



OPEN

DATA DESCRIPTOR

# Time-course RNAseq data of murine AB1 mesothelioma and Renca renal cancer following immune checkpoint therapy

Wee Loong Chin<sup>1,2,3,4,5,12</sup>, Rachael M. Zemek<sup>1,6,7,12</sup>, Caitlin M. Tilsed<sup>1,6</sup>, Alistair R. R. Forrest<sup>8</sup>, Vanessa S. Fear<sup>1,6,7</sup>, Catherine Forbes<sup>1,6,7</sup>, Louis Boon<sup>9</sup>, Anthony Bosco<sup>6,10,11</sup>, Belinda B. Guo<sup>8</sup>, Michael J. Millward<sup>2,3</sup>, Anna K. Nowak<sup>1,2,3,5</sup>, Richard A. Lake<sup>1,5,6</sup>, W. Joost Lesterhuis<sup>1,6,7</sup>✉ & Timo Lassmann<sup>7</sup>✉

Time-critical transcriptional events in the immune microenvironment are important for response to immune checkpoint blockade (ICB), yet these events are difficult to characterise and remain incompletely understood. Here, we present whole tumor RNA sequencing data in the context of treatment with ICB in murine models of AB1 mesothelioma and Renca renal cell cancer. We sequenced 144 bulk RNAseq samples from these two cancer types across 4 time points prior and after treatment with ICB. We also performed single-cell sequencing on 12 samples of AB1 and Renca tumors an hour before ICB administration. Our samples were equally distributed between responders and non-responders to treatment. Additionally, we sequenced AB1-HA mesothelioma tumors treated with two sample dissociation protocols to assess the impact of these protocols on the quality transcriptional information in our samples. These datasets provide time-course information to transcriptionally characterize the ICB response and provide detailed information at the single-cell level of the early tumor microenvironment prior to ICB therapy.

## Background & Summary

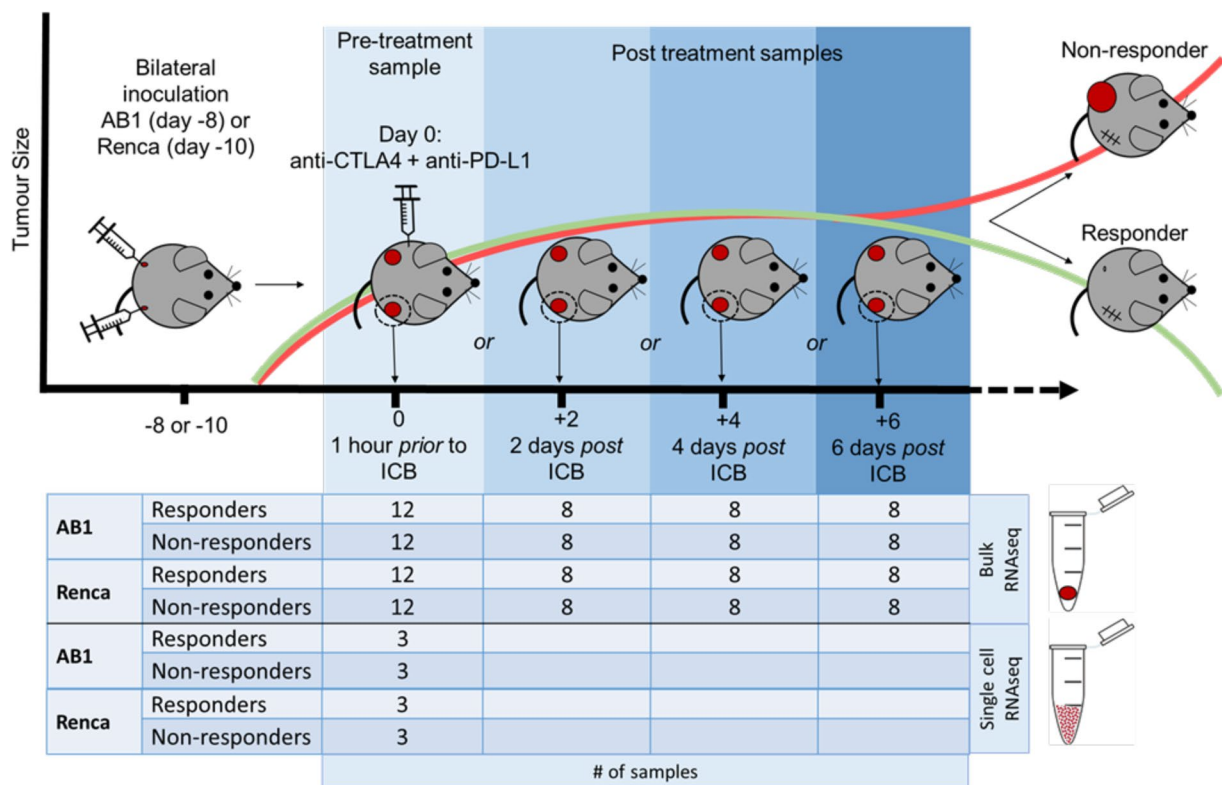
Immune checkpoint blockade (ICB) has been termed the “fourth pillar” of cancer therapy<sup>1</sup>, but the transcriptional events underpinning treatment response remain incompletely understood. Time-critical transcriptional events in the immune microenvironment are important for response to ICB, but such events are difficult to characterise in patients<sup>2</sup>. For one, it is challenging to obtain repeated tumor samples from patients during treatment<sup>3</sup>. Furthermore, the analyses of these patient samples may be confounded by inter-individual differences in germline and cancer genetics, tumor microenvironment (TME) composition and environmental influences.

To acquire reliable tumor transcriptome data across multiple time points after ICB treatment, we created murine models with bilateral tumors derived from syngeneic cancer cell lines. In these models, ICB with antibodies against CTLA4 and PD-L1 leads to either a symmetric bilateral response or a symmetric failure to respond in both tumors<sup>4</sup>. This allows unilateral tumor harvest for RNA sequencing whilst allowing reliable response tracking to ICB treatment on the contralateral tumor.

<sup>1</sup>National Centre for Asbestos Related Diseases, Nedlands, WA, 6009, Australia. <sup>2</sup>Medical School, University of Western Australia, Crawley, WA, 6009, Australia. <sup>3</sup>University of Western Australia, Crawley, WA, 6009, Australia. <sup>4</sup>Department of Medical Oncology, Sir Charles Gairdner Hospital, Nedlands, WA, 6009, Australia. <sup>5</sup>Institute for Respiratory Health, Nedlands, WA, 6009, Australia. <sup>6</sup>School of Biomedical Sciences, University of Western Australia, Crawley, WA, 6009, Australia. <sup>7</sup>Telethon Kids Institute, University of Western Australia, West Perth, WA, 6872, Australia. <sup>8</sup>Harry Perkins Institute of Medical Research, QEII Medical Centre and Centre for Medical Research, The University of Western Australia, Nedlands, Perth, WA, 6009, Australia. <sup>9</sup>JJP Biologics, Warsaw, Poland. <sup>10</sup>Asthma and Airway Disease Research Center, University of Arizona, Tuscon, Arizona, USA. <sup>11</sup>Department of Immunobiology, The University of Arizona, College of Medicine, Tuscon, Arizona, USA. <sup>12</sup>These authors contributed equally: Wee Loong Chin, Rachael M. Zemek. ✉e-mail: [joost.lesterhuis@telethonkids.org.au](mailto:joost.lesterhuis@telethonkids.org.au); [timo.lassmann@telethonkids.org.au](mailto:timo.lassmann@telethonkids.org.au)

Sample type	Number	Description
Bulk	144	72 samples AB1, 72 samples Renca
		For each model:
		- 36 Responders and 36 NR - 12 time point zero, 8 time point 2,4,6
Single cell	12	6 samples AB1, 6 samples Renca
		For each model:
		- 3 responders and 3 non responders
Bulk	8	8 samples AB1-HA:
		- 4 tumor parts immediately submerged in RNAlater for bulk RNAseq
		- 4 tumor parts from same tumors gentleMACs dissociated before bulk RNAseq

**Table 1.** Number of samples, timepoints, type of sequencing data available.



**Fig. 1** Workflow for responder/non-responder sample generation. The bilateral tumor model was used to generate samples, with tumor harvest occurring at successive timepoints pre and post treatment with ICB. Adapted from Fig. 1a in “Temporally restricted activation of IFN $\beta$  signalling underlies response to immune checkpoint therapy in mice,” by Zemek, R. M. *et al.*, Nat Commun 13, 4895 (2022).

By harvesting tumors across multiple time points in parallel experiments, we create biological replicates of reliably time-indexed RNAseq data of the tumor transcriptome, which would not be possible in a patient cohort. Previously, we used our murine models to characterize the response-associated pre-treatment TME in detail<sup>5</sup>. In our current work, we produced and analysed three datasets (Table 1). Firstly, we produced a set of 144 bulk RNAseq samples from AB1 mesothelioma and Renca renal cell cancer across 4 time points prior (day 0) and after treatment (day 2, 4 and 6) with ICB. Secondly, we sequenced 12 samples of AB1 and Renca tumors an hour before ICB administration using single cell sequencing (Fig. 1, Table 2). Additionally, we sequenced 4 AB1-HA tumors with bulk RNAseq that were prepared using tumor dissociation or not to determine the impact of the dissociation process on data quality. We describe this final dataset as part of our technical validation.

Previous datasets in the ICB literature have focused on genomic correlates with non-temporal transcriptional determinants<sup>5,6</sup>, fewer time points<sup>7</sup> or sparser sampling density with fewer biological replicates<sup>8</sup>. To our knowledge, our work represents the first example of densely time-indexed RNAseq data of the early TME in tumors using both bulk and single cell RNA sequencing. In addition, it samples whole tumors prior to therapy, and compares complete responders and non-responders in the same model. We thus provide a detailed picture of the tumor transcriptome in the early stages after ICB therapy.

Sample name	tumor type	ICB Response
NKD180900300	Renca	Non-responder
NKD180900301	Renca	Responder
NKD180900302	AB1	Responder
NKD180900303	AB1	Non-responder
NKD180900304	AB1	Non-responder
NKD180900305	Renca	Responder
NKD180900306	Renca	Non-responder
NKD180900307	AB1	Responder
NKD180900308	AB1	Non-responder
NKD180900309	Renca	Responder
NKD180900310	Renca	Non-responder
NKD180900311	AB1	Responder

**Table 2.** Sample metadata for single cell samples.

## Methods

With the exception of the tumour dissociation protocol described below, these methods have been described in previous work<sup>6</sup>. Here, we restate descriptions of the experimental murine models, cell culture techniques, *in vivo* treatments and tumour preparation for ease of reference.

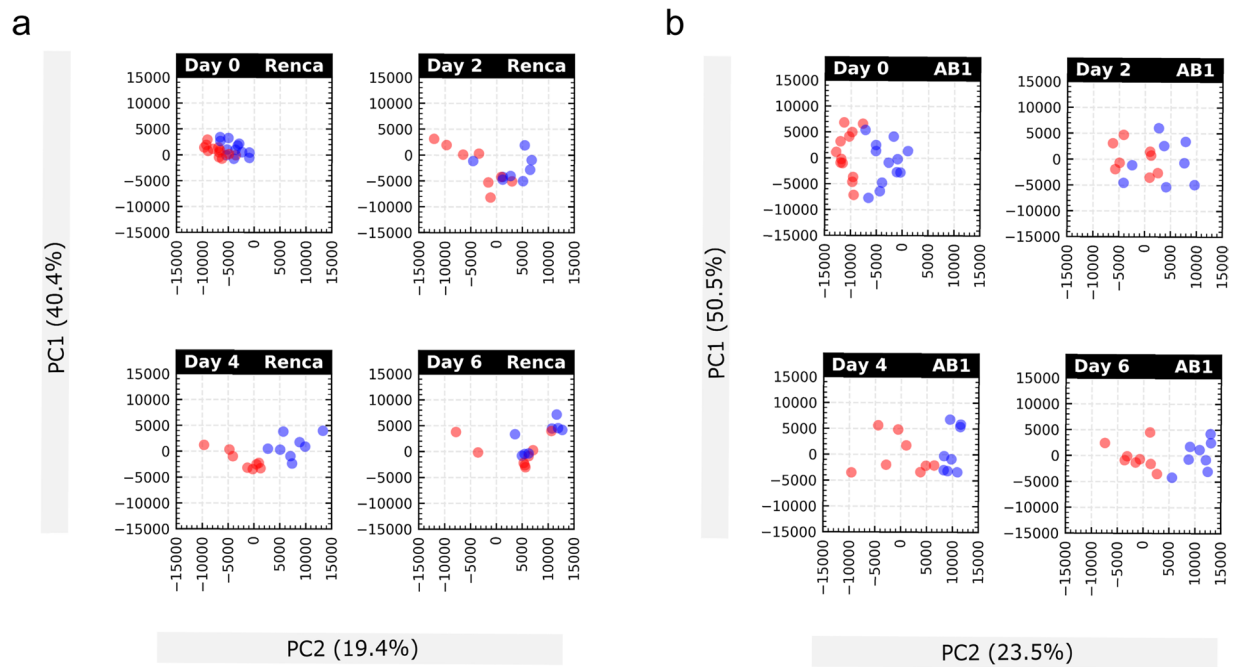
**Mice.** Balb/cAusB mice 8–12 weeks of age were used for all experiments. BALB/cArc mice were obtained from the Animal Resource Centre (Murdoch, WA). All mice were housed at the Harry Perkins Institute of Medical Research Bioresources Facility North under specific pathogen free conditions. Mice were kept in individually ventilated cages (Techniplast, Italy) supplied with filtered air. Cages contained aspen chips bedding (Tapvei, Estonia) and for environmental enrichment were supplemented with wood blocks, tissues, cardboard rolls and were changed every 14 days. Mice were fed Rat and Mouse cubes (Specialty Feeds, Glen Forrest, Australia) and had access to water *ad libitum*. Mice were housed at 21–22 °C with 12-hour light/dark cycle (06:00–18:00). To check SPF conditions were maintained in the animal facility, sentinel mice (n = 3) were screened for a standard panel of bacteria and fungi, non-pathogenic protozoa, endoparasites, ectoparasites and viruses (Cerberus Sciences, Australia) each month. All experiments were conducted in compliance with the institutional guidelines provided by the Harry Perkins Institute for Medical Research animal ethics committee (approval number AE047).

**Cell culture.** Cell line AB1 was obtained from CellBank Australia. Cell line Renca<sup>9</sup> was kindly donated by Dr E. Sotomayor and Dr F. Cheng (University of South Florida, Tampa, FL). Cell lines were maintained in RPMI 1640 supplemented with 20 mM HEPES (Gibco), 0.05 mM 2-mercaptoethanol (Sigma Aldrich), and 10% FCS (Invitrogen, Mulgrave, Australia). AB1 and Renca were supplemented with 100 units/ml penicillin/streptomycin (Thermo Fisher), AB1-HA was supplemented with 100 U/ml benzylpenicillin (CSL), 50 ug/ml gentamycin (David Bull Labs) and 50 mg/ml geneticin (G418; Life Technologies). Cells were grown to 70–80% confluency before each passage and passaged 3–5 times before inoculation. Cells were frequently tested for mycoplasma by PCR and remained negative. Cell lines were validated yearly by flow cytometry for MHC class I molecules H2-Kb (negative) and H2-Kd (positive), and for fibroblast markers E-cad, EpCam and PDGFR $\alpha$  (negative) and by PCR for mesothelin (positive for AB1, negative for Renca).

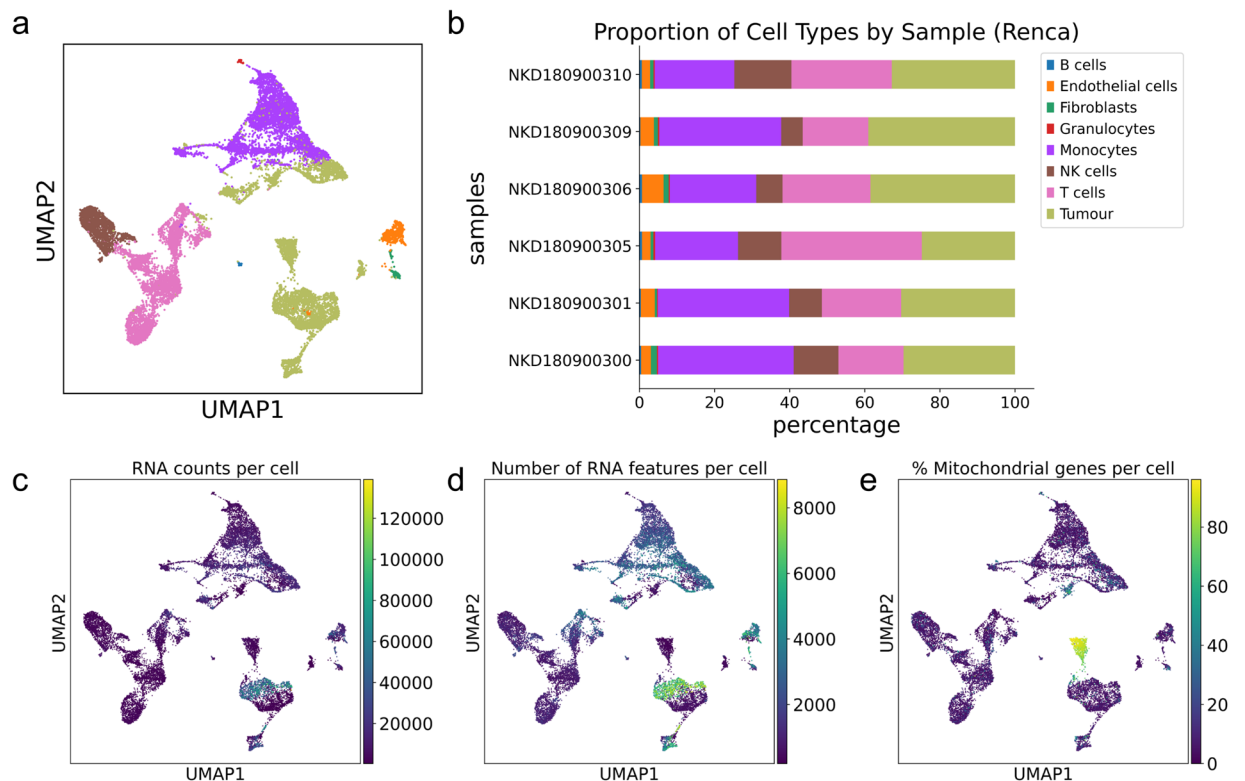
***In vivo* treatments.** We have described the dual tumor model including inoculation and surgery in detail previously<sup>4</sup>. When cell lines were 70–80% confluent, they were harvested and washed 3 times in PBS.  $5 \times 10^5$  cells in 100  $\mu$ l were inoculated subcutaneously (s.c.) onto the lower flank on both sides using a single 26 G needle per injection. Mice were randomized when tumors became palpable, approximately 3–5 days after tumor inoculation. AB1-HA was injected subcutaneously on one flank.

The anti-CLTA4 hybridoma (clone 9H10) and anti-PD-L1 hybridoma (clone MIH5) were cultured in IMDM containing 1% of FCS and gentamycin at Bioceros (Utrecht, The Netherlands). Clarified supernatants were used to purify the antibody using affinity chromatography. The antibodies were sterile formulated in PBS. Mice received an intraperitoneal (i.p.) dose of 100  $\mu$ g of anti-CTLA4 and 100  $\mu$ g anti-PDL1 combined in 100  $\mu$ l PBS. Mice received additional doses of 100  $\mu$ g anti-PDL1 two and four days later.

**Surgery experiments (responder/nonresponders).** For the pre-treatment samples, tumors were resected when they were 9 mm<sup>2</sup>, (eight days for AB1 or 10 days for Renca post tumor inoculation), and mice were administered ICB 1 hour after surgery. For the post-treatment samples, tumors were resected 2, 4 or 6 days after the first administration of ICB. Mice were dosed with 0.1 mg/kg buprenorphine in 100  $\mu$ l s.c. (30 min prior) and anesthetized using isoflurane (4% in 100% oxygen at a flow rate of 2 L/min). Whole tumors and the corresponding draining inguinal lymph node on the right-hand side were surgically removed and immediately immersed in RNAlater (Life Technologies, Australia). The wound was closed with staples (Able Scientific, Australia), and mice were placed in a heat box at 37 °C for recovery. As an indicator of response to the removed tumor, the remaining tumor was monitored for response. Mice were designated as responders if their tumor completely regressed and the tumor did not return for up to 4 weeks after treatment. Mice were designated as non-responders if their

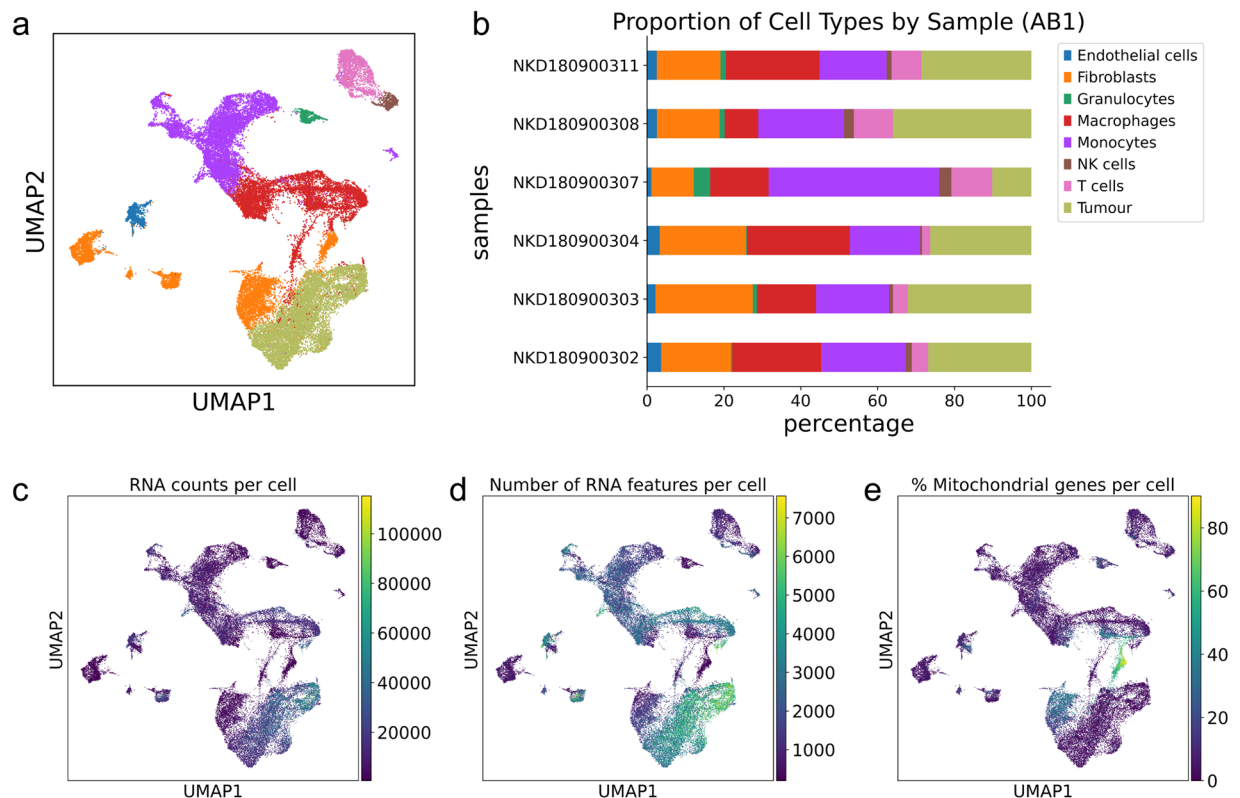


**Fig. 2** PCA plots of AB1 and Renca faceted by day and response. (a) shows PCA plots in Renca and (b) in AB1 (Blue dots represent responders; red dots non responders).



**Fig. 3** Quality control metrics for single cell data from Renca samples. (a) Combined Renca data from 6 single cell samples on UMAP and (b) sample composition for each Renca sample, faceted by cell type (c–e). Quality control metrics (RNA counts per cell, feature counts per cell and percentage mitochondrial content distribution) across all cells on Renca UMAP.

tumors reached 100 mm<sup>2</sup> within 4 weeks after start of treatment, similar to saline-treated controls. Mice were designated as intermediate responders if they had a delay in tumor growth or partial regression and excluded from



**Fig. 4** Quality control metrics for single cell data from AB1 samples. **(a)** Combined AB1 data from 6 single cell samples on UMAP and **(b)** sample composition for each AB1 sample, faceted by cell type **(c–e)**. Quality control metrics (RNA counts per cell, feature counts per cell and percentage mitochondrial content distribution) across all cells on AB1 UMAP.

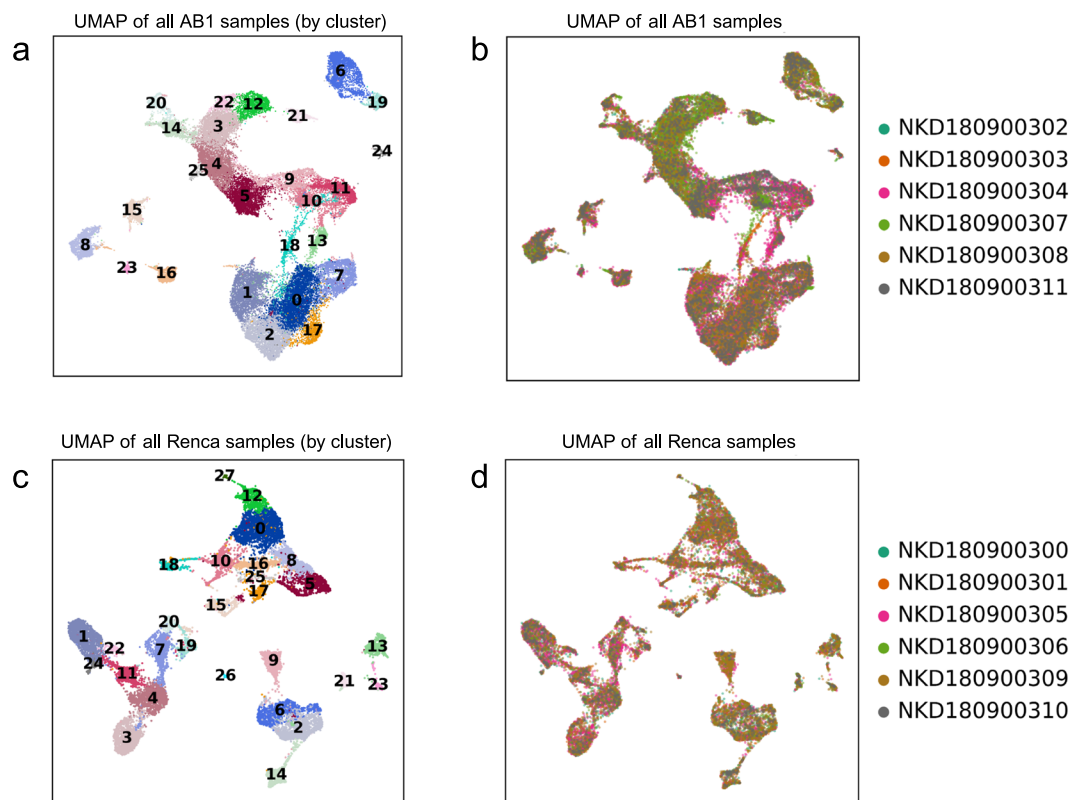
the analysis. For internal consistency, we only used experiments in which both responders and non-responders were present in any cage, i.e., there had to be at least one non-responder amongst responders or vice versa.

**Tumor preparation for bulk RNA sequencing.** Whole tumors once surgically resected, had the surrounding tissue (e.g. skin, fat) removed and were immediately submerged in RNAlater (Life Technologies, Australia). Samples were stored at 4 °C for 24 hours, after which the samples were blotted dry and transferred to a new tube and stored at –80 °C. Frozen tumors were dissociated in Trizol (Life Technologies, Australia) using a TissueRuptor (QIAGEN, Australia). RNA was extracted using chloroform and the aqueous layer was purified on RNeasy MinElute columns (QIAGEN, Australia). RNA integrity (RIN score > 8) was confirmed on the Bioanalyzer (Agilent Technologies, USA). Library preparation and sequencing (50 bp, single-end) was performed by Australian Genome Research Foundation using Illumina HiSeq standard protocols.

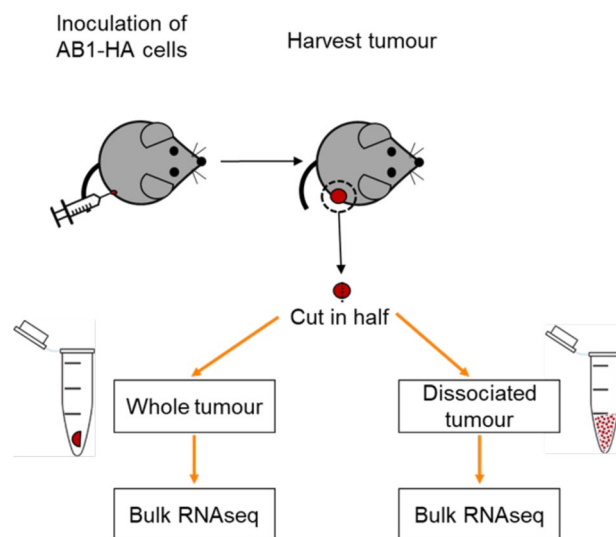
**Tumor preparation for single cell RNA sequencing.** For single-cell RNAseq, tumors were surgically resected as outlined above, and submerged in cold PBS. To obtain a single-cell suspension, tumors were cut into 1- to 2-mm pieces with a scalpel blade and dissociated using the gentleMACS system (Miltenyi Biotec). Single-cell suspensions were stored at –80 °C in 10% dimethyl sulfoxide until it could be determined if they could be designated as responders or non-responders and then they were processed for single-cell profiling. Cryo-stored cells were rapidly thawed and diluted in PBS and pelleted. Pellets were resuspended in PBS and passed through a 40-µm filter to remove cell clumps. Approximately 5000 cells per sample were then loaded onto a 10x Genomics Chromium Controller to generate Chromium Single Cell 3' Libraries. Sequencing was carried out by Novogene Co., Ltd.

**Comparison of tumor dissociation protocols.** To compare bulk vs dissociated sample gene expression, tumors were harvested and cut in half. One half was submerged in RNAlater, and the other was dissociated as described above, pelleted, and resuspended in RNAlater. Samples were stored at 4 °C for 24 hours. RNAlater was removed from bulk samples via blotting and single-cell samples by pelleting, before storing at –80 °C. Both sample types had RNA extracted as described above. Library preparation and bulk RNA sequencing (50 bp, single-end) was performed by Australian Genome Research Facility, using Illumina HiSeq standard protocols.

**Alignment and differential expression of bulk RNAseq data.** We processed a total of 144 RNAseq single-end 50 bp read samples, which comprised of four time points in two mouse models. After reviewing quality control on all samples using FastQC software, we used Kallisto (v0.43.0) for transcript abundance estimation. Gencode



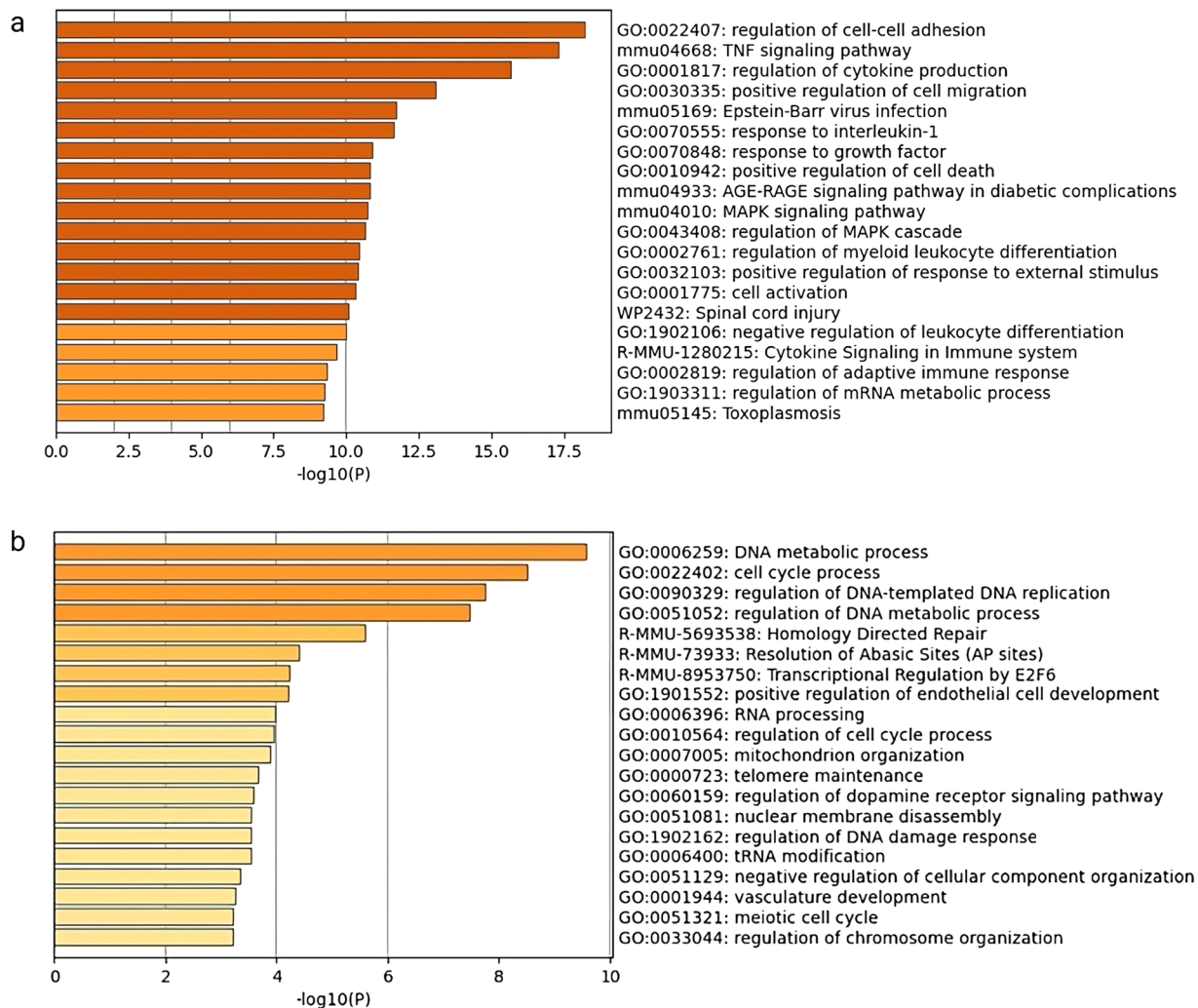
**Fig. 5** Clustering and batch effect correction on single cell data from AB1 and Renca. (a) UMAP of AB1 cells faceted by Seurat clusters and by (b) sample of origin. (c) UMAP of Renca cells faceted by Seurat clusters and by (d) sample of origin. Sample metadata is provided in Table 2.



**Fig. 6** Workflow for whole vs single cell sample generation. Each half was subjected to two different sample preparation protocols. Bulk RNA sequencing was performed on both sample preparations to compare the effect of these protocols on downstream differential analysis.

M21 (GRCm38.p6) transcript sequences were used to create the index for pseudoalignment. For the quantification of single-end reads, we specified an average fragment length of 200 base pairs and a standard deviation of fragment length of 20.

We also processed a further 8 RNAseq single-end 50 bp read samples for the comparison of whole versus dissociated samples using the steps described above. To analyse bulk RNAseq samples from AB1-HA prepared using gentleMACS vs RNAlater, we used the same alignment strategy with Kallisto. Following this, we



**Fig. 7** Functional annotation of DE genes between two sample dissociation protocols (gentleMACs versus maceration of frozen tissue). **(a)** functional annotation for upregulated genes after dissociation using gentleMACs. **(b)** functional annotation downregulated genes using gentleMACs.

performed differential expression analysis using the Wald test with Sleuth (v0.29.0), with sample preparation as a covariate. Genes were deemed differentially expressed at an absolute log fold change of 0.5 and a false discovery rate of less than or equal to 0.05. Enrichment of differentially enriched gene sets was performed with Metascape<sup>10</sup>.

**Single cell sample demultiplexing and UMI counting.** For single cell analysis, we processed FASTQ files from 6 AB1 and 6 Renca samples using cellranger v3.0 (10X genomics). For each sample, we performed demultiplexing and read alignment using the cellranger count function, using cellranger's pre-supplied mm10 reference with an expect-cells parameter of 6000.

### Data Records

All sequencing data have been uploaded to the NCBI GEO database. Project data is available through accession number GSE153943<sup>11</sup>.

### Technical Validation

**QC control on bulk RNAseq data.** We performed FastQC across all bulk RNAseq samples sequencing depth ranging 20 – 30 million reads on all samples, with approximately 40% unique reads per sample. Following alignment and transcript quantification, samples visualised with PCA (Fig. 2) did not demonstrate significant sample-specific batch effects.

**Cell labelling.** We used an automated labelling strategy based on bulk RNAseq references, which we have described in detail in a previous publication. Briefly, the R package SingleR<sup>12</sup> was used in “cluster mode” using species-specific annotation references provided with the package. For annotation of murine data, we used the mouse RNAseq dataset<sup>13</sup>. Clusters were defined from Seurat's FindClusters function at default (0.8) resolution.

Similarly, labelled clusters were merged (Figs. 3a,b, 4a,b). We confirmed that this approach was robust to cluster size by showing that labels were consistent even when cluster size was modified by changing resolution parameters in the FindClusters function and that labelling was consistent with cell canonical markers. Since batch specific differences could introduce issues with cell identification, the consistency of cell labels allowed us to confirm the robustness of our clustering and batch-effect removal strategy.

**Normalisation and removal of batch effects.** In both AB1 and Renca, we used the Seurat (version 3.14) R package to combine samples for downstream analysis. Gene counts were normalized against both sequencing depth and against the percentage of mitochondrial DNA in each cell using negative binomial regression. The resulting Pearson residuals from these processing steps were used for downstream PCA, cluster identification and UMAP embedding and visualization.

The inspection of sample-specific UMAP plots (Fig. 5) showed that this normalisation procedure abolished batch-specific differences in AB1 and Renca samples. After QC normalisation and integration, we used 38,000 cells in AB1 and 18,000 cells in Renca for downstream analysis. We could identify 28 cell clusters that consisted of cells in the range of 58–1877 cells per cluster in Renca and 26 cell clusters in AB1 that consisted of cells in the range of 204–3963 cells per cluster. We did not detect batch specific differences in mitochondrial content or feature count after integration with Seurat (Figs. 3c–e, 4c–e).

**Validation of the single cell RNAseq (gentleMacs) dissociation protocol.** Processing of bulk RNA seq samples involved disruption of RNAlater-preserved frozen tissue in Trizol. In contrast, the single cell sample preparation involves a cell dissociation step at room temperature using the gentleMACs protocol. To assess the impact of these protocols on the quality transcriptional information in our samples, we performed bulk RNA sequencing on 8 samples, with 4 samples prepared using either protocol (Fig. 6). We performed downstream analysis with DE and pathway analysis. In samples prepared using gentleMACS dissociation, 871 DE genes were upregulated and 840 genes downregulated relative to dissociation of frozen macerated tissue in Trizol (thresholded at log-fold change (beta-value) of 1.0, fdr level of 0.05). Pathway analysis using metascape<sup>10</sup> on these differentially expressed genes is shown in Fig. 7.

### Code availability

The code used to reproduce these analyses are available on GitHub [https://github.com/wlchin/scientific\\_data\\_ab1\\_renca](https://github.com/wlchin/scientific_data_ab1_renca) as snakemake<sup>14</sup> workflows.

Received: 20 June 2023; Accepted: 23 April 2024;

Published online: 03 May 2024

### References

1. Binnewies, M. *et al.* Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat Med* **24**, 541–550, <https://doi.org/10.1038/s41591-018-0014-x> (2018).
2. Chin, W. L., Zemek, R. M., Lesterhuis, W. J. & Lassmann, T. Functional genomics in cancer immunotherapy: computational approaches for biomarker and drug discovery. *Molecular Systems Design & Engineering* **4**, 689–700 (2019).
3. Lesterhuis, W. J. *et al.* Dynamic versus static biomarkers in cancer immune checkpoint blockade: unravelling complexity. *Nat Rev Drug Discov* **16**, 264–272, <https://doi.org/10.1038/nrd.2016.233> (2017).
4. Zemek, R. M. *et al.* Bilateral murine tumor models for characterizing the response to immune checkpoint blockade. *Nat Protoc* <https://doi.org/10.1038/s41596-020-0299-3> (2020).
5. Zemek, R. M. *et al.* Sensitization to immune checkpoint blockade through activation of a STAT1/NK axis in the tumor microenvironment. *Sci Transl Med* **11**, <https://doi.org/10.1126/scitranslmed.aav7816> (2019).
6. Zemek, R. M. *et al.* Temporally restricted activation of IFN $\gamma$  signaling underlies response to immune checkpoint therapy in mice. *Nature Communications* **13**, 4895, <https://doi.org/10.1038/s41467-022-32567-8> (2022).
7. Hwang, S. *et al.* Immune gene signatures for predicting durable clinical benefit of anti-PD-1 immunotherapy in patients with non-small cell lung cancer. *Scientific Reports* **10**, 643, <https://doi.org/10.1038/s41598-019-57218-9> (2020).
8. Riaz, N. *et al.* Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab. *Cell* **171**, 934–949.e16, <https://doi.org/10.1016/j.cell.2017.09.028> (2017).
9. Salup, R. R. & Wiltrot, R. H. Adjuvant immunotherapy of established murine renal cancer by interleukin 2-stimulated cytotoxic lymphocytes. *Cancer Research* **46**, 3358–3363 (1986).
10. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Communications* **10**, 1523, <https://doi.org/10.1038/s41467-019-09234-6> (2019).
11. Zemek, R. M. & Lesterhuis, W. J. *GEO* <https://identifiers.org/geo/GSE153943> (2020).
12. Aran, D. *et al.* Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol* **20**, 163–172, <https://doi.org/10.1038/s41590-018-0276-y> (2019).
13. Benayoun, B. A. *et al.* Remodeling of epigenome and transcriptome landscapes with aging in mice reveals widespread induction of inflammatory responses. *Genome Research* **29**, 697–709, <https://doi.org/10.1101/gr.240093.118> (2019).
14. Mölder, F. *et al.* Sustainable data analysis with Snakemake. *F1000Research* **10**, 33, <https://doi.org/10.12688/f1000research.29032.2> (2021).

### Acknowledgements

This work was funded by NHMRC grants 1103980 and 1107043. W.L.C. was supported by a scholarship from the Asbestos Diseases Society of Australia (ADSA), a Raine Foundation Clinician Research Fellowship, and a WA Department of Health Cancer and Palliative Care Network Fellowship. R.M.Z. was supported by a Forrest Prospect Fellowship. T.L. was supported by a Feilman Foundation fellowship and a Stan Perron fellowship. W.J.L. was supported by a Simon Lee fellowship, an NHMRC fellowship and a Cancer Council WA fellowship. A.R.R.F. is supported by an NHMRC fellowship, APP1154524. Part of this research was made possible by the Australian Cancer Research Foundation Centre for Advanced Cancer Genomics, a collaborative cancer research grant provided by the Cancer Research Trust and an enabling grant from Cancer Council WA.



### Author contributions

R.M.Z., C.T., V.F. and C.F. and B.B.G. performed experimental work. A.R.R.F., A.B. and B.B.G. provided technical input. L.B. contributed new reagents. W.L.C., R.M.Z., M.M., A.K.N., R.A.L., W.J.L. and T.L. analysed the data. W.L.C. performed computational analyses. W.L.C., R.M.Z., W.J.L. and T.L. designed the research. W.L.C., R.M.Z., W.J.L. and T.L. wrote the manuscript. W.J.L. and T.L. supervised and managed the overall study. All authors reviewed the manuscript.

### Competing interests

W.J.L.: consultancy Douglas Pharmaceuticals, MSD, research funding Douglas Pharmaceuticals, AstraZeneca. A.K.N. advisory boards Boehringer Ingelheim, Bayer, Roche, BMS; research funding from AstraZeneca. M.M.: Advisory boards MSD, BMS, Roche, AstraZeneca. A.B.: Founder and equity holder of the startup company INSiGENe Pty Ltd that is related to this work. A.B. is a co-founder, equity holder and director of the startup company Respiradigm Pty Ltd that is unrelated to this work.

### Additional information

**Correspondence** and requests for materials should be addressed to W.J.L. or T.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article 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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024