMB with response is significantly stronger than that of its proxy measurements (*APOBEC3B* and *MKI67* expression or mutation in members of the DDR set). *APOBEC3B* and DDR gene set mutation provided no additional explanatory information independent of direct measurement of TMB. Combining TMB with *MKI67* expression marginally improved on TMB alone, possibly through the negative association between *MKI67* and TFG β (Extended Data Fig. 1e, f). To test this hypothesis, we added *MKI67* to a two-pathway model based on TMB and the F-TBRS, and confirmed that *MKI67* does not add independent information to this two-pathway model. Furthermore, there was no benefit from adding *MKI67* to our full three-pathway model, shown in Fig. 2f and Extended Data Fig. 2g. #*P* < 0.1, **P* < 0.05. Exact likelihood ratio test *P* values: TMB,DDR versus TMB, 0.38; TMB,APOBEC3B versus TMB, 0.26; TMB,MKI67 versus TMB, 0.029; TMB,TBRS,MKI67 versus TMB,TBRS, 0.064.



Extended Data Figure 8 | Relationship between different TGF β related gene expression signatures and response. **a**, Correlation between different TGF β related gene expression signatures. Oval size and colour show correlation between signature scores and gene expression), calculated based on the complete RNA-seq dataset of 348 samples. Numbers show Pearson correlation coefficients. Gene-set membership is shown in Supplementary Table 8. See Supplementary Methods for computation of signature scores. **b**, EMT-signature expression is associated with response to atezolizumab in excluded tumours. Scores of three different EMT signatures, EMT1³², EMT2³³ and EMT3³⁴, are significantly higher in non-responders (stable and progressive disease) than in responders (complete and partial response) in excluded tumours (EMT1, $P=0.0102$; EMT2, $P=0.0027$; EMT3, $P=0.0063$); there is no significant difference in signature scores in desert and inflamed tumours (all $P=1$;

two-tailed t -test, P values for each signature are Bonferroni-corrected for three tests). The numbers above the graphs specify sample numbers in each bin. **c**, Explained variance in patient response. Generalized linear models were fitted using binary response (complete or partial response versus stable or progressive disease) as the dependent variable and scores from single input or input combinations (x axis) as independent variables (233 samples). Percentage of explained variance of response is plotted on the y axis. Comparisons between different models were made via likelihood ratio test; a significant P value means that the additional variable contributed some independent information to the model. The association of F-TBRS with response is the strongest among its correlates, that is, three different EMT signatures. None of these signatures provided additional explanatory information independent of F-TBRS.

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-TGF β 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

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

Initial submission Revised version Final submission

S13n: 2 independent studies - only 1 of them had the anti-TGF β 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 referenced in Powles, T. et al. MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer. *Nature* 515, 558-562 (2014).
 anti-human CD8 rabbit monoclonal antibody SP16 (Spring Bioscience; Cat #M3160) Website: <http://products.springbio.com//products/CD8-SP16-/M316>.
 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-TGF β 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.

Antibodies used for flow cytometry: FcR (Clone 2.4G2, 5 μ g/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).

Antibodies used for IHC: CD3 (clone SP7, Thermo Fisher Scientific, Waltham, MA); pSMAD2 (clone 138D4, Cell Signal Technologies, Danvers, MA).

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.

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.