

Table of contents

Contents

Table of contents.....	1
Introduction	4
Cancer immunometabolism.....	5
Cancer and the immune system.....	5
Cellular metabolism in normal cells and cancer cells	5
Glycolysis and lactate production	6
Amino acids and their derivatives	8
Itaconate	10
Adenosine signaling.....	11
The cyclooxygenase and PGE ₂ pathway.....	12
Fatty acids and cholesterol	12
Metabolic programs of trained immunity	13
Metabolic effects of immune checkpoints	14
Mitochondrial regulation of T cells.....	15
Metabolic interventions in ACT.....	16
Common bulk metabolic analyses.....	18
Extracellular flux analysis.....	18
Steady-state metabolomics	19
Fluxomics	19
Key immunometabolism concepts derived from bulk analyses.....	19
Glycolysis supports diverse immune effector functions	19
The pentose phosphate pathway in inflammatory myeloid cells	20
The TCA cycle, ETC as a central hub	20
Fatty acid metabolism supports immune cell phenotypes and functions	20
Amino acid metabolism is differentially regulated in distinct immune cells	21
Limitations of current wet-bench approaches.....	21
The need for in vitro to in vivo validation	21
The need to understand the spatiotemporal aspects of immunometabolism	22
Cell numbers and sorting time limitation	22
Leveraging different 'omics approaches for indirect metabolic assessment.....	22

Metabolic analysis based on single-cell transcriptomics.....	23
Pathway-based approaches	24
Flux balance analysis-based approaches	24
Protein-based single-cell metabolic analysis	25
Single-cell metabolic profiling by mass cytometry	25
Application of metabolic CyTOF in the ex vivo clinical context.....	25
Use of CyTOF-based metabolic profiling in animal model research.....	26
Met-Flow: Single-cell metabolic profiling by flow cytometry	26
Pros and cons of distinct single-cell metabolic profiling techniques and developments.....	27
Single-cell transcriptomics versus cytometry.....	27
Mass cytometry versus flow cytometry for single-cell immunometabolism research.....	28
Spatially resolved immunometabolism	29
MIBI-TOF as protein-based spatial single-cell immune/metabolic analysis	29
Mapping metabolism within microenvironments in situ using dehydrogenase activity assays	30
Computational tools for predicting neoantigens	31
Characterizing tumor-infiltrating immune cells.....	33
Bulk analysis	33
Cell phenotypes from single-cell data	34
Lymphocyte receptor repertoires	36
Spatial cellular phenotyping.....	37
Visualization of single-cell data.....	38
Pathway enrichment visualization and analysis of omics data	39
Stage 1: Definition of a gene list using omics data	40
Stage 2A: pathway enrichment analysis of a gene list using g:Profiler (Step 6A)	41
Stage 2B: pathway enrichment analysis of a ranked gene list using GSEA.....	42
Stage 3: interpretation and visualization of pathway enrichment analysis results	43
METHODS: Pathway enrichment analysis.....	44
Specification of the following parameters:	45
Running GSEA	45
Examination of GSEA results	45
Variation of enrichment results with EnrichmentMap	47
Specification of additional files:	48
Tuning parameters.....	48

Specification of additional files	49
Tuning parameters.....	49
Navigation and interpretation of the enrichment map	50
Exporting figures, creating legends, and saving work	51
OCTAD: virtual therapeutic screening platform.....	52
Procedure 1: Desktop version	52
Setup	52
Case and control samples	53
Select case samples	53
Compute or select control samples	54
Evaluation of synergy-driving gene expression.....	55
Procedure.....	57
Setup	57
Preparing the R environment.....	57
Loading and preprocessing data.....	62
Fitting a linear model.....	64
Assessing differential expression	66
Determining power to detect synergistic effects	67
Determine the extent of synergy	69
Identification and categorization of synergistic genes	70
Enrichment analysis.....	72
Specific gene sets: gene ontology over-representation analysis	75
Anticipated results.....	78
Differential expression analysis.....	78
Synergistic effect analysis	78
Gene set enrichment analysis	79
CHiCANE.....	79
Procedure.....	80
Pre-processing data	80
CHiCANE interaction calling	82
Interpretation and post-processing.....	85
Anticipated results.....	90
Assessing alignments and quality of CHI-C library.....	90

Assessing interaction calls of CHiCANE on promoter CHi-C data	90
Enrichment of enhancer marks	92
Overlap with topologically associating domains	94
Exploring the impact of covariates and statistical distributions	95
Conclusions	97

Introduction

This bioinformatics portfolio is written to fulfill the “Final Portfolio” component of the AS410.635.83 Bioinformatics: Tools for Genome Analysis course at Johns Hopkins University. My chosen topic is “**Computational Tools for Cancer Immunometabolism Research**,” focusing on machine learning tools.

After a succinct introduction to the tumor immunometabolism field, the portfolio will detail recently developed bioinformatics tools that can be employed in a tumor immunometabolism project. Whenever possible, the portfolio will indirectly apply topics taught in the aforementioned class.

Cancer and the immune system

Cancer is a genetic condition characterized by an evolutionary, progressive accumulation of genomic aberrations that predisposing germline mutations may augment. High heterogeneity is a characteristic feature of tumor evolution. Higher intra-tumor heterogeneity has been associated with poorer prognosis and linked with the tumor's ability to resist immune surveillance and therapy.

The intrinsic complexity of the interaction between tumor and the immune system poses considerable challenges and requires comprehensive approaches to interrogate cancer immunity during tumor initiation and progression and following therapeutic modulation. Several high-throughput technologies enable the generation of the necessary data and thereby provide the basis for mechanistic understanding, and ultimately increase the number of patients that benefit from cancer immunotherapy. Next-generation sequencing (NGS) technologies have not only provided large datasets that can be mined for immunologically relevant parameters but are also increasingly used in a clinical setting to inform cancer therapy. Additionally, novel technologies such as single-cell RNA sequencing (scRNA-seq) and mass cytometry by time of flight (CyTOF) have matured and enabled the precise characterization of molecular processes at the single-cell level. The widespread use of NGS techniques and the continuous development of novel high-throughput technologies require an expanded computational toolbox to analyze and visualize heterogeneous data.

Cellular metabolism in normal cells and cancer cells

Metabolism involves a nexus of biochemical reactions that convert nutrients into metabolites. Via these conversions and the resulting metabolites, cells generate energy, redox equivalents, and macromolecules required to survive and sustain cellular functions. Moreover, metabolic profiles reflect the cellular state, and metabolic pathways are intimately entwined with cell signaling and epigenetic networks. Thus, metabolism has a central role in cellular homeostasis and adaptation in response to intracellular and extracellular stimuli. The essential nutrients available to the cells include glucose, amino acids, and fatty acids. These nutrients are mostly converted and used in central metabolism, consisting of catabolic glycolysis and the tricarboxylic acid (TCA) cycle and the connected anabolic pathways that provide precursors for macromolecule synthesis.

Via glycolysis, glucose is broken down to pyruvate, leading to the cellular energy equivalent ATP generation. Glucose can also enter the pentose phosphate pathway or glycogenesis pathway for carbon storage and the production of NADPH, nucleotide sugars, and ribose-5-phosphate; in turn, these metabolites support a variety of macromolecule biosynthesis, anti-oxidant production, and protein glycosylation pathways. Other glycolytic metabolites, such as glycerol-3-phosphate and 3-phosphoglycerate, can be diverted into the fatty acid and serine-glycine biosynthesis pathways, respectively. Pyruvate generated through glycolysis can be further metabolized to lactate and alanine. In addition, pyruvate can enter the TCA cycle via conversion to oxaloacetate or acetyl-CoA, which is crucial for the biosynthesis of fatty acids, amino acids, and ATP.

Amino acids can be synthesized within the cell while others, termed essential amino acids, cannot be derived from food – are indispensable for nucleotide and protein synthesis. Unsurprisingly, owing to the broad physiological range of available extracellular amino acids, some amino acids are catabolized by cells while others are synthesized, depending on the cell type, metabolic state, and microenvironment. Glutamine is the most prominent example of catabolized amino acid. In addition, glutamine can be converted into several other amino acids, such as proline and aspartate, used for fatty acid synthesis, or fully oxidized via glutaminolysis, yielding ATP and NADPH.

Like essential amino acids, essential fatty acids derived from food can be taken up by cells and further modified for cell membranes and signaling molecules. Furthermore, the integration of fatty acid β -oxidation (FAO) and the TCA cycle and the electron transport chain produces the important metabolic cofactor acetyl-CoA, NADPH, and ATP. In addition to the energy, redox homeostasis, and macromolecules resulting from these metabolic conversions, the intermediary metabolites generated regulate cellular signaling and the epigenome.

Cancer cells rely on the same metabolic networks mentioned above; however, various central metabolism pathways can be dysregulated in cancer cells depending on their genetic landscape, cellular origin, microenvironment, and functional phenotype. Targeting vulnerabilities of the dysregulated metabolic pathways in cancer cells is, therefore, an attractive therapeutic strategy.

Notably, the metabolic machinery and nutrient-sensing mechanisms can regulate anti-cancer immune responses. Emerging evidence indicates that cancer cells can suppress antitumor immunity by depleting essential nutrients or reducing the metabolic fitness of tumor-infiltrating immune cells. Thus, metabolic interventions hold promise for improving the effectiveness of immunotherapies. Notably, the similar metabolic needs of cancer cells and immune cells might preclude synergistic effects of such combinations. However, much potential lies in targeting the metabolic pathways that are differentially essential to cancer cells and immune cells and, in particular, those modulated by cancer cells to evade immunosurveillance.

Glycolysis and lactate production

The aberrant bioenergetic activity enables tumor cells to use large amounts of glucose and produce lactic acid via glycolysis even in the presence of adequate oxygen (aerobic glycolysis), with a correspondingly low rate of oxidative phosphorylation (OXPHOS), is a phenomenon known as the Warburg effect. Lactic acid is exported into the extracellular microenvironment via monocarboxylate transporters (MCTs), in particular, monocarboxylate transporter 4 (MCT4), which results in an acidic tumor microenvironment (TME). Both aerobic glycolysis and the resulting acidification of the TME strongly influence T cell-mediated antitumor immune responses and tumor-infiltrating myeloid cell activities. As a result of high glucose consumption rates by tumor cells, tumor-infiltrating lymphocytes (TILs) have decreases in mTOR activity, nuclear factor of activated T cells (NFAT) signaling, and glycolytic capacity, which lead to impaired production of antitumor effector molecules (Chang et al., 2015, Ho et al., 2015a). Moreover, glycolytic activity in tumors can stimulate the expression of granulocyte colony-stimulating factor (G-CSF) and granulocyte-macrophage colony-stimulating factor (GM-CSF) and thereby contribute to the recruitment of myeloid-derived suppressed cells (MDSCs) to the TME (Li et al., 2018b).

Furthermore, glucose deprivation and the accumulation of lactic acid in the TME interrupt the metabolic rewiring and signaling cascades that facilitate dendritic cell (DC) maturation and the pro-inflammatory polarization of macrophages, thus promoting the development of pro-tumorigenic myeloid cells, including tolerogenic DCs and M2-like macrophages (Dietl et al., 2010, Colegio et al., 2014, Tannahill et al., 2013, Gottfried et al., 2006). Interestingly, mice harboring tumors deficient in lactate dehydrogenase A (LDHA), which converts pyruvate to lactate, have a decreased frequency of splenic MDSCs and an improved cytotoxic function of tumor-infiltrating natural killer (NK) cells (Husain et al., 2013), suggesting that production by tumor cells might contribute to tumorigenesis by promoting IL-23-mediated and IL-17-mediated inflammation (Shime et al., 2008). Thus, in addition to modulating immune responses, lactate produced by cancer-associated fibroblasts (CAFs) can be utilized by tumor cells as a surrogate nutrient source (Faubert et al., 2017, Hui et al., 2017).

In agreement with findings from murine tumor models, the results of several clinical studies revealed that aerobic glycolytic in human tumors is negatively associated with host antitumor immune responses and therapeutic outcomes of anti-cancer immunotherapy. For example, human tumors refractory to adoptive T cell transfer (ACT) immunotherapy have elevated levels of aerobic glycolytic activity, and glycolytic tumors have lowered T cell tumor infiltration levels cytotoxicity compared with less-glycolytic tumors (Cascone et al., 2018). Furthermore, LDHA-mediated lactic acid production suppresses IFN γ expression in tumor-infiltrating T cells and NK cells, promoting tumor growth and immune evasion in mouse models (Brand et al., 2016); this suggests that glycolytic activity not only provides an intrinsic growth advantage for tumor cells but also has tumor cell-extrinsic effects that abrogate immunosurveillance of cancer. Hence, targeting glucose metabolism and lactic production and secretion is an appealing strategy for anti-cancer therapy; however, such approaches, particularly those targeting shared glycolytic pathways that support T cell function, might simultaneously blunt immune responses. Conversely, targeting the glycolytic pathway might suppress tumor-promoting inflammation mediated by IL-17, IL-6, and IL-23, thus restricting tumorigenesis (Shime et al., 2008). Therefore, reconsideration of metabolic approaches to anti-cancer therapy is required to ensure that effective antitumor immunity is sustained and explore whether targeting the glycolytic pathway at different stages of tumorigenesis leads to distinct therapeutic responses.

PKM2, an enzyme that converts phosphoenolpyruvate into pyruvate during the final step of glycolysis, is often expressed at high levels in tumor cells. Interestingly, PKM2 is less active than PKM1 in converting phosphoenolpyruvate into pyruvate, supporting the Warburg effect and tumor cell survival and proliferation (Christofk et al., 2008, Elf and Chen, 2014). Correspondingly, PKM2 activators, such as TEPP-46, DASA-58, and ML-265, decrease tumor cell proliferation and tumor growth in mouse models by increasing the conversion of phosphoenolpyruvate into pyruvate (Kung et al., 2012). Intriguingly, PKM2 has been demonstrated to promote expression of programmed cell death 1 ligand 1 (PD-L1), a ligand of the inhibitory T cell immune-checkpoint receptor programmed cell death 1 (PD-1), in tumor and immune cells, and, accordingly, increasing PKM2-mediated phosphoenolpyruvate conversion into pyruvate using TEPP-46 reduces the expression of PD-L1 in tumor and myeloid cells in a mouse CT26 colon carcinoma model (Palsson-Mcdermott et al., 2017). Therefore, PKM2 activators might synergize with PD-1-PD-L1 immune-checkpoint inhibitors (ICIs) by simultaneously reducing metabolic stress and immunosuppression in the TME via abrogation of both aerobic glycolysis and PD-L1 expression in tumor cells (as well as suppressive immune cells).

Phosphofructokinase-2/fructose-2,6-biphosphatase 3 (PFKFB3) promotes glycolytic activity and lactic acid production in tumor cells (Li et al., 2018a). Inhibitors of PFKFB3 have been shown to abrogate the Warburg effect, tumor progression, and metastasis in preclinical models (Li et al., 2017a). Similar to treatment with a PKM2 activator, the PFKFB3 inhibitor PFK-158 has been reported to improve therapeutic responses to antibodies targeting the inhibitory immune-checkpoint receptor cytotoxic T lymphocyte antigen 4 (CTLA-4) in a mouse B16 melanoma model (Chesney et al., 2016). The precise mechanisms by which glycolysis inhibitors induce synergistic responses with ICIs remain delineated, although these findings suggest that modulation of cancer metabolism can unleash host antitumor immune responses reprogramming of the TME.

Inhibition of lactic acid production, and thus the associated acidification of the TME, has also been proposed as a strategy to unleash antitumor immunity. Indeed, LDHA inhibitors, such as FX11 and galloflavin, have been reported to reduce tumor growth in xenograft models (Yang et al., 2010). Whether LDHA inhibitors could be used to enhance immunotherapy remains to be determined; however, treatment of melanoma cells with the LDHA inhibitor GSK2837808A markedly increased cytotoxicity mediated by autologous TILs in an in vitro culture assay (Cascone et al., 2018). Moreover, low serum levels of LDH are associated with better therapeutic responses to the anti-PD-1 antibody pembrolizumab in patients with melanoma (Weide et al., 2016). These findings support the notion that the efficacy of anti-cancer immunotherapy can be enhanced by reducing the production of lactic acid. In addition to LDHA, the high levels of lactate transporters (MCT proteins) in tumor cells provide therapeutic opportunities (Murray et al., 2005). Alternatively, neutralizing the acidity of the TME with bicarbonate has been demonstrated to increase T cell infiltration and improve antitumor immune responses when combined with immune-checkpoint inhibition and ACT in multiple mouse models (Pilon-Thomas et al., 2016).

Amino acids and their derivatives

Glutamine and glutamate. The metabolic demands of tumor cells can also be fuelled through upregulation of glutamine anaplerosis – via glutaminolysis to glutamate – of the TCA cycle intermediate α -ketoglutarate (α -KG). Notably, lactate can promote the expression of both the glutamine transporter ASCT2 and glutaminase 1 (GLS) in tumor cells via stabilization of hypoxia-inducible factor 2 α (HIF2 α), which could potentially reprogramme tumor cells towards increased glutaminolysis (Pérez-Escuredo et al., 2016). Furthermore, increased glutamine anaplerosis in tumor cells leads to increased ammonia release; exposure to ammonia can activate autophagy in neighboring cells, such as CAFs. Intriguingly, ammonia-activated autophagy in CAFs has been suggested to further support tumor cell growth by facilitating the release of glutamine from CAFs, which can then be metabolized by tumor cells (Ko et al., 2011). In addition, the products of glutamine metabolism – glutamate and α -KG and aspartate – can, in turn, modulate cellular metabolism, epigenetic landscapes, nucleotide synthesis, and redox balance in tumor cells. Thus, multiple compounds targeting glutamine anaplerosis have been developed as anti-cancer treatments. Among these agents, the GLS inhibitors BPTES (bis-2-(5-phenylacetamido-1,2,4-thiadiazol-2-yl)ethyl sulfide 3) and compound 968 have been shown to prolong survival in several xenograft tumor models by inhibiting cell proliferation and eliciting cell death (Xiang et al., 2015). Unfortunately, the therapeutic utility of BPTES and compound 968 is limited by their moderate potency, poor metabolic stability, and low solubility (Gross et al., 2014). By contrast, CB-839, an allosteric inhibitor of GLS, effectively inhibits

glutaminolysis and has promising activity in preclinical models of triple-negative breast cancer and hematological malignancies (Gross et al., 2014).

In addition to tumor cells, activated T cells and macrophages upregulate glutamine metabolism to support cell fate determination and immune responses (Tannahill et al., 2013, Carr et al., 2010, Liu et al., 2017). Furthermore, glutamine deprivation can suppress T cell proliferation and cytokine production (Carr et al., 2010). Nevertheless, glutamine restriction during T cell activation in vitro has promoted memory CD8⁺ T cell differentiation (Nabe et al., 2018). Furthermore, genetic ablation of GLS expression has also been demonstrated to promote the differentiation and effector function of CD4⁺ T helper 1 (T_H1) cells, and CD8⁺ cytotoxic T lymphocytes (CTLs) impairs differentiation of T_H17 cells (Johnson et al., 2018). Thus, interventions targeting glutamine metabolism in tumor cells are postulated to simultaneously affect the immune state of the TME and antitumor immunity. In support of this hypothesis, CB-839 has been shown to synergize with PD-1 inhibitors in several clinical trials (Tannir et al., 2018); however, which immune cells are responsible for this synergistic antitumor activity remains to be determined. Transient CB-839 treatment augments CTL-mediated antitumor responses in mouse models (Johnson et al., 2018), which might explain the synergy between this agent and PD-1 inhibitors and supports the therapeutic potential of GLS inhibition in anti-cancer immunotherapy. Of note, glutamine deprivation has also been shown to hamper T_H1 cell differentiation in vitro but favors regulatory T_{reg} cell development in differentiation cultures (Klysz et al., 2015).

These observations of opposite responses in different T cell subsets highlight the need for further investigations to delineate how glutamine metabolism is modulated in the TME and the underlying mechanisms by which glutamine orchestrates T cell responses. In addition to glutamine, glutamate levels can also fine-tune T cell proliferation and cytokine production. Upon T cell receptor (TCR) activation, CD4⁺ and CD8⁺ T cells upregulate glutamate receptors, which correlates with increased expression of activation molecules and production of IFN γ , with evidence of co-stimulatory effects mediated by these voltage-gated potassium channels (Shanker et al., 2018). Conversely, high concentrations of extracellular glutamate can suppress T cell activation (Poulopoulou et al., 2005). Whether glutamate-mediated signaling can be manipulated in order to enhance anti-cancer immunotherapy remains unclear.

Arginine. Arginine metabolism also has vital roles in T cell activation and modulating immune responses. During the resolution of inflammatory responses, immunomodulatory cells promote arginine degradation via the catabolic enzyme arginase 1 (ARG1) expression. The accumulation of ARG1-expressing immunomodulatory cells, including tolerogenic DCs, M2-like tumor-associated macrophages (TAMs), and T_{reg} cells, in the TME, could suppress antitumor immunity by degrading arginine and thereby limiting the availability of this amino acid to T cells. Accordingly, supplementation of arginine stimulates T cell and NK cell cytotoxicity and effector cytokine production and, in combination with anti-PD-L1 antibody treatment, enhances antitumor immune responses and prolongs the survival of osteosarcoma-bearing mice (He et al., 2017).

Furthermore, arginine supplementation during in vitro expansion of T cells facilitates differentiation to central memory-like T cells with superior antitumor activity (Geiger et al., 2016). Thus, replenishment of arginine and prevention of arginine degradation in the TME are attractive strategies to re-invigorate T cell-mediated and NK cell-mediated immune responses. These approaches are currently being tested in a clinical trial in which the ARG1 inhibitor INCB001158 is being used in combination with

the ICI pembrolizumab. In mouse tumor models, INCB001158 treatment increases CD8⁺ T cell and NK cell tumor infiltration and stimulates inflammatory cytokines in the TME (Steggerda et al., 2017).

Treatment with PEGylated arginine deiminase (ADI-PEG 20) to deplete the TME of arginine has been shown to suppress the growth of arginine auxotrophic cancers (Qiu et al., 2014, Kelly et al., 2012, Tsai et al., 2017). Intriguingly, ADI-PEG 20 has been reported to enhance T cell activation and tumor T cell infiltration, moderate T cell exhaustion, and abolish T_{reg} cell accumulation in tumors (Brin et al., 2017). These immune phenotypes form the foundation of ongoing clinical trials in which ICIs are combined with ADI-PEG 20. However, the tumor types or immune TMEs that are most suited to treatment with an ARG1 inhibitor or ADI-PEG 20 remain unclear.

Tryptophan. The resolution of inflammation is also mediated by tryptophan metabolism via enzymes, including indoleamine 2,3-dioxygenase (IDO) and tryptophan 2,3-dioxygenase (TDO). Expression of tryptophan-degrading enzymes in tumor cells facilitates tumor progression and is associated with a worse prognosis in patients with gastric adenocarcinoma (Liu et al., 2016). Furthermore, various stromal cells in the TME, including endothelial cells, TAMs, and DCs, also overexpress IDO and TDO. Thus, high levels of TDO and IDO in tumors may decrease tryptophan availability in the TME, suppressing T cells' tumoricidal functions (Munn et al., 2005, Li et al., 2009). Aside from depriving T cells of tryptophan, TDO and IDO catabolize tryptophan to kynurenine, the accumulation of which can promote increases in the number of peripheral T_{reg} cells and reduce the proliferation of effector T cells (Cronin et al., 2018). IDO expression in macrophages also supports their anti-inflammatory and phagocytic activities by producing kynurenine that fuels the synthesis of NAD⁺ (Minhas et al., 2019). This finding suggests that inhibiting IDO might alleviate the M2-like phenotype of TAMs. Intriguingly, findings in a preclinical model of breast cancer demonstrate that therapeutic vaccination with tumor antigen-loaded, IDO-silenced DCs increases the proliferation and cytotoxic activity of antigen-specific T cells and reduces the abundance of T_{reg} cells, as compared with an IDO-expressing DC vaccine (Zheng et al., 2013). Furthermore, systemic IDO inhibition promotes tumor regression by increasing the production of cytokines, including IL-12 and IFN γ , and tumor infiltration of T cells and neutrophils in mouse metastatic liver tumor and bladder tumor models (Yen et al., 2009, Huang et al., 2011).

The oncometabolite 2-hydroxyglutarate. Many gliomas express neomorphic mutant forms of isocitrate dehydrogenase 1 (IDH1) or IDH2 that support tumorigenesis. *IDH1* mutation has been detected in 80% of patients with grade II-III gliomas and secondary glioblastomas (Yan et al., 2009). In addition, *IDH2* mutations are detected in gliomas, although much less common than *IDH1* mutations (Yan et al., 2009). *IDH1* or *IDH2* mutations are also detected in ~20% of patients with AML. Whereas wild-type IDH enzymes convert isocitrate to α -KG, the *IDH* mutations cause a change in enzymatic activity that results in the conversion of α -KG to the oncometabolite D-2-hydroxyglutarate (D-2-HG), accumulation of which alters epigenetic regulation of gene expression and contributes to mTOR activation (Carbonneau et al., 2016). The discovery of *IDH* mutations and their role in oncogenesis has led to the development of novel therapeutic strategies predicated on either inhibiting mutant IDH or restoring wild-type IDH function to suppress the production of D-2-HG.

Itaconate

Itaconate, a derivative of citrate, is produced by immune-responsive gene 1 (IRG1)-mediated decarboxylation of the TCA cycle intermediate *cis*-aconitate in mitochondria. Stimulation with lipopolysaccharide strongly promotes itaconate production in macrophages, which reduces the production of pro-inflammatory cytokines via inhibition of succinate dehydrogenase and activation of nuclear erythroid 2-related factor 2 (NRF2) and activating transcription factor 3 (ATF3) (Lampropoulou et al., 2016, Mills et al., 2018, Bambouskova et al., 2018). These findings reveal that itaconate orchestrates metabolic and transcriptomic programs favoring M2-like macrophage phenotypes. The phenotypes of macrophages *in vivo* are more varied and complex than *in vitro*-polarized macrophages (Xue et al., 2014), but M2-like phenotypes have generally been shown to support tumor progression. Thus, the production of itaconate in TAMs might confer them with M2-like phenotypes.

Adenosine signaling

The concentration of adenosine in tissues significantly increases within several hours following tissue injury and TME and the hypoxic tissues. The ectonucleotidases CD39 and CD73 are cell surface molecules with essential roles in regulating the production of adenosine via the catabolism of ATP to AMP and AMP to adenosine, respectively. The resulting extracellular adenosine can bind to four G protein-coupled purinergic type 1 receptors (adenosine receptor A₁, (A₁R), A_{2A}R, A_{2B}R, or A₃R), which activates the PKA signaling cascade by facilitating adenylyl cyclase-mediated production of cAMP. In particular, activation of A_{2A}R and A_{2B}R is associated with profound immunosuppression during inflammatory diseases. Moreover, elevated expression of CD39 and CD73 in tumors is associated with poor prognosis in cancer patients (Inoue et al., 2017, Turcotte et al., 2015). Aside from tumor cells, T_{reg} cells can express CD39 and contribute to immunosuppression in the TME via the adenosine-A_{2A}R signaling axis (Maj et al., 2017). T_{reg} cells have also been shown to support the dissemination of melanoma cells in mouse models via CD39-dependent abrogation of NK cell-mediated antitumor immune responses, whereas the ectonucleoside triphosphate diphosphohydrolase inhibitor polyoxometalate-1 suppressed tumor growth (Sun et al., 2010). M2-like macrophages also express CD39 and CD73 and thus exert anti-inflammatory functions via the A_{2A}R and A_{2B}R signaling pathways (Csóka et al., 2012). In T cells, adenosine signaling via A_{2A}R inhibits the expression of the IL-2 receptor and TCR-stimulated proliferation and promotes inhibitory immune-checkpoint receptors (including PD-1 and CTLA-4) (Allard et al., 2013, Ohta et al., 2012), thereby impeding T cell effector function and probably antitumor immunity.

Furthermore, A_{2A}R signaling stimulates the expression of PD-L2 and IL-10 in DCs (Li et al., 2012), which might increase the capacity of these DCs to suppress T cell antitumor responses. Adenosine signaling via A_{2A}R also facilitates the accumulation of MDSCs in mouse tumors and their production of VEGF, and, accordingly, pharmacological inhibition of A_{2A}R reduced angiogenesis and increased T cell accumulation in the TME (Sorrentino et al., 2015). Thus, targeting CD39 and CD73 activity to inhibit adenosine production is an attractive strategy for enhancing antitumor immunity. Indeed, *Cd39*-deficient mice have decreased pulmonary metastasis in the melanoma engraftment model compared with their *Cd39*-wild-type counterparts (Turcotte et al., 2015, Merighi et al., 2002). In addition, pharmacological inhibition or genetic ablation of CD73 reduced the migratory capacity of breast cancer cells in mice in an adenosine-dependent manner (Zhi et al., 2010). Several preclinical studies have revealed that treatment with antagonistic anti-CD73 antibodies improves the outcomes of ICI therapy in preclinical models (Allard

et al., 2013, Iannone et al., 2014). Similarly, combined treatment with A_{2A}R antagonists and ICIs can also elicit synergistic antitumor responses in mouse models (Beavis et al., 2015, Mittal et al., 2014).

The cyclooxygenase and PGE₂ pathway

Prostaglandin E₂ (PGE₂) is a bioactive lipid metabolite derived from cyclooxygenase-mediated arachidonic acid metabolism that elicits a wide range of biological effects associated with inflammatory diseases. In contrast to cyclooxygenase 1 (COX1), which is constitutively expressed in non-malignant tissues, COX2 is overexpressed in numerous cancers and strongly associated with immunosuppression and production of a high level of PGE₂ in the TME (Zelenay et al., 2015). Preclinical studies have revealed that PGE₂ overproduction promotes the development and differentiation of T_{reg} cells (Sharma et al., 2005, Baratelli et al., 2005, Mahic et al., 2006), inhibits IL-2 and IFN γ production in human T cells (Snijdwint et al., 1993), and skews activated T cells towards a phenotype the generation of high levels of anti-inflammatory cytokines, including IL-4, IL-10, and IL-13 (Demeure et al., 1997). PGE₂ also promotes M2-like differentiation of TAMs (Larsson et al., 2015) and the immunosuppressive functions of MDSCs. In mouse models, the abundance of PGE₂ in the TME further impedes T cell infiltration by abrogating NK cell-mediated recruitment of conventional type I DCs, thus contributing to cancer immune evasion (Böttcher et al., 2018). Moreover, PGE₂ signaling has been shown to inhibit CTL survival and function (Chen et al., 2015). Together, these findings suggest that inhibition of PGE₂ production and signaling cascades could improve multiple facets of the antitumor immune response.

In support of this theory, aspirin (acetylsalicylic acid) inhibits COX1 and COX2 and thus PGE₂ biosynthesis, with considerable evidence supporting the potential of this agent to suppress tumorigenesis. Celecoxib, a selective COX2 inhibitor, induces synergistic antitumor immune responses when combined with anti-PD-1 antibody therapy in mouse tumor models (Li et al., 2016c). Clinical trials of COX inhibitors in combination with ICIs have been initiated. In addition to antagonizing the activity of COX2, extensive effort has been devoted to elucidating PGE₂ receptors for use in anti-cancer therapy. Notably, expression of PGE₂ receptor (EP4) is associated with poor prognosis in patients with cancer and suppressive features of the TME (Chell et al., 2006, Buchanan et al., 2006).

Fatty acids and cholesterol

Tumor cells often have increased de novo fatty acid synthesis rates to divert energy production into anabolic pathways to generate plasma membrane phospholipids and signaling molecules. Moreover, adipocytes and adipocyte-derived fibroblasts can be identified in the TME and contribute to the high lipid content of the TME (Bochet et al., 2013, Zhang et al., 2012). Lipid accumulation in tumor-infiltrating myeloid cells, including DCs, MDSCs, and TAMs, has been shown to veer these immune cells towards anti-inflammatory and immunosuppressive phenotypes via metabolic reprogramming (Herber et al., 2010, Cubillos-Ruiz et al., 2015, Al-Khami et al., 2017, Niu et al., 2017). CD8⁺ TILs with high levels of PD-1 expression isolated from patients with non-small-cell lung carcinoma also have a high lipid content compared with CD8⁺ TILs with lower or no PD-1 expression (Thommen et al., 2018). In addition, CD8⁺ TILs with high levels of PD-1 expression isolated from patients with non-small-cell lung carcinoma also have a high lipid content compared with CD8⁺ TILs with lower or no PD-1 expression. Notably, however, these

PD-1-high TILs had a higher capacity to recognize tumor cells and were predictive of favorable survival after treatment with PD-1 inhibitors (Thommen et al., 2018). These findings suggest that lipid metabolism in tumor cells and immune cells in the TME has a crucial role in orchestrating immunosuppression and warrant targeting these metabolic pathways to enhance antitumor immunity.

Effective TCR clustering and formation of the immunological synapse are essential for T cell function and are dependent on the lipid composition of cell membranes (Zech et al., 2009). Accordingly, disrupting cholesterol esterification using the sterol *O*-acyltransferase 1 inhibitor avasimibe has been reported to increase the fraction of cholesterol in the plasma membranes of CD8⁺ TILs improve T cell effector function and proliferation (Yang et al., 2016). Of note, avasimibe synergizes with PD-1 inhibitors in eradicating melanoma in mouse models (Yang et al., 2016). Furthermore, drugs that increase FAO through activation of peroxisome proliferator-activated receptor (PPAR) signaling have similar effects (Zhang et al., 2017, Chowdhury et al., 2018). Indeed, the reported lipid accumulation in CD8⁺ TILs could suggest defective utilization of fatty acids, thus warranting the development of approaches to stimulate the antitumor T cell responses by promoting FAO.

Interestingly, different fatty acids drive the differentiation and proliferation of CD4⁺ T cells belonging to specific lineages. Specifically, the long-chain fatty acid lauric acid supports the differentiation of pro-inflammatory T_H1 cells and T_H17 cells, whereas the short-chain fatty acid propionate promotes T_{reg} cells (Haghikia et al., 2015). Thus, the lipid species present in the TME are likely to orchestrate the infiltration pattern of effector CD4⁺ T cells and determine the outcomes of targeting lipid metabolism to treat cancer.

Lipid metabolism programs also differ between M1-like and M2-like macrophages. Specifically, fatty acid synthesis predominates in M1-like macrophages, whereas M2-like macrophages are dependent on FAO to fuel their bioenergetic demands. To date, whether inhibiting FAO or augmenting fatty acid synthesis improves the antitumor activity of macrophages has not been established. Interestingly, limiting flux through the cholesterol biosynthetic pathway in macrophages induces type I interferon responses that drive antiviral immunity via autocrine and paracrine signaling (York et al., 2015); however, whether the same responses could be exploited to enhance the antitumor activity of macrophages remains unclear. In contrast to type I interferon responses, inhibiting ATP-binding cassette transporter G1 (ABCG1), which mediates cholesterol secretion, shifted macrophages from an M2-like towards an M1-like phenotype, thereby increasing their capacity to kill cancer cells in vitro (Sag et al., 2015). Thus, how cholesterol metabolism fine-tunes macrophage behavior under different conditions remains to be determined. Finally, macrophages of *Map3k8*-null mice showed impaired M2 polarization, associated with decreased lipid catabolism (Kannan et al., 2016). Hence, interfering with lipid metabolism in macrophages, including inhibition of CD36-mediated lipid uptake and FAO (Huang et al., 2014), might enhance antitumor immunity.

Metabolic programs of trained immunity

Trained immunity is a specialized form of the immune response, in which training stimuli, such as β -glucan, stimulate a long-term enhancement of the activity of innate immune cells through metabolite-orchestrate epigenetic reprogramming. For example, by educating innate immune cells to produce high pro-inflammatory cytokines, training stimuli can protect from various infections. Training stimuli have

been tested in various trials to elicit antitumor responses. These trials are primarily supported by the fact that the engagement of trained immunity through Bacillus Camette-Guerin (BCG) vaccination in patients with non-muscle-invasive bladder cancer (Buffen et al., 2014). Given that innate immune cells have essential roles in forming the immunosuppressive TME, stimulating trained immunity in TAMs and tumor-infiltrating DCs might synergize with other immunotherapies TME, stimulating trained immunity in TAMs and tumor-infiltrating DCs might synergize with other immunotherapies by reprogramming the TME to become more immunostimulatory. Given that metabolites, including α -KG, acetyl-CoA, succinate, fumarate, and NAD⁺, are critical orchestrators of trained immunity (Cheng et al., 2014, Arts et al., 2016), other metabolic interventions discussed herein might synergize with training stimuli in augmenting antitumor immunity.

Metabolic effects of immune checkpoints

ICIs are an outstanding advance in the treatment of cancer. These therapies were initially developed to enhance the signaling pathways for T cell activation; however, emerging evidence indicates that ICIs also affect the metabolic fitness of T cells. Immune-checkpoint ligation or inhibition influences the metabolic communication and competition between tumor and T cells in the TME. For instance, the interaction of PD-1 with PD-L1 or PD-L2 impairs metabolic reprogramming, including upregulation of aerobic glycolysis and glutaminolysis, in T cells via suppression of the PI3K-AKT-mTOR pathway (Freemerman et al., 2014). By contrast, PD-1 signaling promotes FAO in T cells by stimulating AMPK activity and inducing the expression of carnitine palmitoyltransferase 1A (CPT1A), a rate-limiting enzyme of the FAO pathway (Patsoukis et al., 2015). In addition to modulating the metabolic profile of TILs, immune checkpoints can also directly affect metabolism in tumor cells. Expression of PD-L1 and B7-H3 in tumor cells has been shown to stimulate aerobic glycolysis by activating the PI3K-AKT-mTOR pathway (Chang et al., 2015, Lim et al., 2016). Thus, inhibition of the PD-1-PD-L1 axis might have synergistic anti-cancer effects by promoting the reinvigoration and metabolic fitness of TILs while simultaneously suppressing aerobic glycolysis in tumor cells. Consequently, PD-1-PD-L1 inhibition might also increase the amount of glucose available to TILs, which could ameliorate the nutrient stress imposed on TILs by metabolic conditions in the tumor. In support of this hypothesis, a preclinical study has revealed that PD-1-PD-L1 inhibition indeed increases glucose availability in the TME and enhances the glycolytic activity of T cells (Cascone et al., 2018).

Other inhibitory immune-checkpoint receptors have been reported to affect the metabolic programs of T cells. For example, CTLA-4 signaling inhibits CD28-mediated co-stimulation at least in part by reducing AKT phosphorylation and activation and might therefore impair the increased glucose metabolism and mitochondrial remodeling that occurs following T cell activation – similar to the effects of PD-1 signaling. However, in contrast with the PD-1 pathway, CTLA-4 signaling does not augment FAO (Patsoukis et al., 2015). In addition, T cell immunoglobulin mucin receptor 3 (TIM3), which is another inhibitory immune-checkpoint receptor that is highly expressed in dysfunctional exhausted T cells, has also been demonstrated to alter T cell metabolism via interruption PI3K-AKT-mTOR signaling. Moreover, lymphocyte activation gene three protein (LAG3)-deficient CD4⁺ T cells have a substantially increased rate of basal respiration and aerobic glycolysis and excess respiratory capacity compared with wild-type CD4⁺ T cells (Previte et al., 2017), suggesting that LAG3 reduces the metabolic fitness of T cells.

In contrast to the metabolic impairments caused by inhibitory immune-checkpoint receptors, co-stimulatory molecules support T cell activation by stimulating pathways that control transcriptional reprogramming as well as metabolic switches. For example, CD28 signaling enhances the metabolic fitness of T cells by simultaneously stimulating aerobic glycolysis (Jacobs et al., 2008) and facilitating mitochondrial fusion, which enables effective production of acetyl-CoA (Klein Geltink et al., 2017). Similarly, 4-1BB signaling, which strongly enhances CD8⁺ T cell proliferation, activates glucose and fatty acid metabolism (Choi et al., 2017). Furthermore, 4-1BB and OX40 dual co-stimulation augment glycolysis in CD8⁺ T cells due to robust induction of glucose transporters (Tsurutani et al., 2016). In addition, activation of glucocorticoid-induced TNFR-related protein (GITR) upregulates nutrient uptake, lipid storage, glycolysis, and oxygen consumption in CD8⁺ T cells (Sabharwal et al., 2018). Furthermore, inducible T cell co-stimulator (ICOS) co-stimulation can drive glycolysis in activated T cells via activation of the PI3K-AKT-mTOR pathway (Zeng et al., 2016, Gigoux et al., 2014). Taken together, current evidence indicates that ICIs and stimulators of co-stimulatory receptors have a significant impact on T cell metabolism.

Mitochondrial regulation of T cells

TCR stimulation induces mitochondrial biogenesis and remodeling, which are necessary to fulfill the metabolic requirements of T cell activation (Menk et al., 2018). During CD8⁺ T cell differentiation, mitochondrial fusion and fission also instruct metabolic programming via currently undefined signaling pathways (Buck et al., 2016). Furthermore, mitochondria physically associate with the immune synapse early after productive TCR activation, leading to the local generation of ATP and stabilizing the immune synapse by modulating calcium signaling. TCR activation also stimulates the production of mitochondrial reactive oxygen species (ROS), which drives the cell expansion phase of T cell activation (Sena et al., 2013). Intriguingly, most CD8⁺ TILs have an effector memory phenotype. Following cognate antigen stimulation, CD8⁺ effector memory T cells are characterized by a rapid mTOR complex 2 (mTORC2)-Akt-dependent upregulation of aerobic glycolysis (Gubser et al., 2013).

Interestingly, activated mTORC2 and AKT colocalize with and inhibit GSK3 β at mitochondria-endoplasmic reticulum junctions; subsequent recruitment of hexokinase 1 to voltage-dependent anion channels on mitochondria promotes pyruvate oxidation, thereby supporting the metabolic requirements and reprogramming necessary for efficient acquisition of effector function by memory T cells (Bantug et al., 2018). However, TILs are characterized by a decreased mitochondrial mass compared with peripheral blood T cells and subsequently show a limited respiratory capacity (Scharping et al., 2016). Therefore, these metabolic processes might be centered on mitochondrial activity, and dynamics are compromised in the TME, thus abrogating CD8⁺ T cell function. The loss of mitochondrial mass in TILs might result from persistent AKT activation, in turn, leading to progressively decreasing expression of PPAR γ co-activator 1 α (PGC1 α), which co-activates several transcription factors, such as PPAR γ , NRF1/2, and ERR α , and thereby stimulates mitochondrial biogenesis and FAO (Scharping et al., 2016). In experimental models, overexpression (Scharping et al., 2016, Menk et al., 2018) or induction of PGC1 α through 4-1BB stimulation rescues mitochondrial function and avoids metabolic exhaustion of TILs, resulting in enhanced antitumor activity (Siska et al., 2017).

Moreover, combination treatment of tumor-bearing mice with an anti-PD-1 antibody and the pan-PPAR- α agonist bezafibrate increases mitochondrial biogenesis and OXPHOS in CD8⁺ T cells isolated from tumor-draining lymph nodes, in association with increases in the abundance and antitumor activity of effector memory T cells (Chowdhury et al., 2018). By contrast, TILs from patients with RCC have been reported to contain small, punctate, fragmented mitochondria with poorly defined membranes and cristae and an increased inner membrane mass (Siska et al., 2017). In addition, these mitochondria are hyperpolarized, leading to excessive production of ROS that suppress T cell antitumor function. Indeed, in vitro treatment with scavengers of mitochondrial ROS improves the activation and proliferation of TILs from patients with RCC.

Despite the evidence that the augmentation of mitochondrial biogenesis and fitness improves the antitumor function of T cells (Scharping et al., 2016, Menk et al., 2018, Siska et al., 2017), whether this strategy stimulates OXPHOS in TILs remains unknown. Notably, the hypoxic conditions of the TME might limit the oxidative capacity of TILs and, in combination with low glucose levels, could force TILs to rely on alternative energy sources. Indeed, compared with their counterparts in the circulation, CD8⁺ TILs have increased expression of the transcription factor PPAR α and its target genes involved in fatty acid uptake, triglyceride turnover, and peroxisomal and mitochondrial FAO (Zhang et al., 2017). Furthermore, enhancement of this metabolic signature using a PPAR α agonist enables TILs to maintain antitumor activity in the TME despite being deprived of glucose and oxygen (Zhang et al., 2017). Autophagy in CD8⁺ T cells also has a role in providing lipid substrates necessary for FAO, thus supporting memory CD8⁺ T cell differentiation and survival.

Similarly, autophagy of mitochondria has been found to facilitate the survival of memory CD4⁺ T cells, mainly by limiting the toxic effects of mitochondrial activity and lipid overload (Murera et al., 2018). Moreover, this selective autophagy of mitochondria, known as mitophagy, is also crucial for the clearance of dysfunctional mitochondria in NK cells and the induction of trained NK cell-mediated immunity upon viral infection (O'Sullivan et al., 2015). Thus, intact mitochondrial function is essential to mounting an efficient antitumor response. Investigations are therefore warranted to clarify whether autophagy or mitophagy in TILs contributes to supplying lipids for FAO or to mitochondrial homeostasis and whether pharmacological activation of these processes could enhance TIL metabolic fitness, survival, and antitumor activity.

Metabolic interventions in ACT

In vitro autologous TILs. ACT of in vitro-expanded autologous TILs has successfully been used to treat cancer patients. The isolated TILs are, however, terminally differentiated and might therefore have limited long-term activity. Interestingly, the infusion of T cells with a self-renewing memory phenotype confers a more robust and more sustained antitumor response in mouse models (Klebanoff et al., 2011). Moreover, exposure to IL-15 during in vitro culturing polarizes tumor-reactive CD8⁺ T cells to a central memory phenotype associated with more potent antitumor activity after adoptive transfer to tumor-bearing mice (Klebanoff et al., 2004). IL-15 actively drives a metabolic shift towards oxidative metabolism via mitochondrial biogenesis and expression of the key regulatory enzyme of FAO – i.e., CPT1A (van der Windt et al., 2012). In addition, IL-15-generated memory CD8⁺ T cells have a fused mitochondrial network compared to the punctate mitochondria of effector cells (Buck et al., 2016). Accordingly, cultured CD8⁺ T

cells with the promoter of mitochondrial fusion M1 and the inhibitor of mitochondrial fission mDivi1 have shown potential as a therapeutic strategy to drive T cells memory phenotype and thereby improve the efficacy of ACT immunotherapy (Buck et al., 2016).

An early observation that mTORC1 inhibition with rapamycin leads to increased memory T cells after viral clearance, indicating a crucial role of metabolic modulation in activated T cells during a primary immune response (Araki et al., 2009). In this study, silencing of the mTORC1 component Raptor phenocopied the effect of rapamycin (Araki et al., 2009). Alternatively, inhibition of mTORC2-AKT signaling or glycolysis (the metabolic signature of effector CD8⁺ T cells) during the expansion of CD8⁺ T cells in vitro can also confer the cells a memory phenotype and increased antitumor activity (Crompton et al., 2015, Zhang et al., 2016). Interestingly, naïve and tumor-reactive TILs isolated from patients can be metabolically modified with AKT inhibitors during in vitro expansion, resulting in increased antitumor activity and a memory-like phenotype allogeneic transplantation into immunodeficient, multiple myeloma-bearing mice (Crompton et al., 2015). These data encourage efforts to integrate such metabolic interventions into current clinical protocols for ACT immunotherapy.

A more precise understanding of the processes inducing metabolic T cell exhaustion in the TME might also reveal new therapeutic targets. For example, glucose deprivation in the TME might cause phosphoenolpyruvate (PEP) insufficiency in TILs (Ho et al., 2015b). Correspondingly, increasing PEP production in melanoma antigen-specific T cells via overexpression of phosphoenolpyruvate carboxykinase 1 (PCK1) suppressed the activity of sarcoplasmic/endoplasmic reticulum ATPase 3 (SERCA), thereby enabling sustained intratumoral TC-mediated calcium-NFAT signaling and T cell effector function upon ACT into B16 melanoma-bearing mice (Ho et al., 2015b). In addition, tumors can also contain high TILs consequently have higher intracellular potassium levels, which inhibits TCR-driven AKT-mTOR signaling and antitumor activity (Eil et al., 2016). Accordingly, increasing potassium efflux via overexpression of the potassium channel K_v1.3 in tumor antigen-specific CD8⁺ T cells during in vitro expansion can enhance ACT outcomes in melanoma-bearing mice (Eil et al., 2016).

Chimeric antigen receptor T cell therapy. ACT using autologous T cells genetically modified to express chimeric antigen receptors (CARs) targeting a specific tumor antigen is a promising therapeutic strategy, with clinical successes resulting in regulatory approvals: the anti-CD19 CAR T cell products tisagenlecleucel and axicabtagene ciloleucel are indicated for the treatment of selected patients with B cell acute lymphoblastic leukemia (B-ALL) or large B cell lymphoma. Similar to the effects observed with an ACT of TILs (Crompton et al., 2015), AKT inhibition during the ex vivo expansion of anti-CD19 CAR T cells has been shown to alter their metabolism, increase their differentiation towards a memory phenotype and improve their therapeutic activity against B-ALL in immunodeficient mice (Klebanoff et al., 2017). Likewise, treatment of CAR T cells with a PI3K inhibitor in vitro resulted in less-differentiated cells with enhanced in vivo persistence and antitumor activity in mice (Zheng et al., 2018). These findings are in keeping with the roles of AKT-mTOR signaling in promoting a terminally differentiated effector phenotype and enhancing glycolytic flux upon T cell activation (Frauwirth et al., 2002). Indeed, inhibition of mTOR or the glycolytic pathway (using 2-deoxyglucose) also favors T cell differentiation towards naïve and memory phenotypes, albeit with a dramatic reductive effect on cell proliferation (Araki et al., 2009, Sukumar et al., 2013), thereby potentially limiting the utility of such metabolic interventions in the context of in vitro CAR T cell expansion. Intriguingly, PI3K inhibition skews T cell differentiation towards naïve and memory phenotypes without suppressing CAR T cell proliferation, suggesting that memory T cell

differentiation does not always reduce proliferative capacity (Zheng et al., 2018). However, the mechanisms underlying these disparate effects of inhibiting PI3K versus mTOR or glycolysis remain to be determined.

In addition to pharmacological intervention, the co-stimulatory domain used in the CAR construct has also been shown to elucidate metabolic fitness and persistence of the resulting T cell product. For example, the inclusion of a CD28 domain stimulates CAR T cell glycolysis and effector differentiation, using a 4-1BB co-stimulatory domain induces mitochondrial biogenesis, OXPHOS, and subsequent memory T cell differentiation, thus resulting in better in vivo persistence (Kawalekar et al., 2016). Together, these findings suggest that immunometabolism is a crucial determinant of CAR T cell therapy outcomes and other ACT approaches. Notably, the clinical protocols of ACT therapy are well suited to rapid elucidation and manipulation of the crucial metabolic machinery of T cells owing to the ability to study candidate compounds during the in vitro T cell engineering and expansion phases.

Targeting cancer and immune cell metabolism can synergize with immunotherapy. Understanding and harnessing metabolic crosstalk in the TME can increase the often low response rates achieved with immunotherapies. While various combinations of metabolic agents and immunotherapies are already in clinical trials, efforts to elucidate the metabolic signatures of tumor immune evasion and the metabolic demands of immune cells are crucial to exploit combination therapies fully. Notably, metabolic programs also influence antigen presentation and recognition. Thus, metabolic interventions might improve immune cell responses against highly immunogenetic cancers and increase the immunogenicity of cancer cells, thus broadening the spectra of cancers that can be effectively treated with immunotherapy.

Common bulk metabolic analyses

Extracellular flux analysis

Extracellular flux analyzers, such as the Seahorse apparatus, perform real-time recordings of the extracellular acidification rate (ECAR) and oxygen consumption rate (OCR) as essential readouts of metabolism, and as such, provide indirect measurements of glycolysis and mitochondrial respiration, respectively. While the use of specific substrates and inhibitors allows one to investigate the relative preference of a cell for glucose, glutamine, or fatty acids, extracellular flux analysis does not yield detailed information about the activity of metabolic pathways outside of glycolysis and the TCA cycle within the mitochondria. Nevertheless, this technology allows the easy, fast, and reasonably affordable profiling of cells in a 96-well format strongly contributed to expanding the immunometabolism field (Pelgrom et al., 2016, Au - Van den Bossche et al., 2015, van der Windt et al., 2016). Some of the key findings obtained using the Seahorse instrument include the glycolytic switch and mitochondrial dysfunction in inflammatory macrophages, increased glycolysis in effector T cells and activated dendritic cells (DCs), and the heightened mitochondrial respiration in T memory cells and IL-4-activated macrophages (Everts et al., 2014, Huang et al., 2014, Jan et al., 2016). Thus, ECAR is a surrogate marker for glycolysis. However, acidification may not always result from enhanced glycolysis. Indeed, CO₂ production within respiring

mitochondria or extracellular release of TCA cycle intermediates (e.g., succinate) also acidifies the culture medium. This limitation can be overcome by performing targeting metabolomics or fluxomics for the glycolysis pathway.

Steady-state metabolomics

For a given time point, metabolomics measures steady-state levels of a broad range of metabolites via liquid or gas chromatography-mass spectrometry (LC-MS or GC-MS). While untargeted metabolomics measures hundreds of metabolites in biological samples, targeted metabolomics assesses pre-selected metabolites and yields data with higher sensitivity and allows for accurate absolute quantification of metabolite concentrations. In addition to measuring small metabolites, high-resolution MS-based profiling has been recently applied in “shotgun” lipidomics to profile the lipidome of differentially activated macrophages (Dennis et al., 2010, Hsieh et al., 2020). Most MS-based approaches still require relatively high levels of input material (~100k cells), yet recent developments provide an avenue to measure metabolites at a single-cell resolution spatially.

Fluxomics

Altered flow through a given pathway does not necessarily correlate directly with metabolite levels within this pathway as measured by metabolomics profiling at a set time point. The use of stable isotope ^{13}C -labeled or ^{15}N -labeled substrates and measurement of isotopologues through targeted MS at serial time points allows one to determine metabolic fluxes and pathway dynamics. While metabolomics provides a “snapshot” of the metabolites present at one particular moment, $^{13}\text{C}/^{15}\text{N}$ label-tracing (or so-called fluxomics) could be regarded as a “movie” that yields insight into the relative flux through distinct metabolic paths. The use of such analyses is still in its infancy at the single-cell resolution. However, the community can be readily used to quantify the rates of specific metabolic reactions in bulk approaches such as the mass isotopomer multi-ordinate spectral analysis (MIMOSA fluxomics) (Alves et al., 2015).

Key immunometabolism concepts derived from bulk analyses

Glycolysis supports diverse immune effector functions

Immune activation is an energy-demanding process that is typically accompanied by an increase in glycolytic flux. Glycolysis starts with glucose transporters' uptake (e.g., GLUT1, which is predominant in immune cells) and subsequent processing in the cytosol to yield pyruvate, generating TP and reducing NAD^+ to NADPH during the process. Immune effector functions are directly controlled by the glycolytic enzymes hexokinase 1 and 2 (HK1 and HK2), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), enolase, and the pyruvate kinase isoenzyme M2 (PKM2) (Chang et al., 2013, Galván-Peña et al., 2019, Moon et al., 2015, Palsson-McDermott et al., 2015). In addition, to maintain glycolytic flux in the absence of mitochondrial oxidative phosphorylation (OXPHOS), cells can reduce pyruvate to lactate via lactate

dehydrogenase (LDH) to recycle NADH into NAD⁺. This fermentation reaction can be measured as ECAR in Seahorse analysis when protons, along with glycolysis-derived lactate, are exported from the cell by monocarboxylate transporter MCT1. While glycolysis is most prominent in inflammatory immune effector cells (Krawczyk et al., 2010, Michalek et al., 2011), it is critical for diverse modes of immune activation, including alternative macrophage activation and the induction of T regulatory (Treg) cells. In addition, glycolytic rates differ between B and T lymphocytes (Caro-Maldonado et al., 2014, De Rosa et al., 2015, Huang et al., 2016, Jan et al., 2016).

The pentose phosphate pathway in inflammatory myeloid cells

The pentose phosphate pathway (PPP) branches off from the glycolytic pathway when hexokinase-derived glucose-6-phosphate is oxidized by G6P dehydrogenase (G6PD), the rate-limiting step of the PPP. This oxidative branch generates reducing equivalents of NADPH that have multiple roles in immune cells. Neutrophils and inflammatory macrophages use the PPP and especially NADPH oxidase to generate reactive oxygen species (ROS) to combat infectious agents. Moreover, NADPH is used to generate fatty acids (through fatty acid synthesis, FAS) and anti-oxidants, such as glutathione, that prevent excessive tissue damage. Nucleotide and amino acid precursors needed to support cell proliferation are produced in the non-oxidative branch of the PPP. Together, both branches of the PPP support anabolic programs and immune effector functions (Baardman et al., 2018, Haschemi et al., 2012).

The TCA cycle, ETC as a central hub

The TCA cycle and mitochondrial electron transport chain (ETC)-based energy production are essential elements of life as most cells rely on energy production via OXPHOS, which is typically evaluated measuring OCR. Compared to glycolysis, OXPHOS is much more efficient in generating energy, and thus this pathway is typically associated with the longevity of homeostatic immune cells such as alternatively activated macrophages or B and T memory cells (Lam et al., 2016, van der Windt et al., 2013). However, during the pro-inflammatory activation of macrophages, the TCA cycle is significantly remodeled via downregulation of IDH and SDH activity (Jha et al., 2015). The latter is achieved through direct inhibition by the immunoregulatory metabolite itaconate that ACOD1/IRG1 specifically produces inactivated myeloid cells (Lampropoulou et al., 2016, Swain et al., 2020).

Fatty acid metabolism supports immune cell phenotypes and functions

Fatty acid oxidation (FAO) is a catabolic pathway converting fatty acids into products like acetyl-CoA, NADH, and FADH₂, which are used in the mitochondria to generate energy. Carnitine palmitoyltransferase 1 and 2 (CPT1 and CPT2) shuttle long-chain fatty acids into the mitochondrial matrix for subsequent oxidation to acetyl-CoA by hydroxyacyl-CoA dehydrogenase (HADHA). Experiments often using too high concentrations of the CPT1-inhibitor etomoxir suggested that FAO is essential in IL-4-induced macrophages, memory T cells, and Tregs (Michalek et al., 2011, van der Windt et al., 2012, Vats et al., 2006). However, studies that employed cell-specific genetic knockouts for CPT1 and CPT2 reports that FAO is primarily dispensable for these processes (Divakaruni et al., 2018, Nomura et al., 2016, Raud et al., 2018, Van den Bossche and van der Windt, 2018).

As opposed to FAO, FAS is an anabolic pathway converting cytosolic acetyl-CoA into lipids. Acetyl-CoA carboxylase (ACC) first carboxylates acetyl-CoA to malonyl-CoA, which is subsequently elongated by fatty acid synthase (FASN). FAS supports cellular proliferation of effector T cells and is key to configuring the plasma membrane for inflammatory signaling in macrophages and for endoplasmic reticulum synthesis to allow cytokine secretion by activated DCs (Everts et al., 2014, Wang et al., 2011).

Amino acid metabolism is differentially regulated in distinct immune cells

Immune activation of T cells is associated with an increased demand for amino acid metabolism, as exemplified by the importance of L-type amino acid transporter 1 (LAT1) and the serine pathway for successful signaling upon TCR engagement (Hayashi et al., 2013, Ma et al., 2017, Ma et al., 2019). Furthermore, amino acids play a role in fine-tuning the specific direction of immune responses. For instance, both Th1 and Th17 cells increase glutamine usage via the transporter protein ASCT2 in response to antigen receptor stimulation, whereas anti-inflammatory Tregs are not affected by altered glutamine supply (Nakaya et al., 2014). In addition, Glutaminase (GLS) converts glutamine into glutamate to fuel the TCA cycle, and this pathway promotes Th17 differentiation while diminishing Th1 and cytotoxic T lymphocyte differentiation (Johnson et al., 2018).

In macrophages, amino acids also play important functional and regulatory roles. For example, glutamine metabolism has been associated with anti-inflammatory polarization programs, and serine metabolism was implicated in the regulation of IL-1 β production (Jha et al., 2015, Rodriguez et al., 2019). Furthermore, amino acids can also be used for the production of effector molecules in macrophages. For example, LPS(+IFN γ)-induced macrophages mainly convert arginine into nitric oxide (NO) via inducible NO synthase (iNOS), while IL-4-stimulated macrophages predominantly metabolize arginine into ornithine and polyamines (Van den Bossche et al., 2012).

Limitations of current wet-bench approaches

The need for *in vitro* to *in vivo* validation

Given the importance of environmental factors, immune cell metabolism in laboratory cell culture differs from *in vivo*. Ma et al. recently used a ¹³C-glucose infusion method in mice to investigate the metabolism of CD8⁺ T cells in response to *Listeria* infection and observed fundamentally different metabolic profiles for activated T cells *in vivo* (Ma et al., 2019). While *in vitro*-activated T cells show a deviation from OXPHOS toward glycolysis, CD8⁺ T cells activated *in vivo* display higher rates of oxidative metabolism and reliance on glucose-dependent biosynthesis of the amino acid serine. The approach developed in this work requires delicate experimental implementation and is not readily scalable or applicable to human samples. This translation to the human setting is vital since animal models have their limitations and do not always replicate human immunobiology. Generally, the further development of *in vivo* or *ex vivo* assessments of cellular metabolism with these and other techniques has a high priority in the field. Meanwhile, the validation of critical concepts in immunometabolism research *in vivo*, especially in humans, requires alternate options for metabolic profiling.

The need to understand the spatiotemporal aspects of immunometabolism

Immune cells heavily rely on a timely supply of nutrients, which can differ depending on the location of the cell within the tumor microenvironment or inflammation site (Van den Bossche and Saraber, 2018). Sorting and bulk measurements inevitably average profiles over cell populations, obscuring the effects of timing and environmental variables. As such, the cell's metabolic state at single-cell resolution can explain the molecular mechanisms underlying their phenotype heterogeneity within complex tissue microenvironments *in vivo*. This could help to elucidate the reasons behind their functioning or their inability to function appropriately in pathological settings. Most recent transcriptional and imaging technologies can resolve the spatial arrangement of individual cells or even metabolites. Their further development will be critical to establishing the next level of understanding of immunometabolic remodeling in human samples and *in vivo* settings (Mazumdar et al., 2020). Timing is another aspect that is often overlooked in the immunometabolism field. Immune cells ramp up glycolysis upon activation, which ultimately leads to the increased metabolic fluxes and abundance of specific metabolites in their effector state (typically assessed after 24 h). However, how this happens over time and how it supports the acquisition of the effector phenotype and disease progression is not well understood. Resolving the kinetics by which immune cells metabolically transit into their effector phenotype is essential to define which events cause and subsequent consequences.

Cell numbers and sorting time limitation

The lack of *in vitro* to *in vivo* transitions and mouse-to-human validations can be at least partially explained by the technological limitations of the commonly used bulk analysis techniques in immunometabolism research. A major caveat of both metabolomics and extracellular flux analysis is the need for relatively high cell numbers often not present in human clinical samples. For example, the Seahorse XF analyzer measures extracellular fluxes in 10k-100k cells per well and typically requires 4-5 replicate wells per condition. For metabolomics, researchers typically measure ~0.5 million pooled cells per replicate, and even with this high input, the metabolic coverage is limited as we typically measure hundreds of metabolites out of the estimated thousands present in the cell. Another drawback of extracellular flux analysis and metabolomics is that both assess metabolic features that are relatively unstable. The obtained results can be affected by lengthy and harsh tissue digestion and fluorescence-activated cell sorting protocols (Llufrio et al., 2018). Therefore, new metabolic profiling methods accompanied by novel, optimized isolation approaches are needed to measure stable metabolic features within small quantities of immune cells *ex vivo*.

Leveraging different 'omics approaches for indirect metabolic assessment

The instability of metabolite levels can be partially overcome by integrating metabolomics with other 'omics data that are more stable during processing, such as transcriptomics and proteomics. Metabolomic fluxes can be regarded as the net result of metabolite levels and the presence and activity of the enzymes that metabolize them. Gene expression analysis and proteomics can already give some insight into the regulation of metabolic pathways but become particularly powerful when integrated with metabolomics data. For example, Jha et al. identified a necessary metabolic break in the TCA cycle that supports inflammatory macrophage responses using a network integration approach (Jha et al., 2015).

Baardman et al. also combined metabolomics and transcriptomics data to demonstrate that a defective PPP is associated with decreased inflammatory responses in lipid-laden macrophages (Baardman et al., 2018). Similarly, integration of bulk proteomics and metabolomics revealed the metabolic requirements for T cell activation (Geiger et al., 2016). Thus, these distinct 'omics approaches, especially when integrated, can reveal (de)regulated metabolic hubs that can be functionally tested using genetic tools or pharmacological compounds.

Single-cell metabolomics is currently too premature for large-scale applications, and thus alternate approaches are needed for immunometabolism research to enter the single-cell era. Previous successes using bulk transcriptomics and proteomics to delineate cellular metabolism show that using the related single-cell 'omics approaches is the way to go at the moment.

Metabolic analysis based on single-cell transcriptomics

With the advent of scRNA-seq, one can resolve transcriptional profiles of individual cells within complex multicellular ecosystems, which is critically important in an immunological context. Using the standard 10x Genomics pipeline, scRNA-seq experiments nowadays profile up to ten thousand cells per sample, routinely detecting two to four thousand genes per cell. Even direct examination of scRNA-seq data in an immunological context can illuminate the vital metabolic processes *in vivo*. For example, we can use data on the immune cells within murine sarcomas treated either with immune checkpoint blockade (ICB) anti-PD-1/anti-CTLA-4 antibodies or with control antibodies (Gubin et al., 2018). In this system, ICB treatment leads to tumor rejection, which is paralleled by the overall activation of the immune compartment and the appearance of activated cells within the myeloid compartment (Gubin et al., 2018).

A similar approach to directly examining specific essential metabolic genes in scRNA-seq data proved robust in describing apoptosis's metabolic aspects during *Drosophila* eye development (Ariss et al., 2018). Together, these examples illustrate the potential utility of scRNA-seq data for studying metabolism. However, methods to comprehensively map the metabolomic landscape at single-cell resolution remain relatively scarce and are in the early stages of development. Most approaches for a metabolic analysis of scRNA-seq data arise from bulk RNA-seq analysis techniques, and considerable challenges have to be overcome when making this transition. Indeed, bulk transcriptional profiling provides an appealing source of data as it yields exhaustive information about transcript levels within the sequenced cells. As such, the absence of a signal typically implies an absence of the corresponding transcript. This contrasts with metabolomics profiling, where coverage is limited, and the absence of a signal often arises due to technological limitations rather than due to the absence of the metabolite itself. However, this advantage of transcriptional profiling is not as pronounced in scRNA-seq data. Due to generally lower depth of sequencing and drop-out events, expression levels of individual genes can appear zero in some cells even though these genes are expressed. The typical depth of sequencing is in 25k-50k reads per cell, which yields confident detection of 3k-4k genes per cell. Shallow sequencing (i.e., on the order of 1k genes per cell) might result in the detection of most common and housekeeping genes and only very few subpopulation-specific genes, which would lead to significant difficulties when comparing the data with other published signatures. Still, even the detection of 3k-4k genes per cell is significantly lower than the estimated 12k-15k genes expressed in purified cell populations based on the bulk RNA-seq

data. In these situations, data imputations have to be used to post-process the data and obtain smooth transcriptional profiles. Within the same scRNA-seq dataset, different cell clusters can have very different RNA content per cell, which results in uneven coverage between cell types. For example, a smaller number of total transcripts is typically detected in neutrophils than T cells or macrophages. Thus, care is required when comparing pathway enrichments between different cells since lower coverage of some clusters can lead to false-negative results.

Most of the approaches for a metabolic analysis of scRNA-seq data can be broadly classified into two main categories that are briefly described below: (1) pathway-based analysis and (2) flux balance analysis (FBA)-based methods.

Pathway-based approaches

Genes comprising distinct metabolic pathways are fairly well-annotated and can be directly probed as gene sets using standard bioinformatics pathway enrichment techniques. This simple yet powerful approach can reveal novel biological insights. For example, the role of serine metabolism in T cells has been investigated based on the initial observation of the transcriptional enrichment of this pathway in published bulk RNA-seq data (Ma et al., 2017). This pathway is enriched in the proliferating T cell cluster within the scRNA-seq data. Xiao et al. also demonstrated how pathway enrichment pipelines could be adjusted for scRNA-seq data analysis and used to characterize the metabolic diversity within the tumor microenvironment. Using a similar analysis strategy, Miragaia et al. highlighted the metabolic heterogeneity among Tregs across different tissues using pathway enrichments (Miragaia et al., 2019).

Flux balance analysis-based approaches

FBA is an elegant way to compute the steady-state distribution of metabolic fluxes in the system to satisfy constraints on input and output fluxes given a specific network architecture. FBA views metabolic wiring in a cell as a system of pipes with a limited number of input and output faucets and requires a steady-state equilibrium established at every node of a system. Incoming metabolic flux toward each reaction should equal metabolic flux going away from that reaction. As such, specific expression levels of individual enzymes are not directly participating in modeling, and the results are primarily dependent on network topology and requirements on input and output reactions of the model. Inputs are typically the most common substrate uptake reactions (e.g., glucose, glutamine), and outputs (usually referred to as an objective function) are associated with the specific studied cell behavior. For instance, an objective can be to maximize biomass production for growing cells or produce specific metabolites such as NO for activated macrophages (Bordbar et al., 2012).

An essential advantage of the FBA-based approaches is that they rely on network analysis that is not associated with predefined pathway knowledge and can reveal novel biological concepts. As such, it provides a powerful technology to understand changes in metabolic flux upon changes in network topologies, e.g., due to gene knockout or tissue-specific expression pattern. Conceptually, this work follows studying tissue-specific metabolism based on bulk RNA-seq data from different tissues (Bordbar et al., 2011). In addition to looking at the dramatically different network topologies, the FBA framework allows the investigation of metabolic remodeling within the same network and the change of the required cellular output induced by stimulation. This is feasible when at least some critical metabolic features are

known. For instance, using this approach, Bordbar et al. modeled activation of macrophages using an increase in NO production as a critical objective of inflammatory macrophages (Bordbar et al., 2012). Key objectives can also be defined based on the differential expression between conditions. This approach for bulk data has previously been adopted to reveal itaconate as an SDH inhibitor (Lampropoulou et al., 2016).

One significant caveat of the current approaches is that they typically focus on differences between two distinct cell populations. In these situations, a more straightforward approach would be to simply perform differential expression analysis between the two populations and perform either pathway enrichment analysis or metabolic subnetwork analysis. The next-generation metabolic scRNA-seq analysis approaches are expected to reveal principal metabolic axes of variation independent of the assignments of the cellular identities. While it is feasible to perform such clustering simply in the space of all metabolic genes or using annotated metabolic pathways as principal components, the most potent clustering approaches would incorporate network-based analysis with no *a priori* knowledge of individual pathways. This would allow for an unbiased definition of individual cells as distinct metabolic entities and will provide an unbiased view of the whole cellular ecosystem from a metabolic perspective.

Protein-based single-cell metabolic analysis

Single-cell metabolic profiling by mass cytometry

While untargeted analysis of proteins within single cells is still in its infancy and has not been applied in immunometabolism research yet, recent literature showcases the use of mass cytometry (cytometry by time of flight, CyTOF) to estimate the metabolic configurations of enzymes within single cells. This technology allows for simultaneous quantification of ~40 parameters at single-cell resolution. In the context of immunometabolism, one should consider the inclusion of antibodies targeting the main metabolic features.

To design an immunometabolic CyTOF panel, Harmann et al. first screened over 100 commercial antibodies against a broad range of metabolic factors. They came to a selection of 41 metabolic antibodies that passed all biological controls and yielded robust, reproducible results in a proof-of-concept study of human white blood cells (Hartmann et al., 2021). After identifying distinct immune cell subsets based on their lineage marker expression pattern, the metabolic states of all immune cell subsets could be estimated and compared without the need for prior sorting. The metabolic state profiles acquired by mass cytometry were in agreement with their described role in specific immune functions as exemplified by high expression levels of targets associated with glucose and fatty acid metabolism in plasmacytoid DCs. High levels of the rate-limiting PPP enzyme G6PD in neutrophils. Furthermore, Hartman et al. have demonstrated a high degree of consistency between metabolic changes in T cells measured by Seahorse and the changes observed at a single-cell resolution on the protein level. This highlights that CyTOF-based single-cell metabolic profiling is a powerful tool to gain new insights into the concepts of metabolic remodeling established by bulk measurements.

Application of metabolic CyTOF in the ex vivo clinical context

Comparing tumor and healthy adjacent tissue from colorectal carcinoma patients, along with healthy donor peripheral blood mononuclear cells and lymph nodes, Hartmann et al. (2020) revealed specific metabolic phenotypes enriched in tumor-associated CD8⁺ T cells (Hartmann et al., 2021). These cells showed increased expression of the large neutral amino acid transporter 1 (LAT1, CD98) and decreased expression of the large neutral amino acid transporter 1 (LAT1), and decreased expression of a broad range of enzymes across distinct metabolic pathways. In addition, these meta^{low} cells displayed distinct signs of T-cell exhaustion, including increased expression of PD1 and CD39, reduced mitochondrial capacity, and low levels of TCF1. Of note, a fraction of the meta^{low} subset did not express PD1 or CD39, indicating that combining metabolic and immune cell markers provides additional dimensions to phenotypically and functionally define tumor-associated immune cells about human disease.

Use of CyTOF-based metabolic profiling in animal model research

In addition to human samples, metabolic CyTOF panels can be used to dissect metabolic remodeling *in vivo*. The metabolic panel is primarily configured based on earlier bulk metabolic measurements and previous literature, and as such, the obtained mass cytometry data nicely recapitulated established concepts in the field. Indeed, effector T cells showed an expected increase in glycolytic activity, as evidenced by GLUT1 and GAPDH induction in the CyTOF analysis and validated by an increased ECAR by Seahorse analysis (Michalek et al., 2011). Conversely, memory T cells showed the induction of CPT1a, a key player in FAO previously associated with CD8 memory T cells (van der Windt et al., 2012). Furthermore, this metabolic rewiring occurred in a stepwise manner that could be followed over time at single-cell resolution. This metabolic change included a uniform glycolytic switch early after activation, followed by heterogeneous GAPDH expression after two days of activation. In addition, these two GAPDH populations showed functional differences, which would not have been revealed in bulk approaches.

Furthermore, this approach revealed a transitional pool of effector T cells that could not be detected when solely using common antibodies against lineage markers and T cell subsets. This newfound subset appeared four days after infection, was highly proliferative, showed strongly induced expression of both glycolysis and mitochondrial oxidation markers, and appeared to be fueled by both fatty acids and amino acids the expression of CPT1a and CD98, respectively. Having detected this novel subset of transitional cells by CyTOF, the authors next sorted these cells as CD62L^{lo} CD44^{hi} CD25^{hi} to confirm their increased glycolytic and oxidative metabolism through extracellular flux analysis. This is an excellent example of how single-cell technologies provide novel insights and how established bulk analysis can validate such novel discoveries.

Met-Flow: Single-cell metabolic profiling by flow cytometry

Using fluorescently labeled antibodies instead of heavy-metal-conjugated ones, conventional flow cytometry can also perform single-cell metabolic profiling. Ahl et al. configured a flow cytometry-based method called Met-Flow to interrogate the metabolic state of immune cells using ten fluorochrome-conjugated antibodies targeting rate-limiting enzymes and critical proteins in distinct metabolic pathways, combined with phenotypic markers to generate a 27-color cytometry panel (Ahl et al., 2020). To acquire this extensive panel, the authors used an X-30 FACSymphony cytometer from BD.

Using the expression profiles of 10 metabolic markers allowed Ahl et al. to cluster and separate blood leukocyte subsets into distinct metabolic islands. Akin to the CyTOF analysis by Hartmann et al., CD4+ and CD8+ T cell subsets were not separated on metabolic proteins alone, but most other immune cell subsets were quite efficiently defined solely based on their metabolic profiles. This indicates that it is probably more beneficial to extend our conventional cytometry panels with antibodies against key metabolic enzymes coupled to specific immune effector functions instead of staining additional immune cell surface markers. The actual function is often unknown. Even when using relatively small cytometry panels, including a few metabolic antibodies, it may be informative to elucidate further the phenotype, function, and metabolism of immune cells using conventional equipment.

Of note, Met-Flow with only ten metabolic proteins identified blood leukocyte subsets with a resolution comparable to that of >500 genes by scRNA-seq analysis, whereas solely analyzing the expression of the same ten genes did not resolve the immune populations. Additionally, Met-Flow demonstrated the well-documented upregulation of glycolysis, OXPHOS, and FAS upon T cell activation, and this metabolic rewiring was confirmed by bulk extracellular flux analysis. Finally, follow-up single-cell Met-Flow analysis indicated that most cells relied on glucose for their metabolic switch and revealed a subset of CD4 central memory cells with an alternative metabolic fuel. As such, single-cell, but not bulk, metabolic profiling can identify specific metabolic characteristics and requirements of individual T cell subsets.

Pros and cons of distinct single-cell metabolic profiling techniques and developments

Single-cell transcriptomics versus cytometry

Single-cell transcriptomics and multicolor cytometry are complementary approaches in terms of information acquired and experimental set-up required for profiling. They can be used simultaneously for maximal resolution, but at the same time, either approach can be more accessible or informative.

scRNA-seq typically yields ~3,000-5,000 genes detected in each cell. This allows for unbiased clustering of the data and potential identification of novel subpopulations. Furthermore, it allows simultaneous detection of multiple genes across many pathways and yields a global picture of metabolic changes from the metabolic perspective. On the other hand, cytometry approaches are currently limited to detecting ~40 proteins, including cell lineage markers and selected metabolic enzymes. Therefore, the utilization of cytometry approaches requires a reasonably specific hypothesis to design the antibody panel.

Simultaneously, the specific number of cells profiled by the two approaches is very different, as scRNA-seq is currently limited to about 10,000 cells per sample, while cytometry can routinely process millions of cells and is readily scalable. This distinction becomes critical when minor cellular populations like innate lymphoid cells or DCs need to be considered. Furthermore, the cost of single-cell transcriptional profiling is nearly an order of magnitude larger than that of a cytometry run, and therefore it is not feasible to profile either large numbers of replicates or increase statistical power by profiling many cells from the same samples. Additionally, scRNA-seq is more sensitive to cellular morphology compared to cytometry

approaches. As a result, several cell types, such as myocytes or neurons, cannot be reliably profiled due to the incompatibility of cellular shapes and instrumental design. In other cases, such as neutrophils, transcriptional profiling is complicated by the fragility of the cells and overall lower RNA content, which introduces significant biases into the data.

On the other hand, one significant benefit of transcriptional profiling is rooted in the utilization of single-nucleus RNA-seq. This approach can isolate individual nuclei and profile corresponding transcripts (Mereu et al., 2020). This approach has several significant advantages, even though the number of detected transcripts is typically much smaller because most RNA is localized in the cytosol and not in the nucleus (Ding et al., 2020). First, this approach avoids morphological complications, and all cell types within the sample can be profiled on the same footing. However, most importantly, this approach allows the profiling of frozen samples, including ones that have been biobanked years prior. This opens up the ability to study cohorts of tissues from clinical samples that typically have to be collected across extended periods.

Despite these advantages, transcriptional data provide only a twice-removed proxy to the metabolic measurements and is, therefore, a less direct approach than cytometry in yielding protein-level measurements. On the other hand, antibody-based approaches directly measure levels of enzymes, including their post-translational modifications (and the modification of their key regulators). An additional benefit of using cytometry-based approaches is that it allows the inclusion of non-protein-based markers of cellular phenotypes. For instance, by including tagged IdU, BrU, and puromycin substrates, metabolic profiles and cellular phenotypes can be functionally linked to anabolic processes like DNA, RNA, and protein synthesis (Kimmey et al., 2019).

Altogether, in the setting of controlled experimental design, it would be ideal for a combinatorial approach where scRNA-seq is run on a limited scale to explore the broad landscape of the changes, followed by large-scale cytometry profiling using a hypothesis-driven antibody panel for validation at the protein level.

Mass cytometry versus flow cytometry for single-cell immunometabolism research

Flow cytometry is still the most commonly used approach for immune cell profiling, but the number of parameters that can be simultaneously tested is relatively limited due to fluorescence spillover. This occurs when the detector of another fluorochrome senses the fluorescence emission of one fluorochrome. Therefore, flow cytometry data need to be carefully compensated, and this can be particularly challenging for large panels. Nevertheless, the number of colors and antibodies used in flow cytometry is growing fast, and panels of 20 and more markers are now standard. So far, it does not reach the +40 parameters that can be assessed in CyTOF, but at this point, flow cytometry is catching up. Unlike fluorescence detectors, CyTOF uses a mass cytometer to detect each metal's TOF set by its mass. As such, both the detection overlap and background (autofluorescence in flow cytometry) are low in CyTOF analyses because heavy metals do not naturally occur in cells. Designing ideal CyTOF panels is still crucial to minimize “spillover” due to impurities of metal tags and to take into account differences in signal intensity. Also, no mass channel in CyTOF is as sensitive as bright fluorochromes such as PE in flow cytometry. Also, no mass channel in CyTOF is as sensitive as bright fluorochromes such as PE in flow

cytometry. While stained cells need to be acquired fast in flow cytometry to prevent photobleaching, metal-tagged samples can be (cryo)preserved for weeks and acquired simultaneously when all samples are collected overtime during a clinical trial. Alternatively, cells can be isolated, fixed, and stored over time and handled together once the clinical trial or experiment is finalized. Indeed, unique barcoding of up to 20 samples allows one to combine and subsequently stain, process, and acquire them as one multiplexed sample. These advantages come with a cost since CyTOF antibodies are more expensive than fluorochrome-coupled ones, and the instrument used to acquire the labeled cells is more costly and advanced in the case of mass cytometry. While multi-color flow cytometers are the workhorses of virtually every lab, mass cytometers are less common and restricted to high-end core facilities.

Nowadays, spectral analyzers are pushing the boundaries of flow cytometry. Employing five lasers to detect the entire emission spectra over 64 channels, instruments like the Cytex Aurora can now detect 30+ colors in a sensitive, fast, and accessible manner (Ferrer-Font et al., 2020). This is particularly valuable for immunometabolism research since it can combine fluorochrome-tagged antibodies against metabolic proteins and immune lineage markers with fluorescent metabolic tools to provide an extra layer of metabolic information (Scharping et al., 2016). Thus, for example, the expression of the glucose transporter GLUT1 can be directly related to the uptake of fluorescently labeled 2NBDG. Similarly, increased expression of proteins that suggest increased mitochondrial oxidation can be directly linked to mitochondrial mass and membrane potential using fluorescent dyes such as MitoTracker and TMRE, respectively. As such, a distinct advantage of fluorescence-based flow cytometry is the applicability of fluorescent metabolic dyes, whereas the most significant benefits of mass cytometry are panel size, signal stability, and barcoding possibilities that allow in-depth single-cell metabolic profiling of immune cells in patient biopsies obtained during clinical trials.

Spatially resolved immunometabolism

A significant drawback of all approaches listed so far is that they require cells in suspension and, as such, lack spatial information. The spatial aspect is essential since specific microenvironmental factors, including the complex mixture of stimuli, availability of nutrients, and interactions with neighboring cells, are critical drivers of cellular metabolic profiles in tissues and critically shape immunity and disease progression. Another drawback of conventional single-cell profiling approaches is that the specific protocols used to digest tissues into single cells induce cell death in some, but not all, cells and might favor the isolation of specific immune cell subsets with associated metabolic profiles. Together, these limitations underscore the need to validate key observations of immunometabolic features of cells within their tissue microenvironment, taking into account their spatial organization and avoiding problems associated with cell sorting. New methods that instantaneously capture spatial context and single-cell transcriptome profiles are now being applied, but their use in immunometabolism research has yet been elucidated. Nevertheless, recent publications highlight how the metabolic configuration of single cells within their microenvironment can be assessed at the protein level in a highly multiplexed manner.

MIBI-TOF as protein-based spatial single-cell immune/metabolic analysis

Hartmann et al. took an essential step toward spatial metabolic profiling by transferring their established metabolic CyTOF panel to the multiplexed ion beam imaging (MIBI-TOF) platform. This recently developed approach couples single-cell metabolic profiles, immune cell phenotypes, and functional states to cell-cell interactions and location within tissues (Keren et al., 2019). To do so, tissue sections are stained with heavy-metal-coupled antibodies (that are also used in CyTOF) and then scanned with an ion beam to release the heavy metals from antibodies recognizing specific metabolic and immune targets. For each acquired pixel, the released ions are quantified by TOF-MS as in CyTOF, with the critical advantage that the obtained high-dimensional data obtained by MIBI-TOF also contain spatial information.

Staining colorectal carcinoma and control tissue sections with lineage markers in combination with a broad range of metabolic antibodies, followed by segmentation of individual cells in the acquired images and subsequent clustering of cells into cell lineages, allowed researchers to couple immune cell phenotypes to metabolic characteristics that are driven by their specific tissue microenvironment (Hartmann et al., 2021). Metabolic profiles were spatially organized in environmental niches with similar metabolic characteristics irrespective of cell type. Cells expressing high levels of metabolic factors associated with glycolysis, mitochondrial respiration, or amino acid metabolism were often surrounded by neighboring cells expressing the same metabolic target. This microenvironment-driven metabolic polarization was especially apparent when comparing immunometabolic profiles of cells near the tumor border with those positioned farther away from the boundary. Typically, metabolically suppressed cells were farther away from the tumor border, while metabolically active cells resided at the tumor-immune edge. Since the MIBI-TOF approach can be performed on existing FFPE material obtained from biobanks, it offers the opportunity to link immunometabolic profiles and locations to clinical outcome and therapeutic success. As such, single-cell metabolic profiling is an essential next step to translate immunometabolism research toward clinical diagnostics and therapy. However, it should be mentioned that the presence of a particular enzyme does not always correlate with its activity.

Mapping metabolism within microenvironments in situ using dehydrogenase activity assays

Miller et al. previously configured an alternate method to assess the configuration of cells within the tissue microenvironment (Miller et al., 2017). Their approach relies on enzyme activity measurements at saturating substrate and co-factor availability combined with staining multiple immune cell markers on consecutive tissue sections. Dehydrogenase activities are measured for five enzymes catalyzing key steps in central metabolic pathways: G6PD in the PPP, GAPDH in glycolysis, LDH in lactate fermentation, and IDH and SDH in the TCA cycle. When these dehydrogenases are active, nitroblue tetrazolium chloride (NBT) is reduced by NAD(P)H into a bright colored state that can be quantified and related to the acquired immune cell marker expression on consecutive sections. The authors employed this technique to interrogate the metabolic signatures of macrophage and T cell subsets in human colorectal cancers. Macrophages producing IL-6 and TNF within the tumor showed increased glycolytic GAPDH activity compared to non-inflammatory macrophages within the TME, but overall tumor-associated macrophages appeared metabolically repressed in comparison to macrophages in healthy tissue. It is, therefore, tempting to hypothesize that an impaired metabolic fitness of macrophages in tumors might prevent their antitumor activities.

Moreover, Tregs showed cancer-specific metabolic features with suppressed glycolytic GAPDH activity and increased mitochondrial SDH compared to Tregs in healthy tissues. Identifying such tumor-specific metabolic properties of immune cell subsets could aid the development of new therapeutic approaches. Of note, the activities of these distinct dehydrogenases are measured at saturated substrate concentrations. Thus, they estimate the optimal *in vivo* scenario rather than a precise reflection of their activity in a given context, such as a hypoxia TME where cells compete for nutrients.

Computational tools for predicting neoantigens

In silico prediction of putative neoantigens from mutated genes consists of three main computational steps:

- (1) Identification of somatic mutations using whole-genome sequencing (WGS) or whole-exome sequencing (WES) data from paired tumor and normal tissue and reconstruction of mutated peptides
- (2) Genotyping of the patient's HLA genes from tumor RNA-seq or WES data
- (3) Prediction of peptides binding to the patient's HLA molecules

Mutated peptides arising from somatic mutations can be predicted by comparing tumor versus normal-tissue NGS data from the same patient. NGS data for neoantigen prediction is generated preferentially from WES, which provides the deepest mutation coverage by restricting the assay only to the genome's protein-coding regions. The computational analysis consists of data pre-processing and quality control, identification of somatic mutations using tools for variant detection, and prediction of the affected proteins and functional impact using public repositories of genomic, transcriptomic, and proteomic sequences.

State-of-the-art methods for HLA typing from NGS data are mature and widely used. They include OptiType (Szelek et al., 2014) and Polysolver (Shukla et al., 2015), which showed high accuracy in 4-digit HLA typing, and seq2HLA (Boegel et al., 2012), which can compute both HLA types and allele-specific expression. More recent methods include Kourami (Lee and Kingsford, 2018), HLA*LA (Dilthey et al., 2019), acrasHLA (Orenbuch et al., 2019), xHLA (Xie et al., 2017), HLA-HD (Kawaguchi et al., 2017), and HLAProfiler (Buchkovich et al., 2017), which also perform class II HLA typing. However, unbiased benchmarking of these recent tools is not available and would be extremely useful for characterizing class II typing accuracy. Minimal validation has so far been carried out.

Tools for predicting peptides binding to HLA molecules use machine-learning methods trained on large *in vitro* peptide-HLA binding data sets. NetMHC (Nielsen et al., 2003) and its pan-allele version NetMHCpan (Jurtz et al., 2017) is based on artificial neural networks and is presently the most commonly used method of high performance. Both tools predict the binding affinity as the half-maximal inhibitory concentration (IC_{50}) expressed in nanomolar units and the rank of predicted affinity compared with a set of random natural peptides to account for allele-specific bias. Strong binders are usually selected considering a binding affinity or rank lower than 500 nM or 0.5%, respectively. Recent advancements in deep learning have fostered new machine-learning methods based on deep convolutional neural networks, such as HLA-CNN (Han and Kim, 2017) and DeepSeqPan (Liu et al., 2019). In parallel, a pan-allele

method called PSSMHCpan has been developed to leverage binding motifs to predict peptide binding affinity also for currently under-represented HLA alleles (Fu et al., 2017).

Notably, only 1-5% of the class I binders predicted *in silico* using different computational tools have been experimentally validated. One potential reason for the discrepancy between predicted and experimentally validated neoantigens is the low sensitivity of mass spectrometry (MS)-based methods to identify binding peptides directly. Despite this limitation, MS measurements of eluted HLA-binding peptides can be used to directly interrogate the human immunopeptidome, namely the set of peptides presented on HLA molecules, and to enable reconstruction of antigen profiles presented *in vivo* that could not be captured from previous *in vitro* affinity studies. Furthermore, novel methods such as MHCflurry 1.2.0 (O'Donnell et al., 2018), ForestMHC (Boehm et al., 2019), MixMHCpred (Bassani-Sternberg et al., 2017, Gfeller et al., 2018), and EDGE (Bulik-Sullivan et al., 2019), as well as the latest version of NetMHC (Andreatta and Nielsen, 2016), were also trained on MS data from HLA-eluted peptides. The increasing amount of HLA-ligand MS data available in databases like IEDB (Vita et al., 2015), PRIDE (Vizcaíno et al., 2016), or SysteMHC Atlas (Shao et al., 2018) can provide rich training data sets for the next-generation predictors.

However, MS measurements have two significant limitations:

- (1) The requirement for a large amount of starting material ($\sim 1 \times 10^8$ cells)
- (2) The dependence on protein sequence databases for data analysis limits peptides' identification to the annotated human proteome.

The latter issue can be overcome with computational approaches for updating reference databases incorporating predicted non-canonical neo-antigens, such as those derived from non-exonic regions, insertions or deletions (indels), gene fusions, alternative splicing, or post-translational modifications. However, when supplementing peptide databases with non-canonical peptides, care must be taken to avoid false positives. The potential relevance of non-canonical neo-antigens was shown in a recent study on patients with head and neck tumors treated with immune checkpoint inhibitors, demonstrating that gene fusions are a source of immunogenic neoantigens mediate responses to immunotherapy in patients with low mutational load and low pretreatment immune infiltration (Yang et al., 2019).

Despite the progress in predicting class I HLA neoantigens, the current tools for predicting class II HLA neoantigens, the current tools for predicting class II HLA neoantigens recognized by CD4⁺ T cells, have limited accuracy and are advancing slowly due to a lack of proper training data.

Overall, recent developments in deep learning algorithms and MS-based immunopeptidomics have created fertile ground for developing next-generation HLA presentation predictors. However, a comprehensive and unbiased evaluation of external data is currently lacking. Unfortunately, results from two recent studies based on public peptide-HLA binding data (Zhao and Sher, 2018) or *de novo* experimentally validated human papillomavirus (HPV) peptides (Bonsack et al., 2019) provided little guidance on method selection. These studies identified either MHCflurry or NetMHC as top performers for class I binding prediction. However, they reported variable accuracy across HLA types and peptide lengths and low agreement between predicted and experimental binding affinities. This hampers a complete characterization of tools in terms of accuracy, positive predictive value, and HLA allele coverage. In this context, the generation of the optimal validation data set is of paramount importance: it should

cover a wide range of HLA alleles, effectively capture the rules of antigen presentation (which is not possible using *in vitro* assay data), and be based on unseen data for the training of the algorithms.

Moreover, even optimal models of the neo-antigen presentation cannot predict their immunogenicity. Great collaborative efforts such as the Tumor Neoantigen Selection Alliance (TESLA) are putting together large data sets to validate these *in silico* approaches. They could help to identify the best predictors of mutated peptides that bind *in vivo*.

Despite the need to further optimize tools for predicting neoantigens, they have already demonstrated preliminary clinical value. For example, in two studies on personalized vaccinations in melanoma, predicted neoantigens elicited effective immune responses (Sahin et al., 2017, Ott et al., 2017). One study used synthetic peptides to immunize six melanoma patients (Ott et al., 2017). Four out of six patients did not have recurrent disease, whereas patients with metastatic disease obtained complete tumor regression with additional anti-PD1 therapy. The other study used RNA-based vaccines based on predicted neoantigens recognized by CD4⁺ and CD88⁺ T cells in 13 patients with melanoma (Sahin et al., 2017). Neoantigen-based vaccination reduced metastatic events and caused an objective response in two of five patients with metastases and a complete response in a third patient treated with the vaccine combined with anti-PD1 immunotherapy.

Characterizing tumor-infiltrating immune cells

Bulk analysis

Bulk-tumor transcriptomics data from microarrays or RNA-seq can be used to quantify different cell types of the TME and identify gene signatures that can predict response to immunotherapy with checkpoint blockers. However, given the multifactorial and complex nature of the anti-cancer immune response and the various immune escape mechanisms, it will be challenging to validate the proposed predictive signatures prospectively (Jiang et al., 2018, Auslander et al., 2018). Quantifying different cell types of the TME can be done using either approach based on marker genes or deconvolution-based methods. Whereas tools based on marker genes such as xCell (Aran et al., 2017), TIminer (Tappeiner et al., 2017), and MCP-counter (Becht et al., 2016) assess only semi-quantitative abundance scores, deconvolution methods can quantitatively estimate individual cell types' fractions a heterocellular tissue by considering the bulk transcriptome as the “convolution” or summation of cell-specific signatures. Deconvolution methods formulate the problem as a system of equations that describe each gene's expression in a bulk sample as the weighted sum of the expression profiles of the admixed cell types. The mathematical problem (so-called inverse problem) of inferring the unknown cell-type fractions is solved using a signature matrix describing sorted cell types' expression fingerprints.

CIBERSORT is a popular deconvolution algorithm based on a signature matrix describing the expression profiles of 22 immune cell phenotypes derived from microarray data of sorted or enriched immune cell types (Newman et al., 2015); it uses support vector regression to identify the solution. More recently, CIBERSORT has been integrated into CIBERSORTx, a deconvolution tool that further enables the building of custom signature matrices from single-cell or flow-sorted bulk transcriptomic data to remove

possible batch effects (Newman et al., 2019). Another tool, TIMER, is a multistep computational approach for quantifying six immune cell types in 32 different cancer types (Li et al., 2016b). TIMER merges the input samples to be deconvoluted with immune cell reference profiles, performs normalization to remove batch effects, derives a cancer-specific signature matrix, and quantifies cell abundance scores using linear least-squares deconvolution. Despite being estimated via deconvolution, TIMER scores cannot be interpreted as cell fractions or compared across different immune cell types and datasets (Li et al., 2016b). The recently introduced method EPIC (Racle et al., 2017) was developed using bulk and scRNA-seq data from circulating and tumor-infiltrating immune cells, CAFs, and epithelial cells. EPIC provides two separate signature matrices for the analysis of blood or tumor data. Unlike previous methods, it estimates cell fractions referring to the total cells in a sample, thus enabling intra- and inter-sample comparison. Finally, quanTiseq is a recent algorithm for deconvolution of blood and tumor data based on a novel signature matrix built from a compendium of RNA-seq data sets for ten circulating immune cell types, including regulatory T cells and classically (M1) and alternatively (M2) activated macrophages (Finotello et al., 2019). quanTiseq is tailored explicitly for RNA-seq data and implements an entire analytical pipeline, from pre-processing raw RNA-seq data until deconvolution of cell fractions. quanTiseq also estimates immune cell fractions that can be compared within and between samples (Finotello et al., 2019). Analysis of samples from two cohorts of patients with melanoma treated with kinase inhibitors or immune checkpoint blockers demonstrated that deconvolution methods could be used to monitor the immunological effects of targeted agents and reveal immune cell composition in response to immune checkpoint blockers (Finotello et al., 2019).

Benchmarking cell-type quantification methods is difficult due to the differences in the cell types and estimated scores/fractions. A recent comparative benchmarking revealed high accuracy in quantifying CD8⁺ T cells across different approaches but limited performance for heterogenous cell types such as dendritic cells (Sturm et al., 2019). The **Tumor Deconvolution Challenge**, organized by the Dialogue on Reverse Engineering Assessment and Methods (DREAM) initiative, can reveal the top performers and provide guidelines for selecting the best method based on the cell type interest. Significantly, besides simple enumeration of cell types, novel methods including CIBERSORTx (Newman et al., 2019) and linseed (Zaitsev et al., 2019) can reconstruct cell- and sample-specific transcriptional profiles and, thus, have the potential to elucidate the functional state of cell subpopulations in the TME.

In summary, the method's selection depends on the questions to be addressed and the type of information expected to be gained. For example, EPIC and quanTiseq are the preferred methods to obtain cell fractions compared both within and between samples. In contrast, MCP-counter and xCell provide higher signature specificity and lower background noise, respectively (Sturm et al., 2019). Specific methods can also be selected considering the cell type of interest (for example, CIBERSORT, xCell, and quanTiseq for M1/M2 macrophages; xCell, EPIC, and TIMER for epithelial cells).

Cell phenotypes from single-cell data

Compared with bulk approaches, single-cell technologies can provide complementary insights into cancer immunity and have been used to study the TME of different cancer types (Tirosh et al., 2016, Puram et al., 2017, Li et al., 2017b, Lavin et al., 2017, Azizi et al., 2018, Lambrechts et al., 2018). Notably,

scRNA-seq techniques open new avenues to study rare or unknown immune cell types (Villani et al., 2017) and shed light on the transcriptional programs that underlie the immune cells' plasticity and functionality. For example, scRNA-seq from tumor-infiltrating CD8⁺ T cells can provide valuable information about their activation state and exhaustion level (Zheng et al., 2017). Besides the investigation of gene expression in single immune cell types, single-cell technologies can uncover the genetic and transcriptomic makeup of tumor cells, cancer stem cells, and cells committed to epithelial-to-mesenchymal transition as the detection of cell-specific genetic variants and estimation of tumor clonality and evolution. However, care must be taken when using scRNA-seq techniques for quantifying the cellular composition of tumors due to the differences in single-cell dissociation efficiency relative to immune cells, which can bias cell-type proportions (Lambrechts et al., 2018).

Analysis of scRNA-seq data shares some analytical steps with bulk RNA-seq (e.g., read mapping) and poses additional challenges due to the peculiarities of these data: high data dimensionality, higher noise, and absence of biological replicates per se, and data sparsity due to gene dropouts. Although the computational strategies for scRNA-seq data are far less mature and standardized than those for bulk RNA-seq, well-implemented and documented frameworks tackling the main analytical steps are already available. They include pipelines using R (for example, Seurat (Stuart et al., 2019), Scater (McCarthy et al., 2017), SINCERA (Guo et al., 2015), and Scraper (Lun et al., 2016)), Python (for example, Scanpy (Wolf et al., 2018)), or user-friendly graphical interfaces (for example, Granatum (Zhu et al., 2017) and ASAP (Gardeux et al., 2017)). The core pipeline usually consists of four main steps:

- (1) Quality control and removal of low-quality cell profiles (for example, stressed cells or doublets)
- (2) Selection of informative genes (for example, genes with highly variable expression among cells)
- (3) Normalization of expression profiles to allow cell comparison
- (4) Annotation of cell types based on their transcriptional profiles

Seurat is currently the best developed and documented framework and allows single-cell, multi-omics data integration, data harmonization, and cell-type identification (Stuart et al., 2019).

Annotation of different cell types is a pivotal step in scRNA-seq data analysis. However, there is currently no consensus on identifying known systematically and novel cell types (or cell states) based on their expression profiles. One common approach is to use unsupervised clustering to group cells with similar profiles and – assuming that each cluster represents one cell type or cell state – to identify the marker genes specific for each cluster. However, this approach has several limitations:

- (1) Clustering approaches may force the partitioning of the data into discrete clusters even when cells cover a continuum of states
- (2) Results strongly depend on the clustering strategy adopted (that is, computational method and parameter settings).
- (3) Standard clustering methods might not identify small clusters or rare cells, and therefore dedicated approaches such as RaceID3 (Herman et al., 2018) and GiniClust (Jiang et al., 2016) have to be used.

- (4) Once cluster-specific marker genes are identified, there is no standard strategy to assign cell identities to clusters.

Most of the scRNA-seq studies published so far involve manual cell annotation based on marker genes and prior knowledge, an approach that is labor-intensive and has low reproducibility.

One alternative approach is to project scRNA-seq data onto reference expression profiles of previously annotated cell types. For example, the tool scmap (Kiselev et al., 2018) maps single-cell profiles onto single cells or clusters of a reference data set. SingleR classifies single-cell transcriptomes by comparing them with expression profiles of sorted cell types using correlation analysis (Aran et al., 2019). SingleR embeds reference data from human and mouse cell populations, including immune cells, and accepts user-supplied references. Other tools leverage the integration and harmonization of scRNA-seq datasets across studies (Stuart et al., 2019) or map marker genes onto cell ontologies (Aevermann et al., 2018). For example, the recently developed scMatch method is based on correlation analysis but can either use reference transcriptomes or cell ontologies to annotate cells. When analyzing sizeable scRNA-seq datasets, a combination of manual and automated cell annotation should be used, whereas, for small datasets, manual annotation can be sufficient. However, be careful when prior knowledge is integrated, specifically for classifying marker-negative cells, as they might be affected by dropouts.

Independent evaluation of the computational tools for scRNA-seq data analysis has not yet been carried out. Hence, the use of consensus approaches using diverse methods to validate the results' robustness is recommended. Soon, the availability of carefully designed, gold-standard datasets such as those recently generated using both droplets- and plate-based scRNA-seq (Tian et al., 2019) will finally enable method benchmarking and definition of guidelines and best practices data analysis.

Lymphocyte receptor repertoires

Interrogation of cancer immunity also requires the search for common clonotypes involved in response to tumor antigens to identify shared BCR and TCR sequences. The specificity of B and T cell responses (that is, which antigens they recognize) depends on the repertoire of receptors they are equipped with. NGS has become a powerful tool to interrogate BCRs and TCRs, and different computational tools now provide simplified access to the analysis of BCR and TCR diversity (Heather et al., 2018). Recently, the analysis of immune repertoires from sequencing data has seen two significant advancements:

- (1) The development of dedicated computational tools for the extraction of immune repertoires from bulk-tumor RNA-seq data
- (2) The possibility to determine pairs of protein chains of individual TCRs and BCRs from single cells

Three tools have been recently developed for the analysis of TCR and BCR repertoires from bulk RNA-seq data. Initially developed for targeted sequencing of BCR and TCR repertoires (Bolotin et al., 2015), MiXCR has been recently adapted to analyze bulk-tumor RNA-seq data with high accuracy and precision (Bolotin et al., 2017). Furthermore, the TRUST algorithm initially developed for TCR analysis of bulk-tumor RNA-seq data (Li et al., 2016a) can also extract BCR repertoires (Hu et al., 2019). However, this approach

can produce incomplete CDR3 sequences mapping to different clonotypes. Therefore, data post-processing is advisable to decrease false positives. Finally, V'DJer is a tool specifically designed to extract BCR repertoires from bulk RNA-seq data (as precomputed files of mapped reads), which can then be quantified in downstream analysis (Mose et al., 2016). Although promising for its applicability to short-read data, the requirement of data pre-and post-processing might restrict the usage of V'DJer to bioinformaticians.

TCRs and BCRs consist of pairs of protein chains that, collectively, determine their antigen specificity. In bulk datasets, the two chains' pairing is lost and cannot be tracked back by computational means. Single-cell approaches retain this information and further allow the joint analysis of transcriptomes and immune repertoires to link the latter to the cell state and practical orientation. Despite their still limited standardization and the lack of unbiased benchmarking, these approaches enable analyses inaccessible to bulk approaches. Several computational tools have been developed to extract TCRs (for example, TraCeR (Stubbington et al., 2016), TRAPeS (Afik et al., 2017), and scTCRseq (Redmond et al., 2016)) and BCRs (for example, BASIC (Canzar et al., 2017), BraCeR (Lindeman et al., 2018), and BALDR (Upadhyay et al., 2018)) or both (for example, VDJpuzzle (Eltahla et al., 2016, Rizzetto et al., 2018)) from full-transcript scRNA-seq data. The relevance of this approach for Immuno-oncology is increasingly being appreciated and demonstrated in numerous studies.

Spatial cellular phenotyping

The quantification of the immune contexture requires images from tissue slides to obtain cellular phenotypes and their spatial distribution. Individual cells are first detected by thresholding and segmenting the raw images. Then their phenotypes are identified and classified by detecting signals from the specific markers in the corresponding cellular compartment (such as nucleus, cytoplasm, or cytoplasmic membrane) used in staining procedures. Besides commercial software packages such as to inform (Perkin Elmer), Halo (Indica Labs), or StrataQuest (TissueGnostics GmbH), a growing number of open-source and free software tools, including ImageJ (Schneider et al., 2012), CellProfiler (Carpenter et al., 2006), and Ilastik (Sommer et al., 2011), are available for this purpose. By combining and extending their core functionalities via plug-ins, macros or scripting, custom analysis pipelines have been created and adapted to fit the different multiplex imaging methods. As an alternative to fully developed imaging software packages, image analysis pipelines are often implemented using the image-processing routines and libraries from MATLAB (imaging toolbox) or Python (scikit-image and OpenCV) (Giesen et al., 2014, Keren et al., 2018). This is the case for novel multiplex imaging techniques including IMC, MIBI-TOF, MERFISH, CODEX, seqFISH, Spatial Transcriptomics, or Slide-seq, which require pre-processing, image restoration, and post-processing tools.

These primary analyses of raw images typically result in datasets that provide information about each individual detected cell, including spatial coordinates, expressed markers, staining intensities of the expressed markers, compartments, and meta structures (that is, tumor or stroma). Different software packages that are either commercial (for example, TIBCO Spotfire and Phenomap) or freely available (CellProfiler Analyst (Dao et al., 2016) and histo-CAT (Schapiro et al., 2017)) implement methods (for example, t-distributed stochastic neighbor embedding (t-SNE), PhenoGraph clustering or SPADE) to explore and analyze these high-dimensional data. By assessing the cellular phenotypes and their spatial

relationship. For example, the distances between the cellular phenotypes can give information about the tumor immune architecture and reveal immune cell types that form local neighborhoods or are dispersed throughout the tumor, indicating tumor-immune cell interactions. Furthermore, analyzing the spatial neighborhood can reveal enrichment or depletion in cell-cell interactions that indicate cellular organization and cell-cell communication (Schapiro et al., 2017). Clustering multiplex images based on their phenotypic signatures and specific cell-cell interactions may reveal distinct groups defining individual disease states (Giesen et al., 2014, Keren et al., 2018, Schapiro et al., 2017).

Despite the available tools' functionality, analyses of whole-slide images and information extraction using computational algorithms still represent a significant bottleneck. In addition, many pre-processing and intermediate steps depend on manual interaction and require in-depth knowledge of image processing. Notably, the correct and reliable segmentation of cells, markers, and other features often require extensive training of machine-learning models (for example, using Ilastik or DeepCell (Van Valen et al., 2016)) or empirical determination of thresholds that are specific to the individual samples.

Visualization of single-cell data

Visualization of complex single-cell data is challenging, and numerous algorithms have been developed for representation and exploration. Although many tumor-immune cell analyses are focused on tumor-infiltrating immune cells, it is essential to bear in mind that the cellular landscape of the TME is highly heterogeneous and includes various additional cell types such as CAFs and endothelial cells. Linear dimension reduction methods (e.g., principal component analysis (PCA)) are often unable to capture single-cell data's complex structure, and therefore, approaches for non-linear transformation in two dimensions such as t-SNE. Its derivatives (viSNE (Amir et al., 2013), Barnes-Hut-SNE, Fourier-interpolated t-SNE, and hierarchical SNE (van Unen et al., 2017)) have been developed. In general, these methods are used to graphically represent functionally related groups of cells as clusters with similar gene expression profiles in 2D plots. However, although these methods are widely used, they need careful interpretation as the results are dependent on parameters balancing global and local aspects. Additional issues are the clustering performance, which can be improved by kernel-based similarity learning (Wang et al., 2017), and accuracy, which can be increased by explicitly modeling the dropout characteristics (Pierson and Yau, 2015).

Non-parametric methods, including t-SNE, have limitations such as loss of large-scale information and intercluster relationships. Still, these limitations can be circumvented by interpretable dimensionality reduction (Ding et al., 2018), as recently demonstrated by clustering immune cell types in the TME using scRNA-seq data from patients with melanoma. Similar reproducibility and preservation of global distances can be achieved with uniform manifold approximation and projection (UMAP) (Becht et al., 2019). Notably, recent studies indicate that in some cases, a cell's state represents a continuum rather than being assigned to several discrete states, which ensures the plasticity of the immune system to respond to pathogens or neoantigens released by the tumor. Pseudo-temporal trajectories can represent the continuous nature of cell states. This approach assumes that cells with similar expression profiles arise from the same lineage and that cells with more similar expressions are more closely related. Once the data have been analyzed and cells have been projected into a low-dimensional space, a minimum spanning tree can be used to

build a backbone for cell state transitions, for example, from naïve to cytotoxic CD8⁺ T cells. This 1-dimensional ordering is referred to as pseudotime.

One of the top-scoring methods concerning the analysis of the complexity of the trajectories and overall performance was partition-based graph abstraction (PAGA). PAGA generates graph-like maps of cells that preserve both continuous and disconnected structures in data at multiple resolutions (Wolf et al., 2019). Furthermore, a very recent and promising addition to the visualization toolbox is based on Markov processes to characterize cell fate probabilities (Palantir) (Setty et al., 2019). Finally, an interesting approach is implemented in the tool *velocyto*, which uses exonic and intronic reads from scRNA-seq data to model the abundances of pre-mRNAs and mature mRNAs to predict gene expression changes over time (i.e., RNA velocity). This information then predicts future cell states and displays cell kinetics in a vector field overlaid onto a dimensionality-reduced representation of the cell populations (La Manno et al., 2018).

The longitudinal single-cell analysis of samples exemplifies the usefulness of such visualization tools (Gubin et al., 2018). Using t-SNE clustering, the major monocyte/macrophage subpopulations that comprise the intratumoral myeloid compartment could be identified, as well as their remodeling upon immune checkpoint blockers. However, there were limited insights into the origins of the cells that populate the individual clusters, and only computational analyses of the pseudo-time-organized sequence of differentiation/activation events with Monocle2 (Qiu et al., 2017) revealed that neither CX₃CR1⁺ macrophages nor iNOS⁺ macrophages are present in a tumor-induced early state and that there is a branching point in the fate of intratumoral myeloid cells.

In comparison with scRNA-seq data, visualization of CyTOF data is more advanced as many methods are further developments of visualization methods for conventional flow cytometry data. Numerous cytometry data visualization and clustering tools have been developed, such as viSNE (Amir et al., 2013), PhenoGraph (Levine et al., 2015), SPADE (Qiu et al., 2011), X-shift (Samusik et al., 2016), ACCENSE (Shekhar et al., 2014), FlowSOM (Van Gassen et al., 2015), and Citrus (Bruggner et al., 2014). An exciting approach for visualization of single-cell CyTOF data is to scaffold maps based on force-directed graphs and reveal immune organization in different tissues (Spitzer et al., 2015).

In general, the choice of a specific visualization tool for scRNA-seq or CyTOF data depends on the functionality, the programming preferences (e.g., R or Python), the size of the datasets to be analyzed, and the computational requirements, and should be made in the context of the problem addressed.

Pathway enrichment visualization and analysis of omics data

Comprehensive quantification of DNA, RNA, and proteins is now a mainstay. The resulting data are increasing exponentially, and their analysis helps unravel novel biological functions, genotype-phenotype relationships, and disease mechanisms. A standard approach to analyze these data is pathway enrichment analysis, which summarizes the extensive gene list into more easily interpretable pathways. Pathways are statistically tested for over-expression in the experimental gene list relative to what is expected by chance. For instance, experimental data comprising 40% cell cycle genes are surprisingly enriched, given that only 8% of human protein-coding genes are involved in the cell cycle process.

In this portfolio, pathway enrichment analysis helps identify immune-exhaustion mechanisms by the PD-1/PD-L1 as the first rational immunotherapeutic target for brain glioma. This pathway is targetable by intravenous PD-1 inhibitors (e.g., pembrolizumab, nivolumab, and Cemiplimab) and PD-L1 inhibitors (e.g., Atezolizumab, Avelumab, Durvalumab).

This protocol encompasses pathway enrichment analysis of large gene lists typically derived from genome-scale (omics) technology. Here, open-sourced human gene expression and somatic mutation data are analyzed (Reimand et al., 2019). The tools used here are g:Profiler, GSEA, Cytoscape, and EnrichmentMap.

Pathway enrichment analysis involves three major stages:

- (1) *Definition of a gene list of interest.* An omics experiment measures the activity of genes in an experimental context. The resulting raw dataset typically requires computational processing, such as normalization and scoring, to identify interest genes, considering the experimental design. For instance, a list of genes differentially expressed between two samples can be derived from RNA-seq data (Anders et al., 2013). In addition, Gene lists derived from other types of omics experiments, such as gene expression microarray (Ritchie et al., 2015), germline and somatic genome sequencing (MacArthur et al., 2014, Gonzalez-Perez et al., 2013, Yang and Wang, 2015), quantitative proteomics (Schubert et al., 2017), and global DNA methylation assays (Assenov et al., 2014) can be used in this protocol. However, each type of data may require pre-processing.
- (2) *Pathway enrichment analysis.* A statistical method is used to elucidate pathways enriched in the gene list from step 1. Then, all pathways in a given database are investigated for enrichment in the gene list. Several pathway enrichment analysis methods are available, and the choice of which to use depends on the type of gene list.
- (3) *Visualization and interpretation of pathway enrichment analysis.* Many enriched pathways can be identified in step 2. In addition, visualization can help elucidate the main biological themes and relationships for experimental evaluation.

Stage 1: Definition of a gene list using omics data

Genome-scale experiments generate raw data that needs to be processed to identify gene-level details for pathway enrichment analysis. The specific processing steps are unique to the omics experiment type and may be standard and thus typically straightforward to implement, or not, in which case advanced computational skills may be necessary for data processing. However, standard processing methods are available for omics technologies and are most conveniently performed by the core facility that generates the data.

There are two major pathways to define a gene list from omics data: list or ranked list. Specific omics data naturally produce a gene list, such as all proteins that interact with bait in a proteomics experiment or all somatically mutated genes in a tumor identified by exome sequencing. Such a list is ideal for direct input into g:Profiler pathway enrichment analysis. Other omics data, by default, produce ranked lists. For instance, a list of genes can be ranked by differential gene expression score or sensitivity in a genome-wide CRISPR screen. Some pathway enrichment analysis approaches to analyze a ranked gene

list filtered by a particular threshold. Alternative approaches, such as GSEA, are designed to analyze ranked lists of all available genes and do not require a threshold. A whole-genome ranked list is suitable for input into pathway enrichment analysis using GSEA. A partial (non-whole-genome) ranked gene list should be analyzed using g:Profiler.

As an example, this portfolio describes the analysis of raw RNA-seq data from hepatocellular carcinoma samples to define a ranked gene list. DNA sequence reads are quality-filtered (e.g., trimming to remove low-quality bases) and mapped to a genome-wide reference set of transcripts to allow counting of reads per transcript, which are aggregated at the gene level. Thus, RNA-seq data for multiple biological replicates for each of multiple experimental conditions are available. First, read counts per gene are normalized to eliminate the unwanted technical variation between samples, such as differences in sequencing lane or total read count per sequencing run (Rapaport et al., 2013, Bullard et al., 2010). Next, read counts per gene are investigated for differential expression across sample groups (e.g., treatment versus control). Software packages such as edgeR (Robinson et al., 2010), DESeq (Anders and Huber, 2010), limma/voom (Smyth, 2004, Law et al., 2014), and Cufflinks (Trapnell et al., 2010) implement RNA-seq data normalization and differential expression analysis procedures.

Differential gene expression analysis results include:

- (i) The P -value of the significance of the differential expression
- (ii) The related Q value (i.e., adjusted P -value) that has been corrected for multiple testing across all genes, for instance, by using the Benjamini-Hochberg false-discovery rate (BH-FDR) procedure
- (iii) Effect size and direction of expression change, so that upregulated genes are positive and at the top of the list and downregulated genes are negative and at the bottom of the list, often expressed as log-transformed fold-change. The gene list is then ranked by one or more of these values and studied using pathway enrichment analysis.

Stage 2A: pathway enrichment analysis of a gene list using g:Profiler (Step 6A)

The default analysis implemented in g:Profiler and similar web-based tools (Chen et al., 2009, Huang et al., 2009, Kuleshov et al., 2016, Mi et al., 2013) search for pathways whose genes are significantly enriched in the fixed list of genes of interest compared to all genes in the genome. The P -value of the enrichment of a pathway is computed using Fisher's exact test, and multiple-test correction is applied.

g:Profiler also contains an ordered enrichment test, which is ideal for lists of up to several thousand genes ordered by a score, whereas the rest of the genes in the genome lack meaningful signals for ranking. For instance, significantly mutated genes may be ranked by a cancer driver prediction method (Kandoth et al., 2013). This analysis replicates a modified Fisher's exact test on incrementally larger subsets of the input genes and reports the subset with the strongest enrichment P -value for each pathway (Reimand et al., 2007). g:Profiler searches a gene set collection representing disease phenotypes, Gene Ontology (GO) terms, networks, pathways, and regulatory motifs.

Pathway enrichment methods that use Fisher's exact test or related tests require the definition of background genes as control. All annotated protein-coding genes are used as default. This leads to P values' improper inflation and false-positive results if the experiment can directly measure only a subset of all genes. For example, setting a custom background is essential in analyzing data from targeted sequencing or phosphoproteomics experiments. The appropriate custom background would include all genes in the sequencing panel or all known phosphoproteins, respectively.

Stage 2B: pathway enrichment analysis of a ranked gene list using GSEA

Pathway enrichment analysis of a ranked gene list is incorporated in the GSEA software (Subramanian et al., 2005). GSEA is a threshold-free method that analyses all genes based on their differential expression rank or other score without prior gene filtering. GSEA is especially suitable and recommended when ranks are available for all or most of the genome's genes (e.g., for RNA-seq data). However, it is unsuitable when only a small subset of genes have ranks available, for instance, in an experiment that identifies mutated cancer genes.

GSEA searches for gene pathways enriched at the top or bottom of the ranked gene list, more so than expected by chance alone. For example, if the topmost differentially expressed genes are implicated in the cell cycle, suggesting that the cell cycle is regulated in the dataset. By contrast, the cell cycle pathway may not be significantly regulated if the cell cycle genes are randomly scattered through the whole ranked list. To calculate a pathway enrichment score (ES), GSEA investigates genes from the top to the bottom of the ranked gene list, increasing the enrichment score if a gene is implicated in the pathway and decreasing the score otherwise. These sum values are weighted, so enrichment in the top- (and bottom-) ranking genes are amplified, whereas gene enrichment with more moderate ranks is not amplified. The ES score is normalized relative to pathway size after being calculated as the maximum value of the running sum, culminating in a normalized enrichment score (NES) that reflects the enrichment of the top and bottom of the list. Finding a permutation-based *P*-value is computed for multiple testing to produce a false-discovery rate (FDR) *Q*-value that ranges between 0 (highly significant) and 1 (insignificant). The same analysis is performed beginning from the bottom of the ranked gene list to identify pathways enriched. The resulting pathways are selected using the FDR *Q*-value threshold and ranked using NES. The 'leading edge' of the GSEA analysis identifies genes that most significantly contribute to the enrichment signal.

GSEA has two methods for identifying the statistical significance (*P*-value) of the ES: phenotype permutation and gene set permutation. The gene set permutation test necessitates a ranked list, and GSEA compares the observed pathway ES to a score distribution obtained by repeating the analysis with randomly sampled gene sets of matching sizes. The phenotype permutation test necessitates expression data for all samples and a definition of sample groups called 'phenotypes' compared with each other. The observed pathway ES is contrasted to a score distribution obtained by randomly shuffling the samples among phenotype categories and repeating the analysis, including computation of the ranked gene list and resulting ES. Gene set permutation is ideal for studies with limited variability and biological replicates. In this scenario, differential gene expression values should be computed independent of GSEA, using variance stabilization methods, and imported into the GSEA software before pathway analysis. Phenotype

permutation should be used with a large number of replicates. The phenotype permutation approach's main advantage is that it maintains the structure of gene sets with biologically significant gene correlations during permutation compared to the gene set permutation approach. This protocol encompasses only gene set permutation because it is appropriate for the most common use case of only gene set permutation. After all, it is appropriate for the most common use case of pathway enrichment analysis. Phenotype permutation is computationally costly and, for the current version of GSEA, requires a custom algorithm to compute differential expression statistics and ESs separately for the phenotype randomizations.

Stage 3: interpretation and visualization of pathway enrichment analysis results

Pathway information is redundant, as genes often participate in several pathways, and databases may organize pathways hierarchically by incorporating general and specific pathways with many overlapping genes (e.g., 'cell cycle' and 'M-phase of cell cycle'). Consequently, pathway enrichment analysis frequently highlights several versions of the same pathway. Collapsing redundant pathways into one biological theme simplifies interpretation. Such redundancy can be addressed with visualization methods such as Enrichment Map and ClueGO. An 'enrichment map' is a network visualization representing overlaps among enriched pathways, whereas 'EnrichmentMap' refers to the Cytoscape application that creates the visualization. Pathways are shown as circles (nodes) connected with lines (edges) if the pathways share many genes. ES colors nodes and edges are sized based on the number of genes shared by the connected pathways. Clustering algorithms and network layout automatically group similar pathways into major biological themes. The EnrichmentMap software inputs a text file containing the pathway gene sets used in the original enrichment analysis and another text file containing pathway enrichment analysis results. Users can use the interactive exploration of pathway ES (filtering nodes) and connections between pathways (filtering edges). Multiple enrichment analysis can be simultaneously visualized in a single enrichment map, in which case various colors are used to differentiate the nodes for each enrichment. If the gene expression data are loaded, clicking on a pathway node will display a gene expression heat map of all pathway genes.

An enrichment map helps identify exciting pathways. First, expected themes can help elucidate the pathway enrichment analysis results (positive controls). For instance, growth-related pathways and other cancer hallmarks are expected to be identified in cancer genomics datasets. Second, pathways not previously correlated with the experimental context are investigated more carefully as potential discoveries. For example, themes and pathways with the strongest ESs should be studied first, followed by progressively weaker signals. Third, interesting pathways are examined in greater detail, elucidating genes within the pathways (e.g., GSEA leading edge genes and expression heat maps). Also, gene expression values can be overlaid on a pathway diagram from databases such as Pathway Commons (Cerami et al., 2012), Reactome (Fabregat et al., 2018), KEGG (Kanehisa et al., 2012), or WikiPathways (Kelder et al., 2012), using tools such as PathVisio (Kutmon et al., 2015). If a diagram is unavailable, tools such as STRING (Szklarczyk et al., 2015) or Gene-MANIA (Warde-Farley et al., 2010) can be used with Cytoscape (Shannon et al., 2003) to elucidate an interaction network among pathway genes for expression overlay to help in the visual identification of the pathway components that are the most altered in the

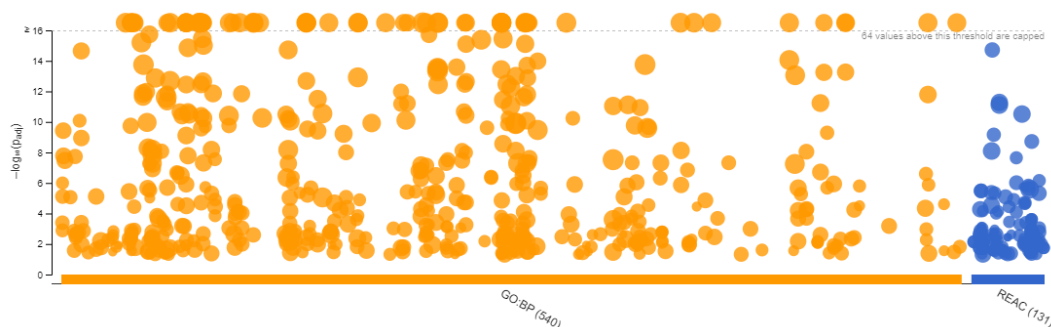
experiment. Also, master regulators for enriched pathways can be searched for by integrating gene sets of miRNA or transcription factor targets using the EnrichmentMap post-analysis tool. Finally, pathway enrichment analysis results can support a scientific conclusion or planning of experiments to support novel pathways' identification or be used for hypothesis generation.

METHODS: Pathway enrichment analysis

Two significant types of gene lists are used in pathway enrichment analysis of omics data. Flat (unranked) gene lists of dozens to thousands of genes can be analyzed using g:Profiler (option A). A statistical threshold is required to compile a gene list from omics data. By contrast, ranked whole-genome gene lists are suitable for pathway enrichment analysis using GSEA (option B). gene lists analyzed with GSEA do not require prior filtering using statistical thresholds. Partial, filtered, ranked gene lists can also be analyzed with g:Profiler.

(A) Use of g:Profiler to conduct pathway enrichment analysis of a gene list

- (i) Open the g:Profiler website at <https://biit.cs.ut.ee/gprofiler/gost>.
- (ii) Paste the gene list into the *Query* field in the screen's top-left corner. The *Homo sapiens* is selected by default as the organism for the analysis.
- (iii) Check the box next to the *Ordered query* to prioritize genes with higher mutation ESs and treat the input as an ordered gene list.
- (iv) Check the box beside *No electronic GO annotations* to discard less reliable GO annotations (IEAs) that are not reviewed manually.
- (v) Set filters on gene annotation data. Here, the initial pathway enrichment analyses include only the GO biological process and biological pathways of Reactome.
- (vi) Click *Run query* to run the analysis. A graphical heat map image will appear, with detected pathways shown along the y-axis (left) and associated genes of the input list displayed along the x-axis (bottom). The resulting pathways are organized into related groups. g:Profiler uses graphical output and switches to textual output when many pathways are found. g:Profiler returns only significant pathways with *P*-values adjusted or multiple testing correction (*Q*-values). By default, results with *Q*-values <0.05 are reported. g:Profiler reports ambiguous and unrecognized gene IDs that can be resolved manually.



- (vii) In the *Detailed output* tab, click the *GEM* button to save the output in Generic Enrichment Map format to be used in Cytoscape for further analysis.
 - (viii) Under the *Data sources* tab, download all the name.gmt.zip files. The gene set file is divided by data source. Next, download and uncompress the ZIP archive to the project folder.
- (B) GSEA Pathway enrichment analysis of a ranked gene list.
- (i) Launch GSEA by selecting the GSEA file.
 - (ii) Under *Method 3: drag and drop files here*, find the *project data folder* and select *Supplementary_Table2_MesenvsImmuno_RNASeq_ranks.rnk* file. Also, select the *Supplementary Table 3 pathway gene set definition (GMT)* file. Click *Load these files!*
 - (iii) Click on *Run GSEAPreranked* in the sidebar under *Tools*. Next, select the *Run GSEA on a Pre-Ranked gene list* tab.

Specification of the following parameters:

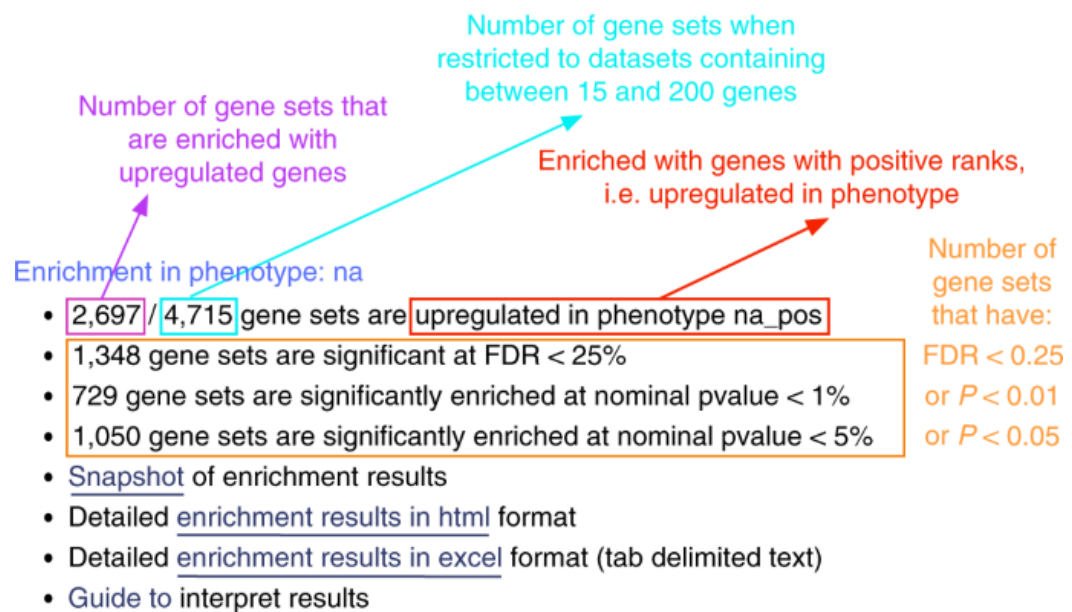
- (iv) *Gene sets database*. Click on the button ‘(...)’ on the right and wait for a few seconds to appear for the gene set selection window. Next, navigate the Gene matrix (local gmx/GMT) tab using the top right arrow. Click on the downloaded local GMT file, and click on *OK* at the bottom of the window.
- (v) *Permutations*. This specifies the number of times the gene sets will be randomized to create the null distribution to calculate the FDR *Q*-value and *P*-value. Use the default value.
- (vi) *Ranked list*. Select the ranked gene list by clicking the right arrow and selecting the rank file.
- (vii) *Collapse/remap to gene symbols*. Set to *no_collapse*.
- (viii) Select the *Show* button next to *Basic fields* to display advanced options.
- (ix) *Analysis name*. Change the default ‘my_analysis’ to a specific name, for example, ‘Mesen_vs_Immuno.’
- (x) *Max size*. Set the maximum size to 200 to remove the larger sets from the analysis.
- (xi) *Save results* and navigate to the *result data folder*.

Running GSEA

- (xii) Run GSEA by clicking the *Run* button. The *GSEA reports* at the bottom left will show the status ‘Running.’ It will be updated to ‘Success’ upon completion. It may take several minutes.

Examination of GSEA results

- (xiii) A 'Success' notification will appear in the screen's bottom-left section when the GSEA analysis is complete. All GSEA output files will be saved and be available in the *result data folder*. Click the *Success* button to open the results in the web browser. Pathways enriched in upregulated genes are displayed in the first set ('na_pos'; 'mesenchymal' in this protocol), and pathways enriched in downregulated genes are shown in the second set ('immunoreactive'; 'na_neg').
- (xiv) In the web browser results summary, click the *Snapshot* link to overview the top 20 findings. For example, significant pathways for the 'na_pos' phenotype should display enrichment in top-ranking (i.e., upregulated) genes (left side of the plot). Conversely, the most significant pathways for the 'na_neg' phenotype should display enrichment in bottom-ranked (i.e., downregulated) genes.
- (xv) In the web browser summary, click on *Detailed enrichment results in HTML format* and use the row to check the number of pathways with FDR Q values < 0.05 to determine appropriate thresholds for EnrichmentMap if no pathways are reported at $Q < 0.05$, more lenient thresholds could be used. Robust analyses should use a cutoff of not more than $Q < 0.05$. Filtering only by uncorrected P -value is inappropriate and not recommended.



a

GSEA Report for Dataset Supplementary_Table2_MesenvsImmuno_RNASeq_ranks

Enrichment in phenotype: na

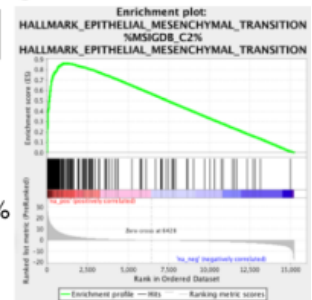
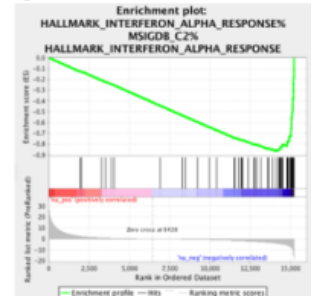
Mesenchymal

- 2697 / 4715 gene sets are upregulated in phenotype na_pos
- 1348 gene sets are significant at FDR < 25%
- 729 gene sets are significantly enriched at nominal pvalue < 1%
- 1050 gene sets are significantly enriched at nominal pvalue < 5%
- [Snapshot](#) of enrichment results
- Detailed [enrichment results in html](#) format
- Detailed [enrichment results in excel](#) format (tab delimited text)
- [Guide to](#) interpret results

Enrichment in phenotype: na

Immunoreactive

- 2018 / 4715 gene sets are upregulated in phenotype na_neg
- 1244 gene sets are significant at FDR < 25%
- 677 gene sets are significantly enriched at nominal pvalue < 1%
- 957 gene sets are significantly enriched at nominal pvalue < 5%
- [Snapshot](#) of enrichment results
- Detailed [enrichment results in html](#) format
- Detailed [enrichment results in excel](#) format (tab delimited text)
- [Guide to](#) interpret results

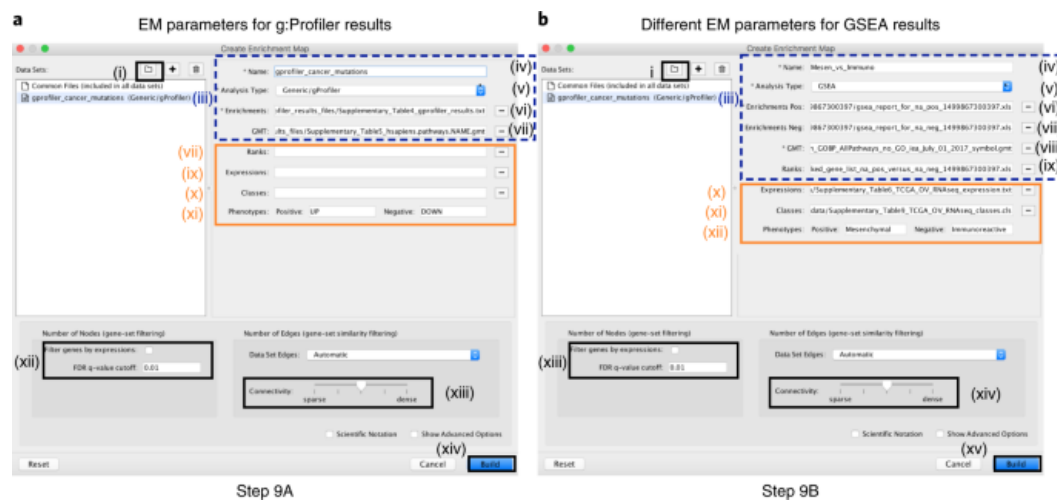
b**c**

Variation of enrichment results with EnrichmentMap

- (xvi) Launch the Cytoscape software.
- (xvii) In the menu, click *Apps*, and then *EnrichmentMap*.
- (xviii) The *Create Enrichment Map* panel will appear. Creation of enrichment maps with g:Profiler (option A) and GSEA (option B) require different input files.

A. Creation of enrichment results for g:Profiler results generated in Step 6A

- (i) In the *Create Enrichment Map* panel, select the folder icon.



Step 9A

Step 9B

- (i) Locate and select the *result data folder* containing the g:Profiler results and click *Open*.
- (ii) In the right-hand panel, g:Profiler output will be automatically populated into their specified fields.
- (iii) Verify that the *Analysis Type* is set to *Generic/gProfiler/enrichr*.

Specification of additional files:

- (xix) *Expressions*. (Optional) Upload an expression matrix for the genes investigated in g:Profiler or upload an expression dataset of all genes. If the expression dataset contains additional genes not used for the g:Profiler search, their expression values will still appear in the enrichment map's heat map.
- (xx) *Ranks*. (Optional) Specify the ranks for the gene list or the expression data.
- (xxi) *Classes*. (Optional) This is a GSEA CLS file defining the phenotype (i.e., biological conditions) of each sample in the expression file. This file is required only for phenotype randomization in GSEA; however, providing it to EnrichmentMap will label the expression file columns in the EnrichmentMap heat map viewer by phenotype.
- (xxii) *Phenotypes*. (Optional) If there are two distinct phenotypes in the expression data, update the phenotype labels so that 'positive' represents the phenotype correlated with positive values (e.g., mesenchymal) and 'negative' represents that correlated with negative values (e.g., immunoreactive).

Tuning parameters

- (xxiii) *Nodes*. By default, g:Profiler returns only statistically significant results ($Q < 0.05$), so the *FDR q-value cutoff* parameter is set to 1 in the EnrichmentMap Input unless more stringent filtering is desired. For this protocol, set the FDR Q-value to 0.01. (Optional) Select *Filter genes by expressions* to exclude genes in the GMT file that are not found in the supplied expression file.
- (xxiv) *Edges*. Keep the connectivity slider in the center. If the network is overcluttered because of too many connections (edges), move the slider to the left to make the network sparser. Contrarily, if the network is too sparse, move the slider to the right to obtain a more densely connected network.

B. Creation of enrichment maps for GSEA results generated in Step 6

- (xxv) In the *Create Enrichment Map*, click on the folder icon and locate the GSEA results folder.
- (xxvi) Click on the GSEA folder to select it. Click *Open*.

- (xxvii) In the right-hand panel, GSEA output will be auto-populated into their specified fields. Alternatively, the '+' symbol can be clicked to specify each of the required files manually.
- (xxviii) Verify that the *Analysis Type* is set to *GSEA*.
- (xxix) *Enrichment Pos.* Verify that the file name is set to '[your_path_to_gsea_dir]/Mesen_vs_Immuno.GseaPreranked.1616060783751/gsea_report_for_na_pos_1616060783751.xls', where '1616060783751' is a unique number generated by GSEA. Alternatively, navigate to the 'Supplementary_Table8_gsea_report_for_na_pos.xls' file.
- (xxx) *Enrichment Neg.* Verify that the file name is set to '[your_path_to_gsea_dir]/gsea_report_for_na_neg_1616060783751.xls'.
- (xxxi) *GMT.* Verify that the file name is set to 'Supplementary_Table3_Human_GOBP_AllPathways_no_GO_iea_July_01_2017_symbol.gmt'.
- (xxxii) *Ranks.* Verify that the file name is set to 'ranked_gene_list_na_pos_versus_na_neg_1616060783751.xls', where '1616060783751' is a unique number generated by GSEA. Alternatively, navigate to the 'Supplementary_Table2_MesenvsImmuno_RNASeq_ranks.rnk' file.

Specification of additional files

- (xxxiii) *Expressions.* (Optional) Upload an expression matrix for the genes investigated in GSEA. For an example file, see 'Supplementary_Table6_TCGA_OV_RNAseq_expression.txt'.
- (xxxiv) *Classes.* (Optional) This is a GSEA CLS file defining the phenotype (i.e., biological conditions) of each sample in the expression file. For an example, see 'Supplementary_Table7_TCGA_OV_RNAseq_classes.cls.' This file is required only for phenotype randomization in GSEA; however, providing it to EnrichmentMap will label the expression file columns in the EnrichmentMap heat map viewer by phenotype.
- (xxxv) *Phenotypes.* (Optional) Replace the 'na_pos' tag with 'Mesenchymal' and 'na_neg' with 'Immunoreactive.' 'Mesenchymal' will be labeled red because it corresponds to the positive phenotype, whereas 'Immunoreactive' phenotypes will be labeled blue.

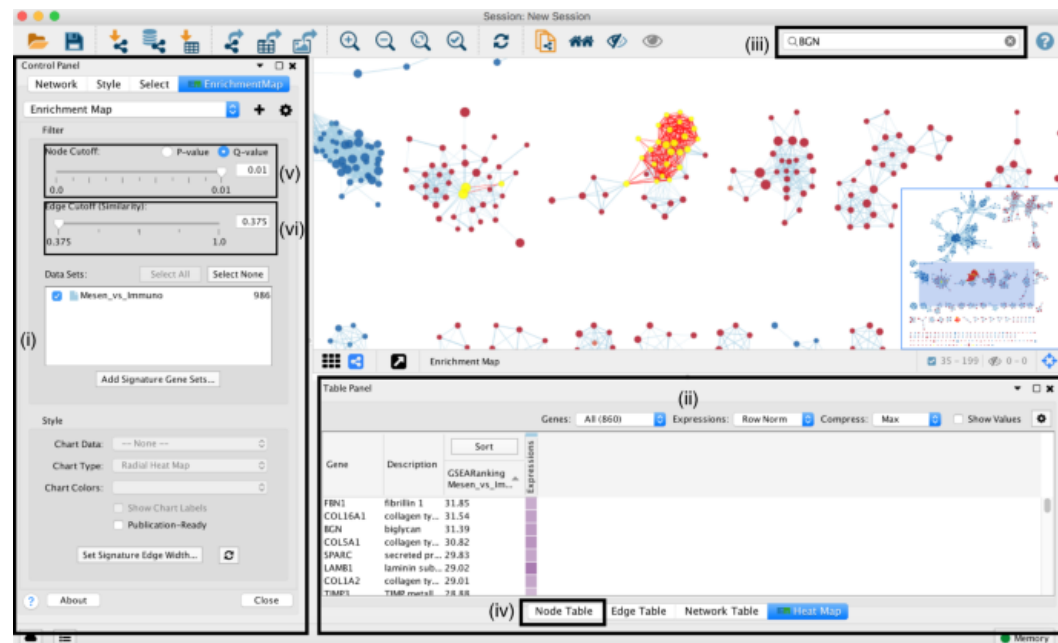
Tuning parameters

- (xxxvi) *Number of Nodes.* Set the FDR *Q*-value cutoff to 0.01. (Optional) Select *Filter genes by expressions* to exclude genes in the GMT file that are not found in the supplied expression file.
- (xxxvii) *Number of Edges.* Keep the connectivity slider in the center. To create networks with fewer edges (a sparser network), move the slider to the left. Alternatively, to create networks with more edges (a denser network), move the slider to the right.

(xxxviii) Click the *Build* button beneath the EnrichmentMap Input panel. The *Building EnrichmentMap* box that appears indicates the progress status.

Navigation and interpretation of the enrichment map

- (xxxix) Select the network in the *Control Panel* located at the Cytoscape window's left side to explore the enrichment map. The *Network Panel* can be selected on the *Control Panel*. The selected network will appear in the main window; navigate to it (zoom and pan) using Cytoscape controls, and explore the pathways by reading the gene set labels. Pathways with common genes often represent similar biological pathways and are grouped as sub-networks. Click on a node to preview the corresponding genes.
- (xl) Type a gene or pathway of interest in the search bar located in the top right corner. All pathways with that gene will be highlighted. For instance, *BGN* and *TP53* are the top genes in g:Profiler and GSEA analyses, respectively.
- (xli) To find the most enriched pathways, select the *Node Table* tab in the *Table Panel*, and select and sort the column named 'EM#_fdr_qvalue' (for g:Profiler) or 'EM#_NES' (for GSEA) by clicking on the column header. Then, highlight a subset of pathways in the network, select rows corresponding to the pathways of interest, right-click on any selected rows in the table, and select *Select nodes from selected rows*.



- (xlii) Depending on the focus of the analysis, different actions can be taken on the resulting enrichment map. Here, we define major biological themes.
 - (i) From the Cytoscape menu, launch AutoAnnotate by clicking *Apps* → *AutoAnnotate* → *New Annotation Set*.
 - (ii) In the *QuickStart* tab, click on *Create Annotations*.

- (iii) Each cluster will have a circle drawn around it and associated with a set of words that correspond to the most common node labels.
- (iv) Manually arrange clusters and move nodes to reduce node and label overlap.

Exporting figures, creating legends, and saving work

- (xlix) Save all work as a Cytoscape session. In the menu, select *File* → *Save As*. Then, navigate to the *results data folder* to save the session and specify the desired file name.

OCTAD: virtual therapeutic screening platform

One common challenge in cancer research is that the resources allocated to them are relatively limited. However, the decreasing cost of sequencing means that it is now more common to generate gene expression profiles of samples from cancer patients (e.g., RNA-Seq). Integrating these profiles with the increasing amount of other available open data (such as the effect of chemical compounds on gene expression) provides a tremendous opportunity to identify new potential therapeutic candidates computationally.

Here, OCTAD is a systems-based approach (Zeng et al., 2021) used to analyze gene expression profiles of disease samples and drug-induced gene expression profiles from cancer cell lines to predict new therapeutic candidates. The idea is to identify drugs that reverse the gene expression signature of disease by tamping down overexpressed genes and stimulating weakly expressed ones. The success of this approach is made possible by multiscale procedures, such as quality control of tumor samples, selection of appropriate reference normal tissues, evaluation of disease signatures, and integration of drug expression profiles from multiple cell lines. A minimum of ten samples of adjacent normal tissue would be preferred to account for typical tissue heterogeneity.

The deep learning (DL)-based method is used to select potential reference tissue samples for the selected case samples based on their expression profiles (Zeng et al., 2019) and implemented in Open Cancer TherApeutic Discovery (OCTAD). OCTAD is available both as a standalone software package in R for bioinformaticians and a web server resource for investigators without a coding background.

Procedure 1: Desktop version

Setup

1. Before running the pipeline, install the required packages by typing the following in the R command:

```
packages=c('dplyr', 'magrittr', 'ggplot2', 'doParallel', 'foreach', 'lme4', 'Rfast', 'httr',  
'data.table')  
  
if (length(setdiff(packages, rownames(installed.packages()))) > 0) {  
  install.packages(setdiff(packages, rownames(installed.packages())))  
}
```

2. Install the required Bioconductor packages:

```
bioconductor_packages=c('RUVSeq', 'edgeR', 'DESeq2', 'limma', 'rhdf5', 'artMS')
```

```

if (length(setdiff(bioconductor_packages, rownames(installed.packages()))) > 0) {
  if (!requireNamespace("BiocManager"))
    install.packages("BiocManager")
  BiocManager::install(setdiff(bioconductor_packages, rownames(installed.packages())))
}

```

3. Install the octad.db package via octad.db:

- Click the Packages tab in the bottom-right panel and then click install. For example, in the Install Packages dialog, write octad.db_0.99.0.tar.gz and then click install.
- Enter the following command:

```
devtools::install_github('Bin-Chen-Lab/OCTAD')
```

4. When the package is installed, load it into R:

```
library(octad.db)
```

Case and control samples

Choosing which cases (tumor samples from the phenoDF data.frame) and controls (corresponding samples such as normal tissue or tumor samples without mutation) are the two most essential factors in achieving optimal results when using this pipeline. Each group of cases needs to be evaluated individually. Visualization of cases in a *t*-SNE plot could help understand their relations with other OCTAD samples. Samples sharing similar transcriptomic profiles tend to cluster together in the *t*-SNE plot. Therefore, the cases scattering in multiple clusters is not recommended to choose as a group.

The case_id and control_id variables must be simple character vectors containing sample IDs. This is because they are most easily generated by subsetting the metadata matrix phenoDF.

Select case samples

5. Phenotype data contains tissue types, such as normal tissue, adjacent normal tissue, primary cancer, recurrent cancer, and metastatic cancer. Here, primary HCC is selected, and the phenoDF data.frame is used to list all available samples from the OCTAD database.

To select HCC samples, use the code below:

```

head(phenoDF)

HCC_primary=subset(phenoDF,cancer=='liver hepatocellular carcinoma' & sample.type
== 'primary') #select data

case_id=HCC_primary$sample.id #select cases

```

The sample IDs will be stored in the character vector case_id.

This code can be easily modified to select other cancers or a set of samples based on mutations and copy numbers (e.g., *MYC* amplification or *TP53* mutation). It is also recommended to use the R package *cgdsr* to select TCGA samples based on more molecular and clinical features.

```
#choose liver hepatocellular carcinoma samples with TP53 mutation
```

```
HCC_with_TP53_primary=subset(phenoDF,cancer=='liver hepatocellular carcinoma' &  
sample.type == 'primary' & grepl('TP53', mutation_list))
```

Compute or select control samples

6. Use the function **computerRefTissue** to compute appropriate normal tissues via comparing gene expression features between case samples and normal tissue samples. Users can select adjacent normal tissue samples by default.

```
#computing top 50 reference tissues
```

```
control_id=computeRefTissue(case_id,output=T,adjacent=T,source="octad",control_size  
=50)
```

7. The list of normal IDs is stored into the variable **control_id**. Use the following code to select adjacent control samples from *phenoDF*. This code can also be used to select any set of OCTAD samples (including cancer samples) as control.

```
HCC_adjacent=subset(phenoDF,cancer=='liver hepatocellular carcinoma' &sample.type  
== 'adjacent' &data.source == 'TCGA')
```

```
control_id=HCC_adjacent$sample.id
```

8. The relationships among case, control and other samples can be visualized by using the following code and a pre-computed *t*-SNE matrix.

```
tsne$type <- "others"
```

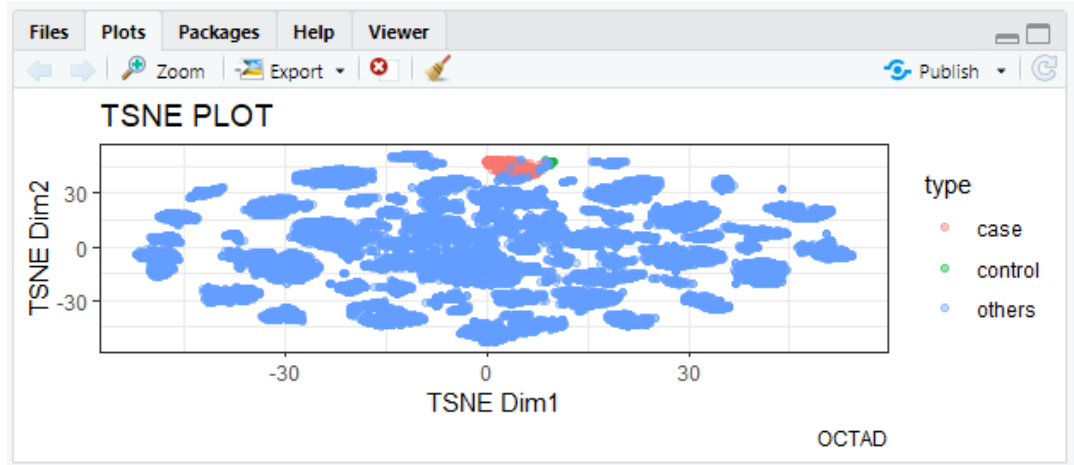
```
tsne$type[tsne$sample.id %in% case_id] <- "case"
```

```
tsne$type[tsne$sample.id %in% control_id] <- "control"
```

```
#plot
```

```
(p2 <- ggplot(tsne, aes(X, Y, color = type)) + geom_point(alpha=0.4)+labs(title = paste  
('TSNE PLOT'), x= 'TSNE Dim1', y='TSNE Dim2', caption="OCTAD") + theme_bw())
```

TSNE Plot:



Evaluation of synergy-driving gene expression

The complex genetic nature of cancer is becoming increasingly clear, with risk arising from the interplay of common and rare variants, together with environmental exposures. Although most findings from genome-wide association studies (GWAS) consider the linear effect of a single genetic variant on the trait of interest, molecular biology is known to be context-dependent. At the molecular level, the effect of a genetic variant on a molecular trait such as gene expression likely depends on the state of other variables, such as another genetic variant or a chemical stimulus. Although findings of gene-gene interactions in GWASs have lacked owing to low statistical power, recent work in identifying stimulus-dependent genetic regulatory mechanisms has demonstrated genome-wide context-dependent effects. Moreover, genes themselves rarely act in an isolated fashion; rather, complex gene-gene and gene-environment interactions determine the transcriptional landscape observed in gene expression experiments. The consequences of these interactions can prove unexpected in light of the personal effects observed, and so new studies should be undertaken to uncover the effects of gene interplay.

New approaches and technologies are urgently needed to casually link disease-associated genes to the cell types, biological pathways, and cellular functions they affect within the context of the polygenic nature of cancer and context-dependence of molecular biology. Simultaneous perturbation of gene pairs to inform on the nature of their interaction has long been standard practice in genetics research, and such approaches have recently become much more systematic through the use of high-throughput genetic screens. However, the complex genetic landscape of cancer requires these approaches on a much larger scale. Towards this, coupling the expanded toolbox of CRISPR-based tools for genetic and genomic screening with cancer cells to produce a large number of patient-specific cancer cells makes possible the multiplexed functional validation of risk variants and genes at an unprecedented scale.

In this protocol, the authors offer detailed considerations for experimental design and an analytic framework for evaluating synergistic effects driving gene expression (Schrode et al., 2021).

Gene expression studies, and even large-scale drug screens, typically compare just two conditions at a time (e.g., genetic perturbation or drug treatment versus control). This experimental framework enables the study of gene-gene interactions (i.e., epistasis), gene-environment interactions, genotype-

specific drug responses, and drug-drug interactions (Fig 1). Although developed to explore combinatorial CRISPR-based perturbations in hiPSC-derived neurons, this experimental strategy is amenable to many cell cultures and animal studies. This protocol range from combinatorial drug screening in cancer cell lines to testing various drug or stress paradigm on wild-type and knockout mice. Overall, with modest changes to experimental design, many genetic and pharmacological studies, both in vitro and in vivo, could incorporate the study of combinatorial and synergistic effects, adding value to research as diverse as addiction (gene-environment), cancer (gene-drug), and toxicology (drug-drug).

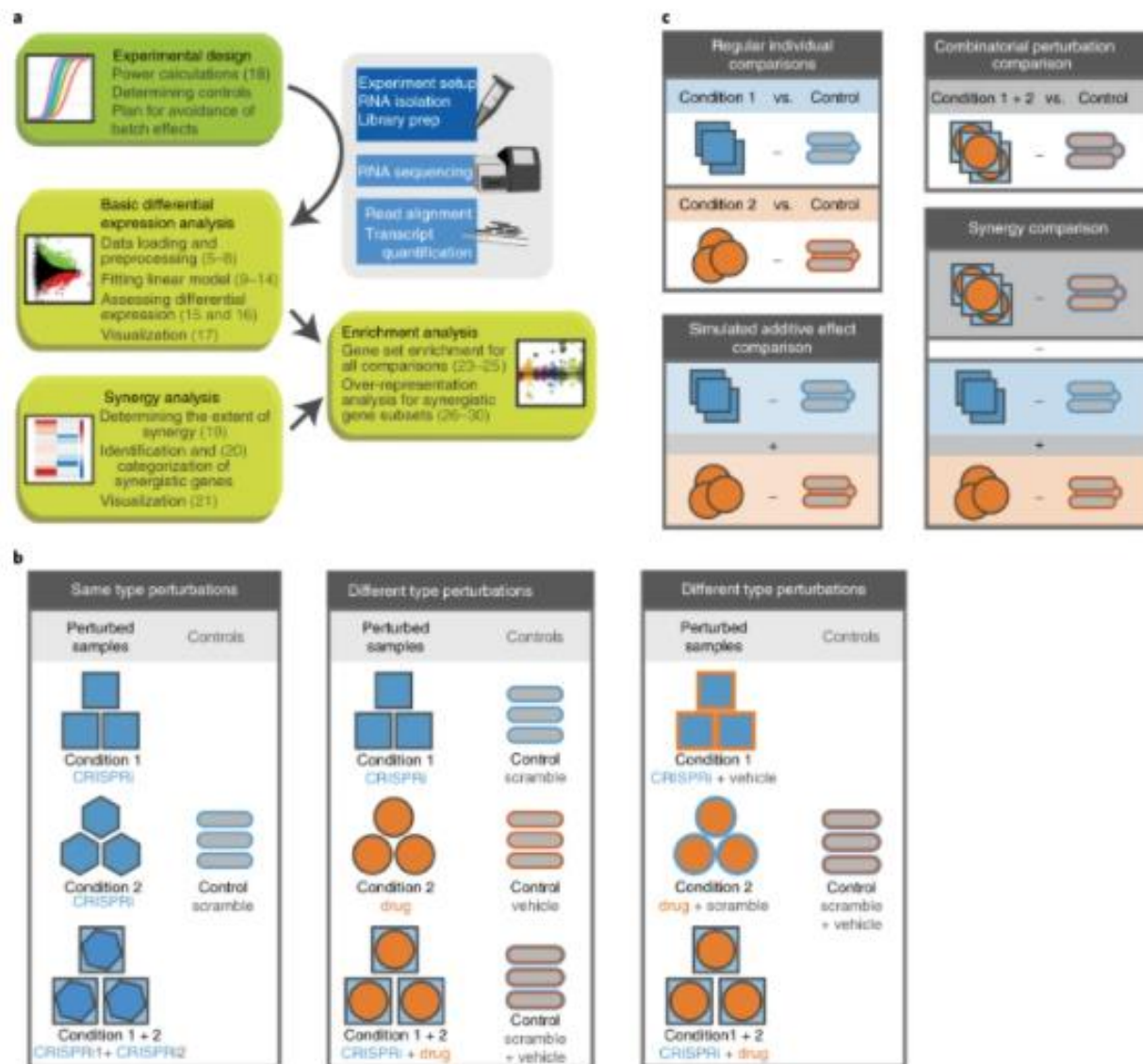


Figure 1: Analysis pipeline, experimental design, and differential expression contrast design (Schrode et al., 2021)

This analysis requires RNA sequencing data from multiple perturbed samples, both separate and in combination. The availability of biological replicates is necessary to improve the power to detect synergistic effects. To take full advantage of the provided script, the RNA sequencing input data should

ideally be provided as raw read counts instead of normalized reads per kilobase of the transcript, per million mapped reads (), as a series of custom normalization steps are built into our analysis pipeline.

The dataset used in this analysis investigated combinatorial drug treatment of BET and MEK inhibitors as therapy for MAPK and checkpoint inhibitor-resistant melanoma, owing to the availability of perturbations with single agents alone and in combination as well as biological replicates. This melanoma study reported that combining BET and MEK inhibitors synergistically inhibited the growth of NRAS-mutant melanoma and prolonged the survival of tumor-bearing mice refractory to MAPK inhibitors and immunotherapy. In addition, our analysis found significant synergistic downregulation of genes associated with various cancers, suggesting that conjunction therapy with these chemotherapeutic agents synergistically improved cancer targeting instead of the additive effects of each agent alone. This improves understanding of biological responses to treatment with multiple chemotherapeutic drugs in conjunction.

Continued progress toward precision medicine requires improvements in genotype-based diagnosis and predictions of drug treatment response. Understanding how polygenic risk adds within and across pathways will improve the calculation of polygenic risk scores (PRSs). In addition, evaluating the effect of drugs on complex genotypes will improve the ability to match patients to treatment. Moreover, understanding convergence and synergy among risk variants could identify novel therapeutic targets to prevent or treat disease. Overall, the translational effect of this protocol includes potential improvements to additive PRS scores and the prioritization of convergent and synergistic genes for mechanistic follow-up and pathways for potential therapeutic targets.

Procedure

Setup

1. Create directories:

Create a 'results' folder by using the `mkdir results` command. All tables (.csv) and plots (.pdf) will be saved in the current folder.

Create a 'genesets' folder using the `mkdir genesets` command. It contains gene set groups in the gene matrix transposed (gmt) file format). Possible gmt files for gene sets of interest can be found at <https://www.gsea-msigdb.org/gsea/msigdb.collections.jsp>. Custom gene sets can be created by assembling a spreadsheet in the gmt format and naming the file using the '.gmt' file extension.

Preparing the R environment

2. Name the experiment. This will be used to save results and generate plot titles. In this example dataset, BET and MEK inhibitors were combined to test their synergistic effect on NRAS-mutant melanoma. Therefore, the experiment is named "BM." In the R environment:

```
experiment.title="BM"
```

3. Load R packages

```
pacman::p_load(edgeR, limma, RColorBrewer, pheatmap, ggplot2,
ggpubr, qvalue, plyr, GSEABase, wesanderson, scales, grid,
webGestaltR, stringr)
```

4. Define custom functions in advance. Run each of these function definitions to define them for later use. Each function description explains what the function will do once applied to the data, starting in Step 9.

The *mds()* function is based on limma package's *plotMDS()*. When giving a DGE object and a column in the metadata table containing groups of interest, it produces an MDS plot colored by the provided groups.

```
mds <- function(normDGE, metacol, title) {
  mcol <- as.factor(metacol)
  col <- rainbow(length(levels(mcol)), 1, 0.8, alpha = 0.5) [mcol]
  plotMDS(normDGE, col = col, pch = 16, cex = 2)
  legend("center", fill = rainbow(length(levels(mcol)))), 1, 0.8, legend = levels(mcol), horiz
= F, bty = "o", box.col="grey", xpd=TRUE)
  title(main=title)
}
```

The *cameraplusplots()* function is based on *camera()* in the limma package. When given a contrast in the form of a named vector, such as a column of the contrast matrix- ('cont.matrix', which defines the comparisons being made, such as combinatorial versus additive), a list of gene set groups, a voom-transformed object (such as object 'v,' which generates a precision weight based on the mean-variance relationship of the counts), a color palette and a design matrix for the gene set groups, it produces a scatter plot of all tested gene sets and their adjusted *P* values as well as a bar graph often most significant gene sets, colored by gene sets, colored by gene set group/category.

```
Cameraplusplots <- function(contrast, genesetlist, vobject, design, catcolors, title) {
  tmp.list <- list()
  cam <- data.frame(matrix(ncol = 5, nrow = 0))
  for (i in 1:length(genesetlist)) {
    cam.s <- camera(vobject, genesetlist[[i]], design, contrast = contrast, inter.gene.cor = 0.01)
    tmp.list[[i]] <- cam.s
    names(tmp.list)[i] <- names(genesetlist)[i]
    tmp.list[[i]]$category <- names(tmp.list[i])
    colnames(cam) <- names(tmp.list[[i]])
    cam <- rbind.data.frame(cam, tmp.list[[i]])
    print(paste0("Gene set categories: ", i))
  }
  cam$neglogFDR <- -log10(cam$FDR)
  ## for plotting purposes only: cam$dirNeglogFDR <- cam$neglogFDR
```

```

cam[(cam$Direction == "Down"), "dirNeglogFDR"] <- -cam[(cam$Direction ==
                                "Down"), "neglogFDR"]
grob <- grobTree(textGrob(c("UP", "DOWN"), x = c(0.94, 0.89), y = c(0.95, 0.05), gp = gpar(fontsize
= 13), hjust = 0))
q <- ggplot(aes(x = cam$category, y = dirNeglogFDR, color = category), data = cam) +
  geom_jitter(aes(size = NGenes, alpha = neglogFDR), pch = 19, show.legend = F) +
  scale_color_manual(values = catcolors) + scale_alpha_continuous(range = c(0.4, 1)) +
  scale_size_continuous(range = c(4, 16)) + geom_hline(yintercept = c(-1.3, 1.3), color = "red",
alpha = 0.5) + labs(x = "Gene set categories", y = "-log10(FDR)", title = title) +
  geom_hline(yintercept = 0) + scale_y_continuous(limits = c(-10, 10), oob = squish, labels = abs)
theme_bw(14) + theme(axis.text.x = element_text(angle = 45, hjust = 1, vjust = 1), axis.ticks.x =
element_black(), panel.grid.minor = element_blank(), panel.grid.major = element_blank()) +
  annotation_custom(grob)
print(q)
cam$geneSet <- row.names(cam)
cam10 <- as.data.frame(cam)
cam10 <- cam10[order(cam10$FDR),]
cam10 <- cam10[1:10,]
grob <- grobTree(textGrob(c("DOWN", "UP"), x = c(0.03, 0.9), y=c(0.025), hjust = 0, col = "grey60",
gp = gpar(fontsize = 9)))
g <- ggplot(aes(x = geneSet, y = dirNeglogFDR, fill = category), data = cam10) +
  geom_col() + geom_hline(yintercept = c(-1.3, 1.3), color = "red", alpha = 0.3) +
  aes(reorder(stringr::str_wrap(geneSet, 60), -FDR), dirNeglogFDR) + xlab(NULL) +
  geom_hline(yintercept = 0) + scale_y_continuous(limits = c(-10, 10), oob = squish, labels = abs) +
  labs(y="-log10(FDR)", title=title) + scale_fill_manual(values = catcolors) +
  coord_flip() +
  theme_bw() +
  theme(panel.grid.minor = element_blank(), panel.grid.major.y = element_blank()) +
  annotation_custom(grob)
print(g)
return(cam)

```

The *oraplot()* function will take the data frame resulting from over-representation analysis using WebGestaltR (done in Step 30) and a color palette as input and returns a bar graph of the ten most significant gene sets.

```

oraplot <- function(oraes, catcolors, name) {
  oraes.n <- oraes[order(oraes$FDR), ]
  oraes.n <- oraes.n[1:10, ]
  oraes.n$neglogFDR <- -log10(oraes.n$FDR)
  oraes.n$geneSet <- gsub("_", "", title = paste0(name)) + scale_fill_manual(values = catcolors) +
  coord_flip() + theme_bw(11)
  return(g)
}

```

The *power.compare.logFC()* function will take the variances of the combinatorial perturbation, and the additive model is usually higher, in proportion to the number of individual perturbations). Furthermore,

the number of samples used to determine the variances (N), a vector of the sample (n_{tests} , usually the number of transcripts). N_{other}/N is the relative sample size, and the variance of $\log\text{FC}_1 - \log\text{FC}_2$ is $\text{sig1}^2 + \text{sig2}^2$. Because the mean standard error is inversely proportional to \sqrt{N} , multiplying the sample size by F decreases the standard error by \sqrt{F} . On the variance scale, this corresponds to dividing by n_{scale} .

```
power.compare.logFC <- function(sig1, sig2, N, N_other = c(2, 4, 6, 8, 10), alpha = 0.05,
n_tests = 20000) {
  d <- seq(0, 3, length.out = 1000)
  alpha_multiple <- alpha / n_tests
  df <- lapply(N_other/N, function(n_scale) {
    sigSq <- (sig1^2 + sig2^2) / n_scale
    cutoff <- qnorm(alpha_multiple/2, 0, sd = sqrt(sigSq), lower.tail = FALSE)
    p1 <- pnorm(-1*cutoff, d, sqrt(sigSq))
    p2 <- 1-pnorm(cutoff, d, sqrt(sigSq))
    data.frame(n_scale, power=p1+p2, d)
  })
  df <- do.call("rbind", df)
  ggplot(df, aes(d, power, color = as.factor(n_scale*N))) +
    theme_bw(14) +
    geom_line() +
    scale_color_discrete("Samples") +
    theme(aspect.ratio = 1, plot.title = element_text(hjust = 0.5)) + ylim(0, 1) +
    xlab(bquote(abs(logFC[observed] - logFC[expected]))) +
    ggtitle("Power versus difference in logFC")
}
```

The *categorize.synergy()* function is conferred the combined matrix of $\log_2\text{FC}$ values of the additive model, the synergistic effect, and the combinatorial perturbation differential expression results. It creates a new column in the resulting data frame, assigning synergy categories to each gene.

```
categorize.synergy <- function(logFCmatrix, meanSE) {
  m <- logFCmatrix
  m$magnitude.syn <- NA
  for (i in 1:length(m$Gene_name)) {
```

```

if (m$Synergistic.logFC[i] > meanSE) {
  if (m$Additive.logFC[i] < -meanSE) {
    if (m$Combinatorial.logFC[i] > meanSE) {
      m$magnitude.syn[i] = "more.up"
    } else m$magnitude.syn[i] = "less.down"
  } else m$magnitude.syn[i] = "more.up"
}
else if (m$Synergistic.logFC[i] < -meanSE) {
  if (m$Additive.logFC[i] > meanSE) {
    if (m$Combinatorial.logFC[i] < -meanSE) {
      m$magnitude.syn[i] = "more.down"
    } else m$magnitude.syn[i] = "less.up"
  } else m$magnitude.syn[i] = "more.down"
} else m$magnitude.syn[i] = "same"
}
m$magnitude.syn <- as.factor(m$magnitude.syn)
return(m)
}

```

The *stratify.by.syn.cat()* function will take a subset of interest from the table created by the *categorize.synergy()* function, above, as input. It creates a list containing vectors of genes by synergy category, which is used as input for over-representation analysis with WebGestaltR in Step 30.

```

stratify.by.syn.cat <- function(log2FC.matrix.sub) {
  synergy.cat.list <- list("less.down" =
    as.character(log2FC.matrix.sub[log2FC.matrix.sub$magnitude.syn == "less.down",
      "Gene_name"]), "less.up"
    as.character(log2FC.matrix.sub[log2FC.matrix.sub$magnitude.syn == "less.up",
      "Gene_name"]),
    "more.down" = as.character(log2FC.matrix.sub[log2FC.matrix.sub$magnitude.syn ==
      "more.down", "Gene_name"]),
    "more.up" = as.character(log2FC.matrix.sub[log2FC.matrix.sub$magnitude.syn ==
      "more.up", "Gene_name"]),

```

```

"same" = as.character(log2FC.matrix.sub[log2FC.matrix.sub$magnitude.syn == "same",
"Gene_name"])))

return(synergy.cat.list)

}

```

Loading and preprocessing data

5. Load data. Read in metadata and expression data and match the order of genes listed in the files.

```

meta <- read.csv("meta.csv", row.names = 1)

counts <- read.csv("counts.csv", row.names = 1)

meta <- meta[match(colnames(counts), row.names(meta)), ]

```

6. Filter lowly expressed genes. Plot counts over counts per million (cpm) and visually inspect the graph. Here, ~10 counts in ≥ 4 samples (number dependent on a total number of samples/replicates; Figure 2a) are kept. Adjust v (in the *abline()* function), which currently corresponds to counts > 0.25 , based on the y value of the intercept between the horizontal and the plotted counts/cpm (counts) line.

```

pdf(paste0("results/", experiment.title, "-1_cpm-counts.pdf"))

plot(cpm(counts) [, 1], ylim = c(0, 50), xlim = c(0, 3), counts[, 1])

abline(h = 10, col = "red")

abline(v = 0.25, col = "red")

dev.off()

keep <- rowSums(cpm(counts[]) > 0.25) >= 4

gExpr <- counts[keep, ]

dim(gExpr)

```

Fig. 3: Differential expression analysis output.

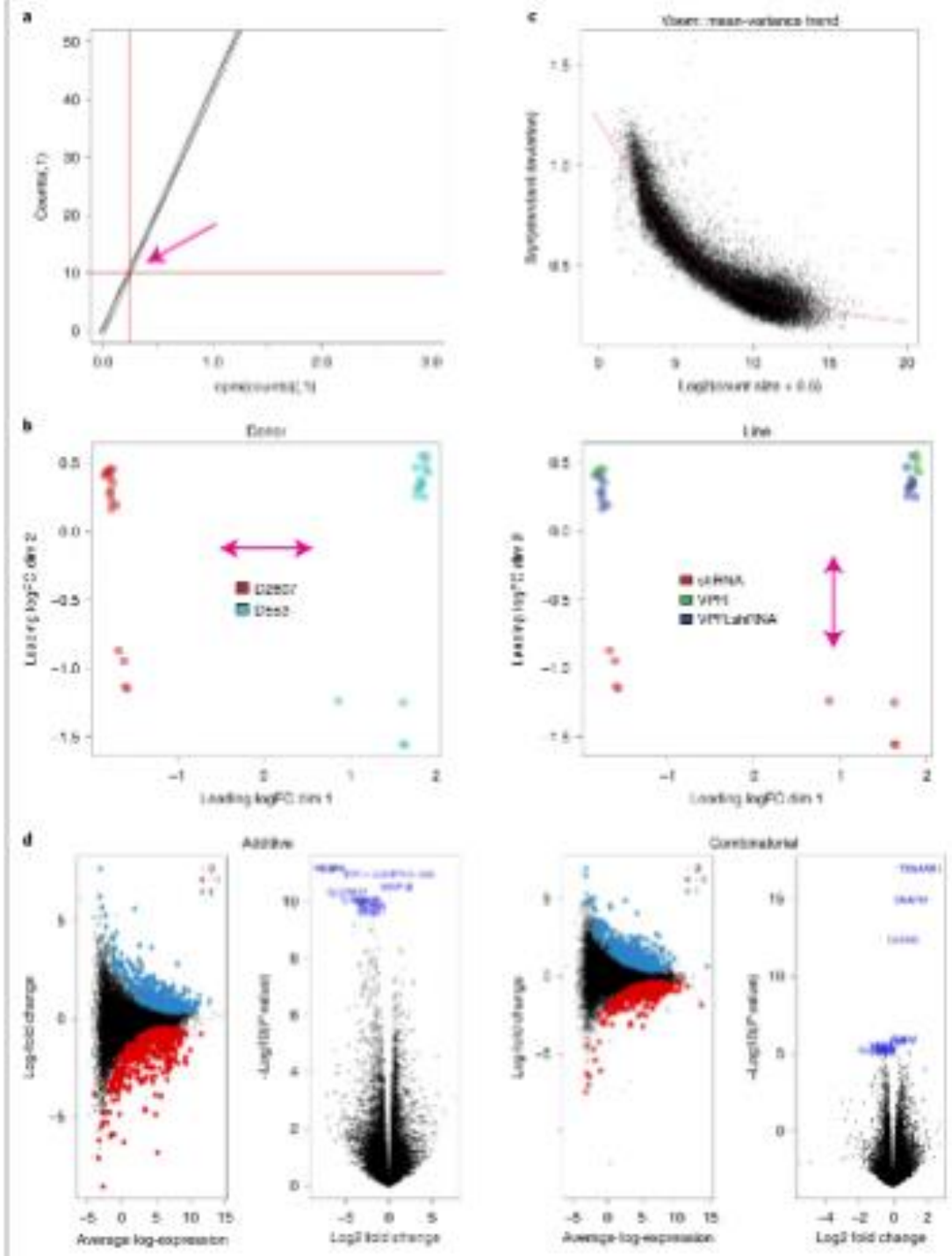


Figure 2: Differential expression analysis output (Schrode et al., 2021)

1. Create a DGEList object, a data format that holds the read counts, normalization factors, experimental group data, gene annotation, and library size information. In this step, gene annotation is manually added to the DGEList object. This step differs if the user's counts file is already annotated with gene names rather than Ensembl IDs. In this case, the annotations (last four lines) do not need to be manually added.

```
y <- DGEList(gExpr)

y <- calcNormFactors(y)

anno <- read.csv("anno.csv")

row.names(anno) <- anno$ensembl

anno <- anno[match(row.names(y), row.names(anno)), ]

y$genes <- anno
```

2. Create a diagnostic plot. Plot an MDS plot to assess variables that should be added as covariates (Fig. 2b). To do this, visually inspect each plot and examine the clustering of points. Ideally, points should cluster primarily by experimental group and less by technical/uninteresting covariates, such as batch, donor, line, or any other variable defined in your metadata. If points cluster by an undesired variable, that variable should be considered a covariate and will need to be added to the linear model in Step 9 to account for its effect on the data.

```
pdf(paste0("results/", experiment.title, "-2_mds.pdf"))

for (i in 1:length(colnames(meta))) {

  mds(y, meta[,i], colnames(meta)[i])

}

plotMDS(y)

dev.off()
```

Fitting a linear model

3. Design the model. Define the linear model for the differential expression analysis. Add the variable of interest and any variables (from the metadata columns analyzed in Step 8) that showed clustering from visual inspection of the MDS plot in Step 8. In this example, 'mod.gene' represents our variable of interest, whereas the MDS plot showed 'line' and 'donor' as additional covariates. Only use variables of interest as identified by clustering in the MDS plot.

```
design <- model.matrix(~ 0 + mod.gene + line + donor, meta)
```


Subsequently, remove the column name from the design matrix to ease its use. R will automatically create column names that unnecessarily combine the variable name and its value, making them cumbersome to type downstream. Here, the variable name is removed.

```
colnames(design)

colnames(design) <- gsub("mod.gene", "", colnames(design))

colnames(design)
```

4. Voom transforms data. The *voom()* function preps the data for linear modeling by converting counts to logCPM and computing weights for heteroscedasticity adjustment. It also generates a diagnostic plot for the mean-variance trend (Fig. 2c). Finally, visually inspect it for fit. Increase the cutoff for marginally expressed genes to improve fit if necessary.

```
v <- voom(y, design, plot = TRUE, save.plot = TRUE)
```

Save the plot.

```
pdf(paste0("results/", experiment.title, "-3_voom.pdf"))
plot(v$voom.xy, type = "p", pch=20, cex=0.16,
     main = "voom: Mean-variance trend",
     xlab = "log2(count size + 0.5)",
     ylab = "Sqrt(standard deviation)")
lines(v$voom.line, col="ref")
dev.off()
```

5. Fit model.

```
fit <- lmFit(v, design)
```

6. Define group comparisons (contrasts).

Define your comparisons of interest. Each comparison is assigned a name and a function using algebra. In the example below, the authors looked at perturbations of four genes (*SNAP91*, *TSNARE1*, *CLCN3*, and *FURIN* alone and in combination)). In addition to comparing the individual and combined perturbations, the authors also add equations for the additive and synergistic model.

```
cont.matrix <- makeContrasts (
  SNAP91a = sanp91 - ctrl, TSNARE1a = tsnare1 - control, CLCN3a = clcn3 - ctrl,
  FURINi = furin - ctrl, Additive = sanp91 + tsnare1 + clcn3 + furin - 4*ctrl,
  Combinatorial =
  all - all.ctrl, Synergy =
  all - tsnare1 - sanp91 - clcn3 - furin - all.ctrl + 4*ctrl, levels = design
)
```

7. Visualize the contrasts in a heat map.

```
cont.p <- t(cont.matrix)
```

```
h <- pheatmap(cont.p, display_numbers = T, number_format = "%.0f", color =
colorRampPalette(rev(brewer.pal(n = 10, name = "RdY1Bu")))(12), breaks = seq(-3, 1, by
= 0.5), cluster_cols = F, cluster_rows = F)

print(h)
```

8. Calculate coefficients and standard errors for each set.

```
fit.cont <- contrasts.fit(fit, cont.matrix)
```

Assessing differential expression

9. Performing empirical Bayes moderation by executing the *eBayes()* function from the limma package and computing statistics using the *decideTests()* function.

```
fit.cont <- eBayes(fit.cont)

plotSA(fit.cont, main = "Final model: Mean-variance trend", ylab = "Sqrt(standard
deviation)")

summa.fit = decideTests(fit.cont, adjust.method = "fdr")
```

10. Save DEG result tables created in Steps 12-15.

Create a list of all results. This will be used in Steps 19, 20, and 25 to run analyses for all comparisons.

```
res.list <- list()
for (i in 1:length(colnames(fit.cont$contrasts))) {
  x <- topTable(fit.cont, coef = i, sort.by = "p", n = Inf, confint = T)
  res.list[[i]] <- x
  names(res.list) <- colnames(fit.cont$contrasts)[i]
```

11. Save the DEG result plots.

Create mean difference (MA) and volcano plots for each contrast, with all significant DEGs and the top ten genes highlighted, respectively (Fig 2d).

```
pdf(paste0("\results/\",experiment.title,\"-4_volcano-md-plots.pdf\""))
par(mfrow=c(1,2))
for (i in 1:length(colnames(fit.cont$contrasts))) {
  plotMD(fit.cont,coef=i,status=summa.fit[,i],values=c(-1,1))
  volcanoplot(fit.cont,coef=i,highlight=10,
    names= fit.cont$genes$Gene_name)
}
dev.off()
par(mfrow=c(1,1))
```

Plot the expression of the top three DEGs in each contrast in all samples:

Change the `meta$mod.gene` variable (in “`p <- qplot(meta$mod.gene [...], fill = meta$mod.gene, [...])`”) to your variable of interest.

`scale_x_discrete` and `scale_fill_manual` are commented out and are optional addition for aesthetics. If used, their variables have to be changed to reflect the data.

```
pdf("results/", experiment.title, "-5_top3-expression-plots.pdf"))
par(mfrow = c(1,1))
for (i in 1:length(colnames(fit.cont$contrasts))) {
  x <- topTable(fit.cont, coef = i, sort.by = "p", n = Inf)
  cat("\n\n####", colnames(fit.cont$contrasts)[i], "\n\n")
  for(j in 1:3) {
    deg <- as.character(x[j, "ensembl"])
    p <- qplot(meta$mod.gene, v$E[deg, ],
      geom = boxplot, fill = meta$mod.gene, ylab = "Normalized expression",
      xlab = "group",
      main = paste0(j, ". DEG: ", as.character(x[j, "Gene_name"]))) +
      geom_jitter() +
      #scale_x_discrete(limits = c("ctrl", "sanp91", "tsnare1", "#clcn3", "furin",
      all.ctrl, "all")) +
      #scale_fill_manual(values = (c("orchid4", "grey", "#steelblue", "grey",
      "firebrick", "blue", "darkblue")) + rotate_x_text(angle = 45) +
      theme_bw(14) +
      theme(legend.position = "none", axis.text.x = element_text(
      angle = 45, hjust = 1, vjust = 1))
    print(p)
  }
}
dev.off()
```

Determining power to detect synergistic effects

12. Calculate power. Power calculations from this analysis can inform further studies. Ideally, power calculations should be performed on a similar dataset in advance.

Calculate the mean, standard error for all measured comparisons.

```
SE ← sqrt(fit.cont$s2.post) * fit.cont$stdev.unscaled
```

Calculate power. Choose the standard error matrix column names representing the additive and the combinatorial perturbation to calculate the median, standard error. Then, run the *power.compare.logFC()* function to create a power plot (Fig. 3a).

```
colnames(SE)
sig1 ← median(SE[, "Additive"])
sig2 ← median(SE[, "Combinatorial"])
g ← power.compare.logFC(sig1, sig2, N=4, N_other=c(4, 6, 8, 10, 14)),
alpha=0.05, n_tests=20000)
pdf(paste0("results/", experiment.title, "-6_synergy-power.pdf"))
print(g)
dev.off()
```

Fig. 4: Synergistic effect analysis output.

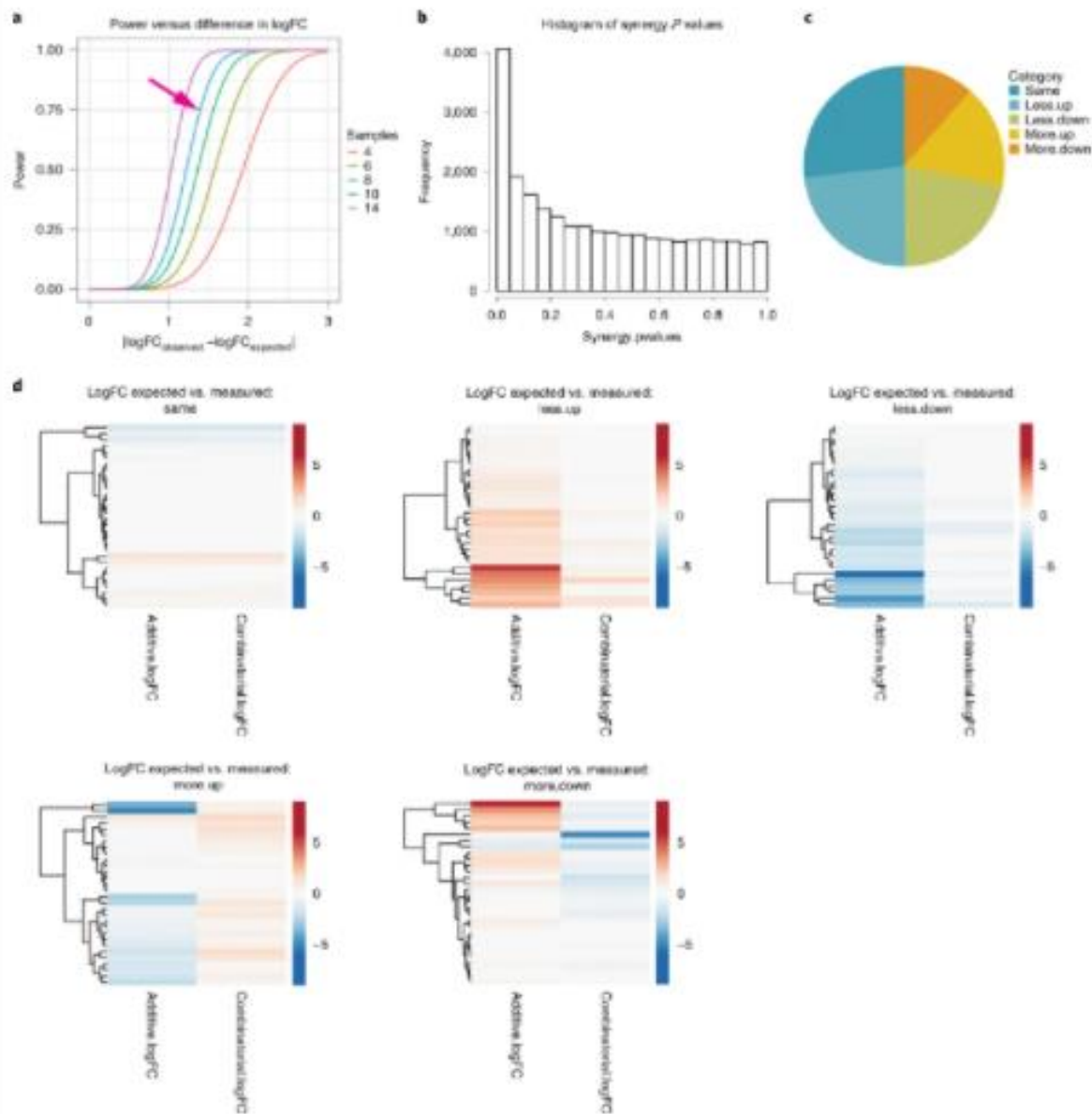


Figure 3: Synergistic effect analysis output. (Schrode et al., 2021)

Determine the extent of synergy

1. Calculate the synergy coefficient and percentage of synergistic DEGs (FDR <10%). Plot a histogram of all synergistic P values to visualize the distribution (Fig 3b).

```
synergy.pvalues <- res.list$Synergy$P.Value
```

```
pi1 <- 1 - qvalue(synergy.pvalues)$pi0
```

```

print(pi1)
pdf(paste0(results/", experiment.title, " - 7_synergy - coefficient.pdf"))
plot.new()
text(0.4, 0.75, labels = paste0("\n, round(pi1 * 100, 2),
"% non-null \np-values and \n",
round(sum(res.list$Synergy$adj. P.Val < 0.1) * 100/length(res.list
$Synergy$ensembl), 2),
" % of genes with \nsynergy FDR < 0.1"))
hist(synergy.pvalues)
dev.off()

```

Identification and categorization of synergistic genes

2. Define synergistic effect categories.

Determine an expression cutoff range. Here, \pm the mean, standard error for all empirically measured comparisons is used.

```
meanSE = mean(SE[, c(1, 2, 3, 4, 6)])
```

Create a table combining \log_2 FC columns of additive, combinatorial and synergy contrasts. Ensure the vectors (here: c(1, 2, 4, 9)) refer to the column indices of Ensembl ID, gene name, logFC, and adjust *P* value in the DEG result tables. If they do not, adjust the numbers to reflect these column indices correctly.

```

colnames(res.list$Additive)
log2FC.matrix ← Reduce(function(x,y)merge (x,y,by=c("ensembl\"",
"Gene_name"),all = TRUE),
list(res.list$Additive[,c(1,2,4,9)],res.list$Combinatorial[,c(1,2,4,9)],
res.list$Synergy[,c(1,2,4,9)]))
colnames(log2FC.matrix)
colnames(log2FC.matrix) ←c("Ensembl\"", "Gene_name\"", "\"Additive.logFC\"",
"Combinatorial.logFC\"", "\"Combinatorial.FDR\"", "Synergistic.logFC\"",
Synergistic.FDR")
rownames(log2FC.matrix) ← log2FC.matrix$Ensembl

```

Add a column assigning synergistic criteria to each gene using the *categorize.synergy()* function, which takes the expression cutoff and the previously created matrix into account.

```

log2FC.matrix ← categorize.synergy(log2FC.matrix, meanSE)
write.csv(log2FC.matrix, paste0("results/", ex))

```

3. Visualize the categories in a pie chart.

Create a table for the sums of synergistic category genes.

```
genes.per.category$category ←  
factor(genes.per.category$magnitude.syn,  
levels = c("same", "less.up", "less.down", "more.up", "more.down"))  
genes.per.category$percent ← paste0(round(genes.per.category$freq *  
100/sum(genes.per.category$freq), 0), "%")  
write.csv(genes.per.category, paste0("results/", experiment.title,  
"_gene – count_synergy – categories.csv"))
```

Plot a pie chart (Fig. 3c).

```
zissou ← wes_palette("Zissoul", 6, type = "continuous")  
q ← ggplot(genes.per.category, aes(x =, y = freq, fill = category)) +  
  geom_col() +  
  coord_polar("y", start = 0) +  
  scale_fill_manual(values = zissou) +  
  theme_void()  
pdf(paste0("results/", experiment.title,  
" – 8_synergy – categories_pie – chart.pdf"))  
print(q)  
dev.off()
```

4. Visualize the categories in heat maps.

Create heat maps of the log₂FC in the additive and the combinatorial comparisons for each synergy category (Fig. 3d).

```
pdf(paste0("results/", experiment.title,  
" – 9_heatmaps_log2FC – Add – vs – Combi.pdf"))  
for (i in 1:length(levels(log2FC.matrix$magnitude.syn))){  
  breaks ← c(seq(-6, -0.3, by = 0.1), seq(0.3, 6, by = 0.1))  
  breaks ← append(breaks, -9, 0)  
  breaks ← append(breaks, 9)  
  tmp ← log2FC.matrix[log2FC.matrix$magnitude.syn ==
```

```

      levels(log2FC.matrix$magnitude.syn)[i],
      c(Additive.logFC, "Combinatorial.logFC")]
    h ← pheatmap(tmp,
      kmeans_k = 30, cellwidth = 70, cellheight = 5,
      border_color = NA,
      breaks = breaks,
      cluster_cols = F,
      show_rownames = F,
      color = colorRampPalette(rev(brewer.pal(
        n = 9, Name = "RdBu")))(117),
      main = paste0("logFC expected vs. measured: \n",
        levels(log2FC.matrix$magnitude.syn)[i]))
    print(h)
  }
  dev.off()

```

Enrichment analysis

All comparisons: gene set enrichment analysis

5. Create a list containing the gene sets. Ensure that the gene sets are in the required format (gmt) using the functions `ids2indices()`, `geneIds()`, and `getGmt()`.

In this example, these are previously manually curated gene set groups, saved in the 'genesets' folder: disorder.gmt, connectivity.gmt, behavior.gmt, neural.gmt, head.gmt, presynapse.gmt, postsynapse.gmt, and synapse.gmt.

```

gs.list ← list("disorder" = ids2indices(geneIds(getGMT(
  ("genesets/disorder.gmt"))), id = v$genes$Gene_name), "behavior" =
  ids2indices(geneIds(getGmt("genesets/behavior.gmt")),
  ID = v$genes$Gene_name),
  "connectivity" = ids2indices(geneIds(getGmt("genesets/connectivity.gmt")),
  ID = v$genes$Gene_name),
  "head" = ids2indices(geneIds(getGmt("genesets/head.gmt")),
  ID = v$genes$Gene_name), "neural" =

```



```

ids2indices(geneIds(getGmt("genesets/neural.gmt")),
ID = v$genes$Gene_name),
"postsynapse" = ids2indices(geneIds(getGmt("genesets/postsynapse.gmt")),
ID = v$genes$Gene_name),
"presynapse" = ids2indices(geneIds(getGmt("genesets/presynapse.gmt")),
ID = v$genes$Gene_name),
"synapse" = ids2indices(geneIds(getGmt("genesets/synapse.gmt")),
ID = v$genes$Gene_name))

```

6. Create a custom color palette.

```

catcols = c("behavior" = "#8c3800", #brown "disorder" = "#3C4347", #darkgrey
"connectivity" = "#e0a81c", #mustard "head" = "#702658", #purple "neural" =
"#004878", #blue "postsynapse" = "#486030", #darkgreen "presynapse" =
"#a8c018", #lightgreen "synapse" = "#5c9340") #green
"#6CA7AD", "#B72415", "#A88C05", "#E06C03", "#653C82",
"#0287AA", "#AD5A60", "#9B0420")
show_col(catcols)

```

7. Run gene set enrichment using *camera*.

Loop through all contrasts in the `cont.matrix` object. Perform enrichment and visualize as scatter (Fig. 4a) and bar plots (Fig. 4b) using the custom *cameraplusplots()* function.

```

camera.res.list <- list()
for (j in 1:length(colnames(cont.matrix))) {
  print(paste0("Contrast: ", colnames(cont.matrix)[j]))
  pdf(paste0("results/", experiment.title, "-10_GSEA-",
            colnames(cont.matrix)[j], "-plots.pdf"))
  camera.res <- cameraplusplots(contrast = cont.matrix[,j],
                                vobject = v, genesetlist =
gs.list, design = design, catcolors = catcols,
                                title =
paste0(colnames(cont.matrix)[j]))
  dev.off()
  camera.res.list[[j]] <- camera.res
  names(camera.res.list)[j] <- colnames(cont.matrix)[j]
  write.csv(data.frame(camera.res), paste0("results/",
experiment.title, "_GSEA-", colnames(cont.matrix)[j], ".csv"))
}

```

Plot legend (Fig. 4a).

```
pdf(paste0("results/", experiment.title, " - 10_GSEA - plot - legend. pdf"))
plot(1, type = 'n', xlab = '', ylab = '', xaxt = 'n', yaxt = 'n', bty = 'n')
legend("center", names(catcols), cex = 1.2, fill = catcols)
dev.off()
```

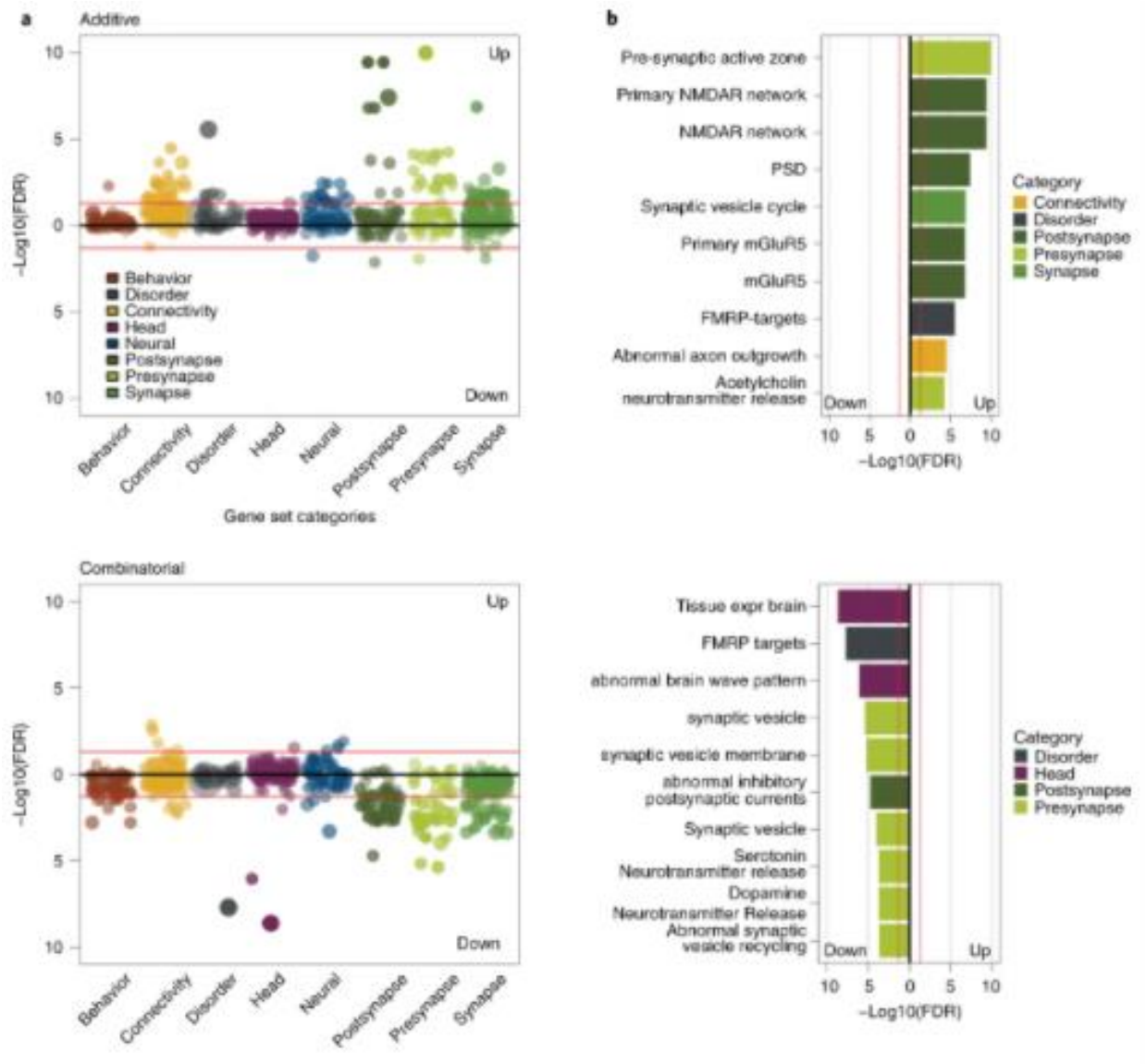


Figure 4: Gene set enrichment analysis output (Schrode et al., 2021)

Specific gene sets: gene ontology over-representation analysis

8. Create synergistic gene subsets.

Adjust FDR cutoff based on the subtlety of the synergistic effect. Here, a cutoff of synergistic FDR <1% was chosen.

```
log2FC.matrix.sub ← subset(log2FC.matrix, Synergistic.FDR < 0.01)
```

9. Stratify subset by synergy category using the custom *stratify.by.syn.cat()* function.

```
syn.cat.list ← stratify.by.syn.cat(log2FC.matrix.sub)
```

10. Define reference genes as all genes analyzed. Because lowly expressed genes were filtered out at the beginning of the protocol in Step 6, this can be interpreted as expressed genes.

```
allgenes ← as.character(y$genes$Gene_name)
```

11. Create a list of file paths that specify the directories in which the gene sets of interest are saved.

```
gs.list <- list("disorder" = "genesets/disorder.gmt", "behavior" =  
"genesets/behavior.gmt", "connectivity" = "genesets/connectivity.gmt", "head" =  
"genesets/head.gmt", "neural" = "genesets/neural.gmt", "postsynapse" =  
"genesets/postsynapse.gmt", "presynapse" = "genesets/presynapse.gmt",  
"synapse" = "genesets/synapse.gmt")
```

12. Run over-representation analysis for 'more up' and 'more down' synergy categories using the *WebGestaltR()* function. Here, 'more up' refers to synergistically upregulated genes when upregulation in the combinatorial condition is greater than that predicted by the additive model. In contrast, 'more down' refers to synergistically downregulated genes. These categories of interest were created in Step 20 using the *categorize.synergy()* function.

Loop through the 'more.up' and 'more.down' vectors in the previously created (Step 27) list object *syn.cat.list*.

Create a data frame for all results.

Visualize using the custom *oraplot()* function (Fig. 5).

```
for (i in 3:4){  
  tryCatch({  
    ora ← data.frame(matrix(ncol = 11,nrow = 0))  
    ora.list ← list()  
    goi ← syn.cat.list[[i]]  
    for (j in 1:length(gs.list)){
```

```

tryCatch({
  ora.s ← WebGestaltR(enrichMethod = "ORA",organism = "hsapiens",
    interestGene = goi,interestGeneType = "genesymbol",
    referenceGene = allgenes,referenceGeneType = "genesymbol",
    enrichDatabase = "others",enrichDatabaseFile = file.path(gs.list[j]),
    enrichDatabaseType = "genesymbol",sigMethod = "top",
    topThr = 50,minNum = 3,isOutput = F)

  ora.list[[j]] ← ora.s
  names(ora.list)[j] ← names(gs.list)[j]
  ora.list[[j]]$category ← names(ora.list[j])
  colnames(ora) ← names(ora.list[[1]])
  ora ← rbind.data.frame(ora,ora.list[[j]])
},error = function(e){cat("ERROR : ",conditionMessage(e),"\\n")})
}

write.csv(ora,paste0("results/",experiment.title,"_ORA-",
  levels(log2FC.matrix$magnitude.syn)[i],".csv"))

g ← oraplot(ora,catcols,paste0(levels(log2FC.matrix$magnitude.syn)[i]))
pdf(paste0("results/",experiment.title,"-11_ORA-",
  levels(log2FC.matrix$magnitude.syn)[i],"-plots.pdf"))
print(g)
dev.off()
},error = function(e){cat("ERROR : ",conditionMessage(e),"\\n")})
}

```

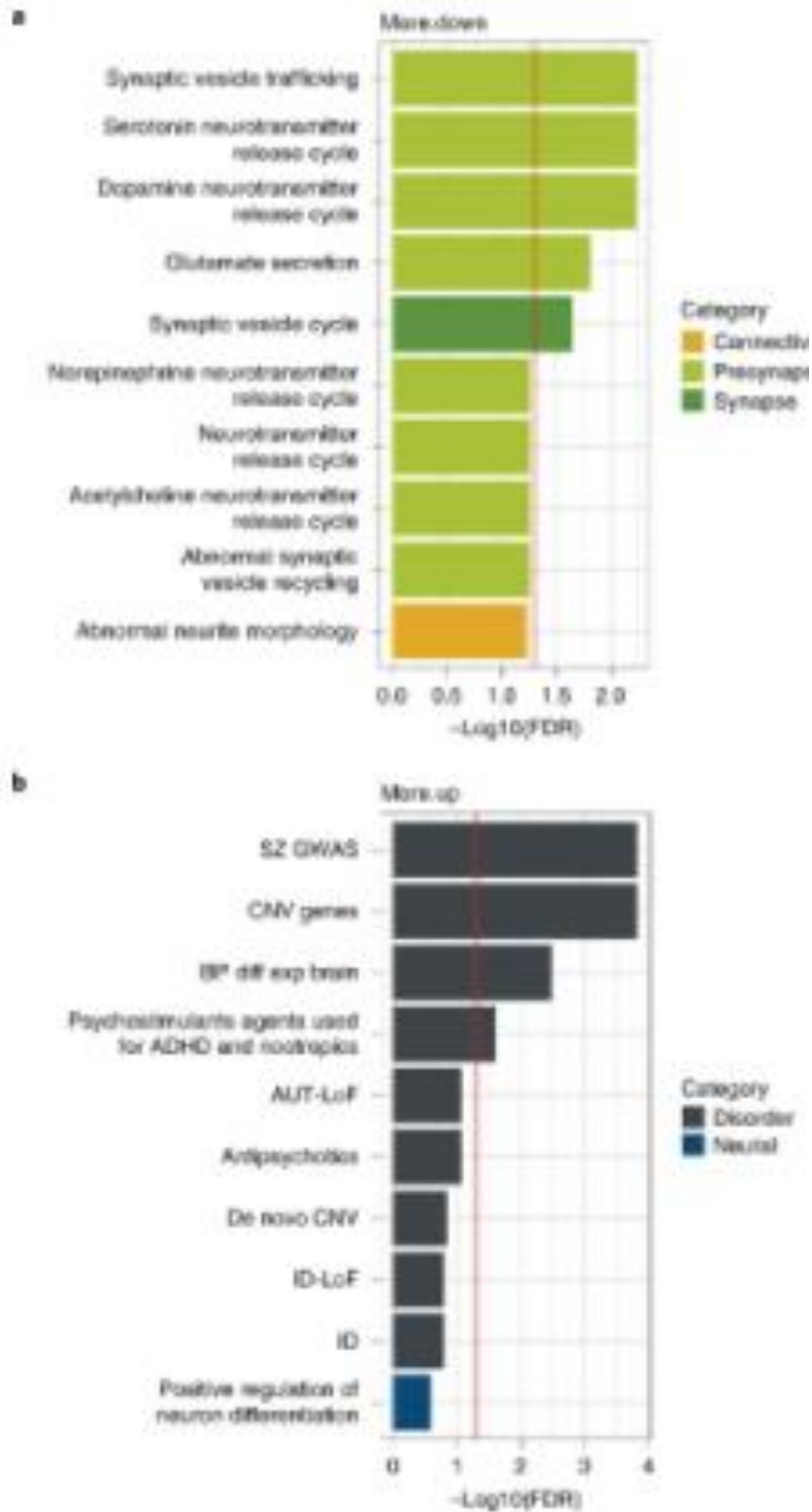


Figure 5: Over-representation analysis output. (Schrode et al., 2021)

Anticipated results

To demonstrate the functionality of this protocol, the authors analyzed the results from a study of the interactions between BET and MEK inhibitors to inhibit the growth of NRAS-mutant melanoma synergistically.

Differential expression analysis

Step 6 plots a visual correlation of gene counts and CPM (Fig 2a). The authors aimed to retain only genes with more than ten counts in at least four samples, as this allows them to visually determine the CPM value corresponding to ten counts (here: 0.25). This value is subsequently entered to define the 'keep' variable.

In step 8, an MDS plot displaying sample names is created and color-coded MDS plots for every metadata column. Figure 2b shows two examples of the four plots created from the example data. They show clearly that sample clustering was determined mainly by the donor and cell line, which will be subsequently added as covariates to the linear model design.

The `voom()` function used in Step 10 creates a mean-variance plot (Fig 2a). The fitted trend line is shown in red and can be used as a visual diagnostic tool to assess fit.

After linear modeling, empirical Bayes moderation is performed in Step 15, and the mean-variance trend for the final model is plotted.

The differential expression results for each comparison are saved as .csv files in Step 16, including Ensembl ID, gene symbol and description, logFC, left and right limit of the confidence interval for the logFC, average expression, *t*-statistic, *P*-value, adjusted *P*-value, and *B*-statistic for each gene.

Step 17 then creates the respective mean difference and volcano plots for each comparison (Fig. 2d).

Synergistic effect analysis

Step 18 calculates the power to detect synergistic effects, given the variances of the data at hand. The resulting power plot (Fig. 3a) visualizes the power (*y*) to resolve a specific synergistic logFC (*x*) for several sample sizes.

In Step 19, the synergy coefficient, π_1 , is calculated to determine a synergistic component in the data. Similarly, the fraction of genes with a synergistic FDR smaller than 10% is computed to determine the extent of synergy, and a histogram of all synergistic *P* values is created to visualize their distribution (Fig. 3b). Finally, all three are written into a .pdf file.

In Step 20, after creating a table to combine the logFCs and FDRs of the combinatorial, additive, and synergistic comparisons, genes are assigned a synergy category based on these variables. The results

are saved into a .csv file. These categories are then visualized in a pie chart (Step 21; Fig. 3c), and their logFCs are plotted in separate heat maps (Step 22; Fig. 3d). These categories describe how gene expression varies synergistically between the combinatorial and the additive model. For example, upregulated genes ('more up') have expression levels in the combinatorial condition more significant than the regular expression based on the additive model.

Gene set enrichment analysis

GSEA of all comparisons is performed in Step 25. The *cameraplusplots()* function used here also creates a scatter plot of all gene sets tested and their $-\log_{10}(\text{FDR})$ for each comparison (Fig. 4a). It also plots bar charts of the $-\log_{10}(\text{FDR})$ for the top five gene sets in each comparison (Fig. 4b). In the same loop, the results are saved as .csv files for each comparison. Finally, Step 30 creates bar charts for the over-representation analysis of upregulated ('more up') and downregulated ('more down') genes and saves the results as .csv files (Fig. 5).

CHiCANE

Chromosome conformation capture (3C) methods are techniques for identifying interactions between chromosomal regions based on chromatin conformation in the cell. The initial 3C methodology has been expanded to include chromosome conformation capture-on-chip (4C), chromosome conformation capture carbon copy (5C), and Hi-C. Hi-C utilizes proximity ligation and biotinylation of chromatin followed by high-throughput sequencing to produce a genome-wide interaction map. Compared to other 3C methods, the advantage of Hi-C is that it provides genome-wide coverage of interactions; however, high-resolution genome-wide coverage comes at very high sequencing costs. To achieve relatively low cost, high-resolution interaction data for a set of regions of interest, 3C, and Hi-C is combined with a capture enrichment step (Capture-C, Next Generation Capture-C, or Capture Hi-C). This targeted approach offers a powerful tool to delineate spatial and functional relationships at specific regions of the genome (rChi-C), such as those identified by genome-wide association studies (GWAS) and particular functional elements such as promoters (pChi-C, NG Capture C). For example, rChi-C has been used to identify putative target genes at breast cancer, colorectal cancer, and rheumatoid arthritis GWAS loci; pChi-C has successfully enabled linkage between genetic variants and their target genes, prioritizing new disease-associated genes and pathways. As an emerging tool, Chi-C offers an excellent system to generate high-resolution interaction maps of selected loci. However, Chi-C approaches generate data matrices representing unbalanced many-to-all interaction profiles between baits and target fragments (also referred to as a non-captured end), respectively, where baits are pre-defined, but the target fragments could be anywhere in the genome. This is inherently different from Hi-C protocols, where the interaction profiles are not biased by pre-defined baits, so both baits and other ends could be anywhere in the genome. Therefore, it remains unclear how best to model CHi-C data and handle the unique error profile of this methodology.

Identifying chromosomal regions that interact more than expected by chance (interaction peaks) requires an accurate model for the background rate of interactions. The capture enrichment step in CHi-C further complicates this model compared to a typical Hi-C experiment, as the capture dynamics will

differ depending on whether one or both ends of the interacting fragment pair are targeted in the capture enrichment step. To address this, CHiCANE is explicitly developed to analyze CHi-C data (Holgersen et al., 2021). CHiCANE is designed to handle the most frequent tasks involved in CHi-C analysis from pre-processing, intermediate data manipulation, interaction peak calling, and subsequent downstream analysis of interaction peaks. The protocol starts with an overview of pre-processing of sequencing data generated from CHi-C libraries using a combination of third-party tools to create alignment BAM files. Using CHiCANE, BAM files are processed into interaction tables by removing all reads where neither end maps to a capture bait nor those in which both ends map to the same fragment. Next, the restriction fragment identifies the map (requiring an overlap of $\geq 51\%$ of the read), and the number of reads linking each combination of restriction fragment is calculated. To identify interaction peaks, CHiCANE models the expected number of reads linking two restriction fragments as a function of the distance between the loci and the 'interactability' of the bait fragment; its inherent propensity to interact with other fragments.

While CHiCANE was conceived to call interaction peaks from rCHi-C experiments, it is also applicable to pCHi-C data. As CHiCANE utilizes BEDTools to process inputs into interaction data tables, it accepts BED files as the first two inputs, a baitmap and a restriction digest of the reference genome. HiCUP is a user-friendly pipeline commonly used in pre-processing CHi-C libraries, including mapping and pairing sequencing reads and creating alignment BAM files. HiCUP also includes a script for generating an in silico digest of the reference genome, which is used for filtering out common experiment artifacts. This in silico digest and the BAM files created by HiCUP is readily consumable by CHiCANE as starting inputs. Another beneficial mapping pipeline from which CHiCANE can utilize output is that from Arima Genomics, which is helpful for processing libraries generated using Arima Hi-C kits and Dovetail's Omni-C kits when combined with the target enrichment protocols. BAM files from either of these pipelines or any other mapping method and pairing read to produce BAM files can be used as CHiCANE input. The flexible toolkit built into CHiCANE is designed to adapt to various CHi-C and sequencing library processing protocols.

This protocol can be easily adapted to any species with appropriate input and annotation files. However, users must consider species-specific topological architecture and, number and length of chromosomes when interpreting the distribution of *cis* and *trans* interactions.

Procedure

Pre-processing data

Initial pre-processing steps are performed in the command line except where otherwise specified.

1. CHiCANE is designed to function with any BAM files from any processing pipeline. HiCUP is an example of one such pipeline used to obtain mapped and filtered read pairs and enzyme-digested genome. This pipeline takes FASTQ files with sequencing data as the input, and the output is Bowtie2-aligned BAM files with corresponding QC summary reports. It also provides a script for generating an in silico digest of the genome required for HiCUP filtering if necessary. A brief example for running the pipeline is given in the code below.

```
bowtie2-build GRCh38_latest_genomic.fna Human_GRCh38
```


2. Create fragments digest BED file. CHiCANE includes a helper function to convert the digested genome output from HiCUP into a BED file, `convert.hicup.digest.bed()`. Otherwise, CHiCANE can use any BED file containing the coordinates of the restriction enzyme digested fragments of the desired genome build. Throughout this protocol, the BED format is used for several operations.

```
hicup_digester --genome Human_GRCh38 --re1 A^AGCTT,HindIII *.fa
```

3. Create a baitmap. Intersect the digested fragment BED file with a BED file containing the coordinates of the probes which were included on the capture array probe using the `bedtools` 'intersect' command. Limit these results to a ≥ 20 bp overlap and sort for only unique fragments to speed up processing as follows:

```
bedtools intersect -wo -a
//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplemen
tary_Data/dataGRCh38_HindIII_fragments.bed -b
//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplemen
tary_Data/data/captured_fragments/MKmifsud2015_captured_fragme
nts.bed
//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplemen
tary_Data/data/captured_fragments/MKmifsud2015_fragments_overla
p_full.bed awk '$7>=20 {print $1 "\t" $2 "\t" $3}'
//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplemen
tary_Data/data/captured_fragments/MKmifsud2015_fragments_overla
p_full.bed | uniq
//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplemen
tary_Data/data/captured_fragments/MKmifsud2015_captured_fragme
nts.bed
```

4. Create an interaction data object. Rather than running CHiCANE as a wrapper function on the BAM and BED files directly, the user can pre-process data by running the `prepare.data()` function on the BAM and BED files to obtain a data table of all interactions containing details of the fragments and the number of reads linking the two fragments. This step is helpful if a user wants to adjust for a covariate that is not already present in the data table (e.g., GC content) and will also speed up the model fitting, although the pre-processing will take some additional time. Another use case is to obtain interaction data on individual replicates for replicating comparison with CHiCANE's `compare.replicates()` helper function before running interaction calling with the replicate merging option. Pass the BAM(s) as the 'bam' parameter, the baitmap as the 'baits' parameter, and the restriction digested fragments as the 'fragments' parameter.

```
interaction.data <- prepare.data(
  bam =
    '//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Su
    pplementary_Data/data/bams/MK/sample.bam',
```

```

baits =
'//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Su
plementary_Data/data/baits/MK/baitmap.bed',

fragments =
'//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Su
plementary_Data/data/captured_fragments/MKmifsud2015_c
aptured_fragments.bed')

write.table(

    interaction.data,

    file =
'//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Su
plementary_Data/data/chicane/MK/interaction_calls.txt',

    row.names = FALSE,

    quote = FALSE,

    sep = '\t'

)

```

- Using a four-cutter enzyme (e.g., DpnIII) or multiple restriction enzymes rather than a six-cutter enzyme (e.g., HindIII) might improve the resolution of a library by generating shorter (DpnII) and more random restriction fragments. However, fewer reads are mapped to each pair of interacting fragments for the same amount of data. As the number of di-tags at any pair of fragments will influence the power of the analysis, at a given sequencing depth, these libraries might lack power for a single restriction fragment-based analysis. If sequencing to a greater depth is not an option, the interaction data from these libraries could be further processed by merging neighboring fragments in pre-defined window size. To determine the window size, the fragment-wise count's distribution of an experiment should be assessed. The size of this window is likely driven by the sequencing depth of the library and, therefore, the power to detect significant interactions. It is helpful to try a few thresholds between 1-5 kb and assess the impact of the varying thresholds on the count's distribution per fragment. Users can accomplish this by creating a fragments BED file based on the desired window size. This file simply specifies fragments according to the window size. This window-based fragments file intersects with the bait coordinates to create a baitmap based on the desired window size. These window size-based BED files can then be input as usual when running CHiCANE. Similarly, the baitmap and fragments BED files could also be designed to cluster interactions within loci as specified.

CHiCANE interaction calling

- The main function `chicane()` is typically run as a wrapper call on the BAM and BED files directly, as it will internally convert each BAM file to a text file then use BEDTools to overlap this file with the captured fragments and identify the restriction fragments corresponding to each read.

Alternatively, `chicane()` can be run in a stepwise fashion on a pre-processed interaction data table generated through the `prepare.data()` function. To use a pre-processed text file, pass either a data table object resulting from `prepare.data()` or the path to such a table to the 'interactions' parameter as shown below:

```
chicane.results <- chicane(interactions = interaction.data)
```

7. While the negative binomial distribution is the default, CHiCANE supports several different distributions which can be specified with the 'distribution' parameter.

```
chicane.results <- chicane(
  bam =
    '//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplementary_Data/data/bams/MK/sample.bam',
  baits =
    '//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplementary_Data/data/baits/MK/baitmap.bed',
  fragments =
    '//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplementary_Data/data/captured_fragments/MKmfisud2015_captured_fragments.bed'
  distribution = 'poisson'
)
```

8. As the default, only interactions that are detected at least once are included in the data. To also include zero counts, use the 'include.zeros' parameter: `include.zeros = 'cis'` includes all zero counts for bait/target combinations on the same chromosome.

```
chicane.results <- chicane (
  bam =
    '//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplementary_Data/data/bams/MK/sample.bam',
  baits =
    '//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplementary_Data/data/baits/MK/baitmap.bed',
  fragments =
    '//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplementary_Data/data/captured_fragments/MKmfisud2015_captured_fragments.bed'
  include.zeros = 'cis'
)
```

9. CHiCANE tries to split data into an optimal number of distance bins based on the size of the dataset, such that the resulting datasets are large enough for the model to be fit. To override this default behavior, set the desired number of distance bins with the 'distance.bins' parameter. For example, the following call will result in the count model fitted separately in each of the 50 data bins (*trans* interactions are fitted separately from *cis* interactions). Specifying multiple CPU cores through the 'cores' parameter would run multiple model fits in parallel to decrease the runtime.

```
chicane.results <- chicane (  
  bam =  
    '//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supple  
    mentary_Data/data/bams/MK/sample.bam',  
  baits =  
    '//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supple  
    mentary_Data/data/baits/MK/baitmap.bed',  
  fragments =  
    '//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supple  
    mentary_Data/data/captured_fragments/MKmfisud2015_captured  
    _fragments.bed'  
  distance.bins = 50  
)
```

10. CHiCANE includes all baits and targets when fitting the model; however, for users wanting to filter out fragments with high or low 'interactivity,' the `chicane()` call supports 'bait.filters' and 'target.filters' parameters. These parameters accept a vector of length two, the first element corresponding to the lower limit as a proportion of fragments and the second element corresponding to the upper limit and a proportion of fragments. Filtering the fragments will affect multiple testing corrections by changing the number of tests performed. In the following example, the baits with the lowest 20% or highest 30% of trans counts will be removed before fitting the model.

```
chicane.results <- chicane (  
  bam =  
    '//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supple  
    mentary_Data/data/bams/MK/sample.bam',  
  baits =  
    '//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supple  
    mentary_Data/data/baits/MK/baitmap.bed',  
  fragments =  
    '//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supple
```

```
plementary_Data/data/captured_fragments/MKmifsud2015_captured
_fragments.bed'

bait.filters = c(0.2, 0.7)

)
```

11. Users who would like to adjust for additional covariates can add them to the model through the 'adjustment.terms' parameter. This parameter will accept any column already present in the data and also accepts more complex expressions such as log transformations and multiple adjustment terms passed as a vector. To adjust for a covariate not already present in the data, use the prepare.data() function, then add extra columns for the desired covariates to the resulting data table. Next, run chicane() with this table passed to the 'interactions' parameter and specify the column name(s) through the 'adjustment.terms' parameters as shown in the example below, where an additional column (gc.content) for GC content (of bait or target fragments) was separately added to the interactions data table file.

```
chicane.results <- chicane(

  interactions =
    '//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Su
    pplementary_Data/data/chicane/MK/interaction_calls.txt',

)
```

12. The chicane() function will create model fit statistics and plots, stored in the directory path to the 'interim.data.dir' parameter. The default 'NULL' will skip the intermediate results. Therefore, it is essential to set this parameter to the desired directory to evaluate the quality of the model fits across distance bins.
13. CHiCANE's default setting is to perform multiple testing corrections separately for each bait; however, global multiple testing corrections is also supported by setting the 'multiple.testing.correction' parameter to 'global' when calling chicane().

Interpretation and post-processing

14. CHiCANE calls can readily be converted to a BED file by selecting columns 3-5 for target fragments and 6-8 for bait fragments. These columns help perform bedtools 'intersect' with several genomic data formats for annotation and enrichment testing of target and bait fragments. For instance, to annotate the genes to which the target fragments map, a BED file can be made from the interaction calls' target fragments, which can then be intersected with a genome annotation file of choice. The following calls are an example of how to pre-process the human refGene file into a BED file to then intersect the CHiCANE interaction peaks' target fragments BED file with the annotation BED file made from refGene. The bedtools intersect -wao option shown below keeps data from both BED files in the output and reports the number of base pairs by which the intersect files overlap. Furthermore, data for which the -a file has no overlap with the -b file is still reported with overlap = 0.

```

zcat refGene.txt.gz | awk '{print $3 "\t" $5 "\t" $6 "\t" $13}' | sort -u > refGene.bed

bedtools intersect -wao -a
//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplementary_Data/data/
chicane/MK/interaction_calls.txt -b
//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplementary_Data/data/
chicane/refGene.bed
//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplementary_Data/data/
chicane/interaction_calls_significant_targets_genes.bed

```

15. To examine functional associations between a locus of interest and the rChi-C identified target genes, a user can perform an eQTL analysis to model the relationship between a GWAS SNP and mRNA abundance of its putative target gene in a cohort of interest. This step requires accessing matched genotype data and mRNA abundance profiles. This relationship is modeled directly in R by fitting a linear model as shown below:

```
lm(mrna.abundance ~ genotype)
```

where 'mrna.abundance' is a continuous variable, and 'genotype' is either a categorical variable (e.g., 'A/A,' 'A/G,' and 'G/G') or a discrete variable encoding genotypes as 0, 1, and 2.

As variation in mRNA abundance is often only in part explained by the genotype, it should be associated with the residual mRNA abundance by regressing out the effect of other covariates, which may also explain variance in mRNA abundance. Hence, when copy number data and DNA methylation data are also available, an appropriate eQTL model in R is specified as:

```
lm(mrna.abundance ~ genotype + copy.number.state + methylation)
```

where 'copy.number.state' is either a categorical variable (e.g., 'loss,' 'neutral,' and 'gain') or a discrete variable specifying the total number of gene copies, and 'methylation' is either a continuous variable (e.g., methylation beta values) or a categorical variable (e.g., 'methylated' and 'unmethylated').

16. For genome-wide visualization of interactions and integration with public epigenomics datasets, a user can readily convert CHiCANE's results to a format compatible with the WashU Epigenome Browser with the create.standard.format() helper function, which simply takes the path to your *q*-value-filtered CHiCANE interaction calling results file and a path to where you would like the Epigenome Browser file to be output to.

```

create.standard.format(

    chicane.results =
    '//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplementary_D
ata/data/chicane/interaction_calls_significant_standard.txt'

)

```

Once the standard format interaction files created with the call above have been saved, upload them to the WashU Epigenome Browser via the dropdown Tracks > Local Text Tracks, then select 'long-range format CHiCANE' as the text file type. Once the track is loaded, right-click on the track and set the 'Display mode' to ARC.

17. CHiCANE's results are also readily compatible with the Gviz package for visualizing locus-specific data alongside genomic annotation tracks. Furthermore, the CHiCANE package includes a helper function to create a multi-track locus plot. The function `create.locus.plot()` creates a locus plot with ideogram, genome axis, genome annotation, feature annotation, counts, and interaction tracks. Specify the genome build with the 'genome' parameter and set the coordinates for the locus to be examined with the 'chr,' 'start,' and 'end' parameters. Pass the preferred genome annotation file through the 'gene.data' parameter as the path to the desired GTF file and set the feature annotation track by passing the path to the desired GTF file and set the feature annotation track by passing the path to a BED file containing coordinates of the selected genomic feature, such as topologically associating domains (TADs), regulatory motifs, or the baitmap created as CHiCANE input to display the baits. Unfiltered CHiCANE interaction calls are included by passing the appropriate path to the 'interaction.data' parameter, which creates a data track to display the raw fragment counts. The function filters these calls for significance with the threshold set in the 'fdr.filter' parameter (default = 0.05) to plot selected interactions on the interactions track. Finally, the function saves the locus plot figure to the path specified in the 'file.name' parameter. The `create.locus.plot()` function also allows users to optimize figure layout through additional graphics parameters. The example code below assumes using the default FDR, height, width, and track heights and uses the baitmap to include baits as the feature track to examine the locus chr2:111100000-112200000 of hg38:

```
create.locus.plot (  
    genome = 'hg38',  
    chr = 'chr2', start = 111100000, end = 112200000,  
    gene.data =  
    '//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplementary_D  
    ata/data/chicane/gene_data.gtf',  
    genomic.features =  
    '//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplementary_D  
    ata/data/baits/MK/baitmap.bed',  
)
```

18. To facilitate a more precise understanding of histone enrichment, a helper function can be used to randomize interactions matching the number and proportion of observed interactions in distance bins of 0-100 kb, 100 kb-1 Mb, 1-10 Mb, >10 Mb, and *trans* to serve as a background. The function `stratified.enrichment.sample()` takes the path to a data table with

the non-significant interactions as the 'nonsignificant.results' parameter and the path to a data table containing the significant interactions as the 'significant.results' parameter:

```
interaction.calls.random <- stratified.enrichment.sample (
  nonsignificant.results =
    '//ws1$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_
    Supplementary_Data/data/chicane/MK/interaction_calls_n
    onsignificant.txt',
  significant.results =
    '//ws1$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_
    Supplementary_Data/data/chicane/MK/interaction_calls.t
    xt'
)

write.table (
  interaction.calls.random[, c('target.chr',
    'target.start', 'target.end', 'q.value')],
  file =
    '//ws1$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_
    Supplementary_Data/data/chicane/MK/interaction_calls_r
    andom_targets.bed',
  row.names = FALSE,
  col.names = FALSE,
  quote = FALSE,
  sep = '\t'
)
```

In practice, this is run multiple times to generate a large number of background distributions in order to infer a stable estimate of background enrichment. Once random partitions are created, save them as separate BED files and overlap them with the BED file of histone marks using the bedtools 'intersect' command below. Note that the '-c' option reports a count of the number of overlaps including 0 for -a entries with no overlap with -b:

```
bedtools intersect -a
//ws1$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supple
mentary_Data/data/chicane/MK/interaction_calls_random_targe
ts.bed -b
//ws1$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supple
mentary_Data/data/chicane/MK/histone_coordinates.bed -c
//ws1$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supple
mentary_Data/data/chicane/MK/interaction_calls_random_targe
ts_histones.bed
```

The proportion of interactions overlapping with histone marks across multiple randomly sampled background datasets is summarized as the mean or median, compared with the observed enrichment of histone marks in the original dataset.

19. TAD boundaries are another useful data type that can be stored in a BED file. However, users should intersect them with the full-length coordinates of an interaction rather than individual fragments to test if the interaction is within or across TAD boundaries. To obtain a BED file for this purpose from CHiCANE calls, first, limit interactions to *cis* interactions only by removing any calls with 'NA' in the 'distance' column. Then for each *cis* interaction, add to a data table the chromosome, the minimum of the fragment starts, and the maximum of the fragment ends. Write these three columns as a BED file to correspond to the coordinates of each interaction. The following is a bedtools call suited to intersect the resulting full interaction coordinates BED file with TAD coordinates, where '-f 1.0' ensures complete overlap between the coordinates and the '-c' option reports a count of the number of overlaps:

```
bedtools intersect -a
//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplementary_Data/data/
chicane/interaction_calls_significant_full_length.bed -b
//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplementary_Data/data/
chicane/TAD_coordinates.bed -c -f 1.0
//wsl$/Ubuntu/home/melissachua/hicup/chicane/CHiCANE_Supplementary_Data/data/
chicane/interaction_calls_full_length_TADs.bed
```

20. Since the negative binomial and Poisson models are nested in structure, with Poisson being a particular case of negative binomial, these models can be quantitatively compared by assessing their log-likelihood estimates. Of the two models, the one with the higher log-likelihood estimate can be considered better. To estimate the statistical significance of the difference between the two likelihood estimates, apply a likelihood ratio test of the two estimates; however, this should be done on matched distance bins. For instance, to test whether the log-likelihood estimate for the first distance bin of the negative binomial model is significantly superior to the Poisson model using the likelihood ratio test.

```
# Extract the log-likelihood estimates and degrees of freedom
```

```
loglik.nb <- -100.123
```

```
df.nb <- 4
```

```
loglik.poisson <-- -120.456
```

```
df.poisson <- 3
```

```
# Estimate the difference in log-likelihood estimates and compare with the chi-squared
distribution
```

```
models.difference <- 2 * (loglik.nb - loglik.poisson)
```

```
p.value <- pchisq(models.difference, df = df.nb - df.poisson, lower.tail = FALSE)
```

21. The workflow above can be automated through a workflow engine such as Common Workflow Language, Nextflow, or Snakemake. These workflow engines will help create a reusable workflow that can run in an automated and parallel fashion. Further, these workflows are easy to share and deploy to HPC and cloud environments, enabling reproducible research.

Anticipated results

Assessing alignments and quality of CHi-C library

If using the HiCUP pipeline for processing sequencing data, a report will be generated for each sample. This report includes assessing the truncation of reads at the restriction enzyme sites and subsequent alignment of truncated reads. It also reports the percentage of di-tags deemed ‘valid’ by the HiCUP filter and those for each of the common artifacts removed by the filter. Additionally, the distribution of di-tag lengths is shown, and de-duplication results are presented and broken down by *cis*-close, *cis*-far, and *trans*. In an excellent library, the majority of both forward and reverse reads should result in unique alignments and successful reading pairs.

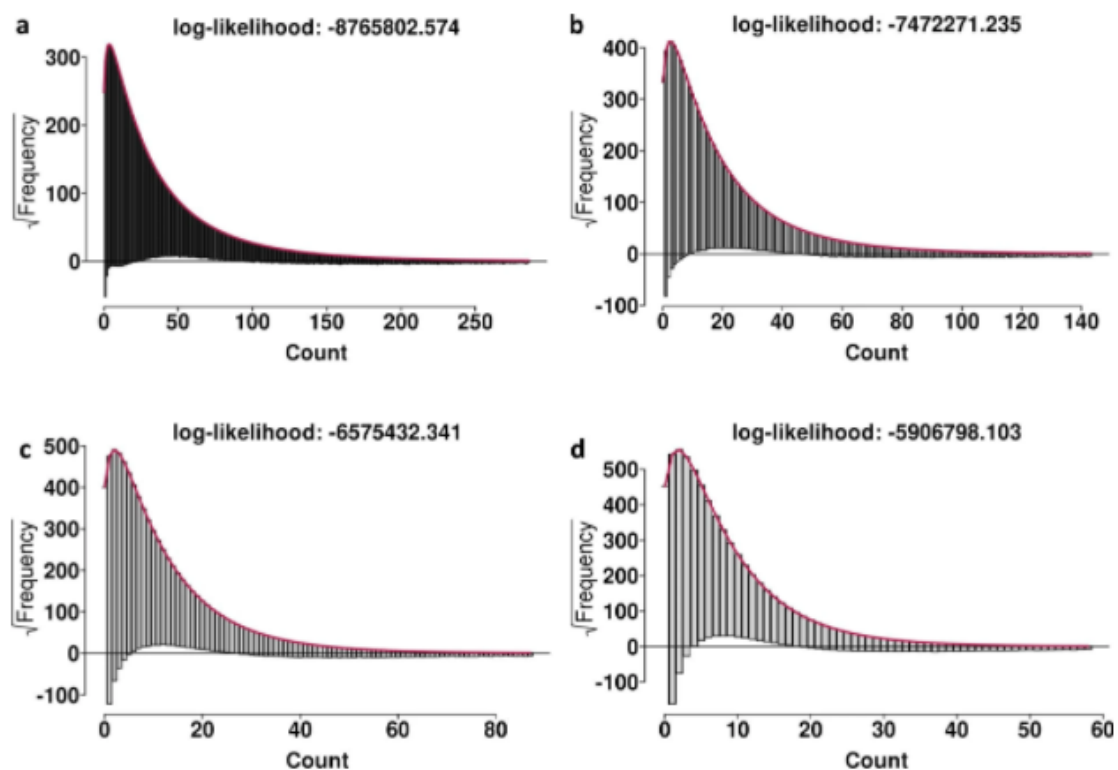
Conversely, an insufficient number of unique alignments or read pairing indicates poor libraries. Another indicator of an excellent library is a high proportion of valid di-tags and a high proportion of *cis* interactions. However, these suggested proportions should be interpreted alongside the overall sequencing depth and coverage per captured fragment. Further statistics on library and sequencing characteristics are detailed in the HiCUP report, and guidelines on interpreting these reports are covered in the HiCUP documentation.

Assessing interaction calls of CHiCANE on promoter CHi-C data

CHiCANE is well-suited to promoter CHi-C analysis. Here, the megakaryocytes (MK) dataset from the Javierre pChI-C study is analyzed. Following replicate merging and interaction calling, this dataset results in 82,830 interaction peaks (q -value < 0.05). As a representative example, the head() of *cis* interactions for chromosome 9 of MK libraries with the most relevant columns selected for readability are shown in table 1. CHiCANE’s default settings using the negative binomial distribution demonstrated a good model fitting to the observed count's data (Fig 1). In analyzing the interaction peaks by type, it can be observed in this pChI-C data that the vast majority are non-bait-to-bait, *cis* interaction within the mid-range distance of 100 kb-5 Mb (Fig 2). Table 1 shows an example of top *cis* interaction peaks called by CHiCANE from chromosome 9 of the MK library using the head() function in R, which defaults to the top six records. Selected columns are shown in the table.

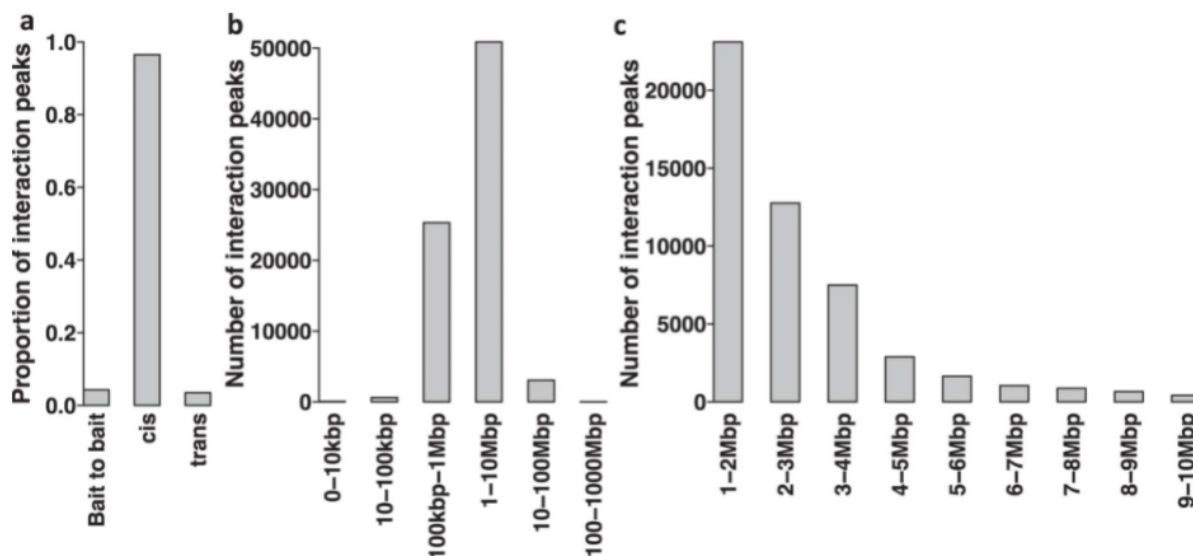
target.id	bait.id	bait. to. bait	bait. trans. count	target.trans. count	distance	count	expected	p.value	q.value
chr9:90100334–90101004	chr9:89602099–89606520	FALSE	14,561	224	496359.5	798	14.4906228	2.25E–43	4.50E–39
chr9:104741849–104745548	chr9:1015333300–101533731	FALSE	20,469	397	3210183	49	2.52376298	5.43E–41	1.67E–36
chr9:115593657–115594456	chr9:116565182–116568871	FALSE	9,599	189	972970	207	4.48286706	1.13E–38	1.70E–34
chr9:104741035–104741849	chr9:1015333300–101533731	FALSE	20,469	226	3207926.5	46	2.52469703	1.30E–37	2.00E–33
chr9:104741849–104745548	chr9:101533731–101535624	FALSE	12,139	397	3209021	42	1.94458688	1.26E–37	2.56E–33
chr9:104745548–104753293	chr9:101533731–101535624	TRUE	12,139	20309	3214743	101	3.83488463	1.27E–36	1.29E–32

Table 1: Megakaryocytes (MK) interaction peaks (Holgersen et al., 2021)



Representative examples of hanging rootograms depicting the negative binomial model fits on the Javierre MK library. Observed counts are shown as histogram bins (gray bars) while the CHiCANE fitted expected counts distribution is in red. The y-axis represents square root transformed density estimates of observed (gray bars) and expected (red line) counts. For observed counts, the height of the bars is shifted to align the top of the bar with the expected counts fit. Bars above and below the reference line (x-axis) indicate over- and under-prediction by the CHiCANE model, respectively.

Fig 1: Model fitting of counts data (pChi-C library) (Holgersen et al., 2021)



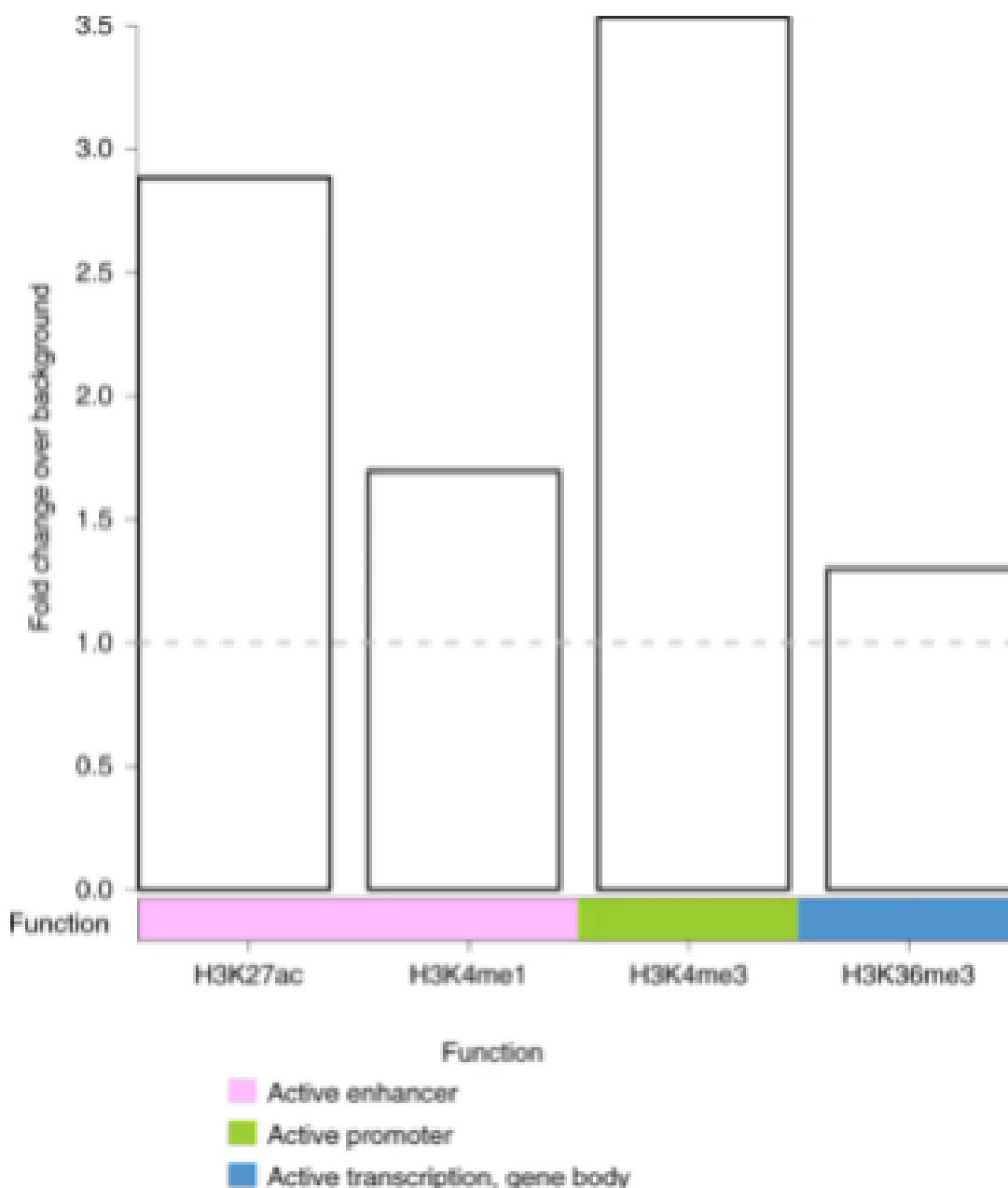
Examples of interpretation of interaction calling on Javierre MK library. **a**, Bar plots showing the proportion of interaction peaks (q-value < 0.05) by type (cis interactions include bait-to-bait interactions). **b**, Bar plots showing the number of interaction peaks (q-value < 0.05) across distance bins. **c**, Bar plots showing breakdown of region 1-10 Mb shown in (b).

Fig 2: Interpretation of interaction peaks by distance (Holgersen et al., 2021)

Enrichment of enhancer marks

Histone modifications are associated with epigenetic regulation of transcription. As such, the authors examined histone marks that colocalize with interacting fragments to determine both locus-specific properties, e.g., enhancer-promoter contacts, and global properties, such as enrichment of histone marks. Using the MK dataset, which had histone mark data from the International Human Epigenome Consortium for the same cell type, we performed enrichment analysis of four histone marks (H3K27ac, H3K36me3, H3K4me1, and H3K4me3). These marks are associated with active enhancers, promoters, and transcribed regions. As the baits in the MK dataset were designed to capture promoters, baits and all bait-to-bait interactions were excluded such that the non-bait interacting fragments (target fragments) of each interaction peak were isolated from CHiCANE calls. The coordinates of the target fragments can be overlapped with the coordinates of histone marks using the bedtools 'intersect' command. A similar analysis can then be repeated on random sets of interactions created with the `stratified.enrichment.sample()` function, creating an estimate of the background distribution of enrichment of a given mark. The resulting data can be examined either for a specific histone mark in the observed dataset and, likewise, the mean proportion estimated from 100 distance-matched randomly sampled datasets. For example, a fold change of >1 shows that the observed data is enriched for the tested enhancer mark compared to the randomly sampled datasets (Fig 3). A similar analysis could be repeated for coordinates of other vital regulatory features of interest, such as the Polycomb-mediated transcriptional repression mark H3K27me3 or the binding motifs of chromatin conformation mediator CTCF transcription factor and RAD21 double-strand break repair protein.

Fig. 7: Enrichment of enhancer marks.



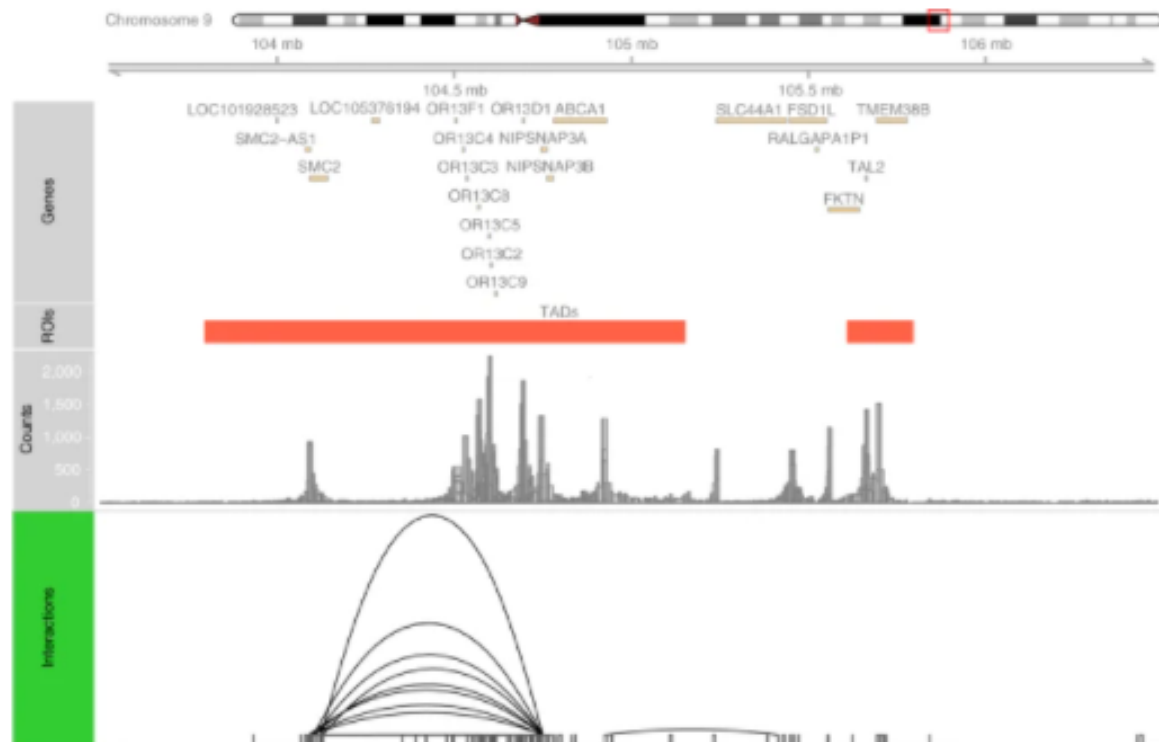
Fold change of enrichment of four histone marks compared to the enrichment of these marks in a distance-matched (to observed data) average background distribution. Fold change ≥ 1 means enrichment in the MK library compared to a random set of interactions.

[Full size image >](#)

Figure 3: Enrichment of enhancer marks (Holgersen et al., 2021)

Overlap with topologically associating domains

TADs are key subunits of chromatin compartmentalization and define the regulatory landscape. Genes tend to be coregulated when they are within the same TAD boundaries, and hence TADs serve as a useful chromatin feature to evaluate the functional activity of ChIP interactions. A BED file containing coordinates of TAD boundaries can be intersected with a BED file with full-length coordinates of interaction peaks, or TAD boundaries can be included as features in a locus plot created with the `creating.locus.plot()` function. To include TADs in a locus plot, we set the 'genomic.features' parameter to a BED file with the TAD coordinates for MK and assess chromatin interactions chr9:103500000-106500000 locus alongside TAD boundaries. This region demonstrates interactions falling largely within a TAD region rather than crossing the TAD boundaries (Fig 4). TAD boundaries are known for the enrichment of CTCF binding proteins, and any aberrant or disruption in the orientation of CTCF sites can lead to changes in TAD boundaries. Therefore, intersecting with CTCF binding site coordinates from the matching cell type or one with a similar genetic background as can further help to prioritize truly functional interaction peaks.



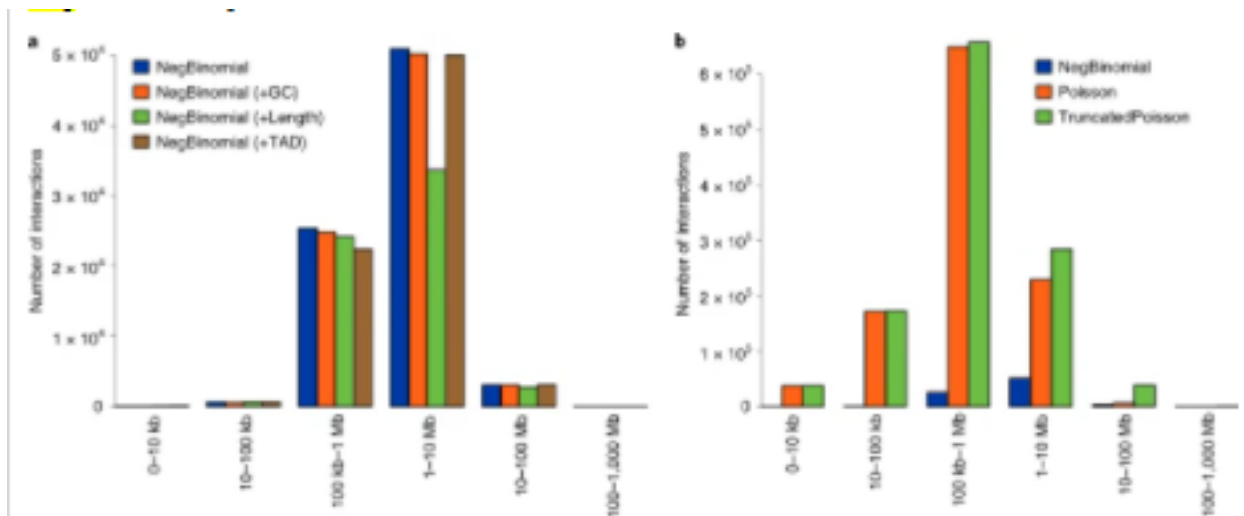
Interaction peaks called by CHiCANE in the Javierre MK library. Regions of interest (ROIs) are TADs (red) included in the Javierre study. The 'Counts' track shows raw counts of fragments from all interactions spanning this region. Interaction peaks are shown as arcs, where the height of the arc is proportional to the $-\log_{10} q$ -value of the interaction peak. Only CHiCANE interaction peaks with q -values < 0.05 entirely within boundaries of the locus are shown.

Figure 4: MK locus plot with TADs. (Holgersen et al., 2021)

Exploring the impact of covariates and statistical distributions

As CHiCANE was conceived as a toolkit for processing and modeling CHi-C count data, it provides several additional parameters to interrogate CHi-C libraries beyond the default settings entirely. Of these, two key parameters are: (i) 'adjustment.terms', which allows for adjustment for potentially confounding variables, and (ii) 'distribution,' which serves to fit alternative distributions to model count data. Adjusting the model is often necessary if either the known global properties of the genome act as potential biases (e.g., high genome instability, heterogeneity in GC content) or when annotation of the resulting interaction peaks highlights potential biases (e.g., selection of longer non-captured/target fragments). These scenarios pose a question as to whether the observed interaction peaks are genuinely independent or observed confounders. To adjust the model fitting process, thereby identifying interaction peaks independent of potential confounders, the `chicane()` call can be parameterized using 'adjustment.terms' specifying any number of additional variables. Using the pCHi-C MK library, this parameter is exemplified

by adjusting the default negative binomial model with three variables separately: % GC content, length of interacting fragments, and presence/absence of interaction within a TAD boundary (Fig 5a). Compared to the default settings (blue bars), adjusting for the bait and target fragment lengths resulted in a substantially reduced interaction peak in the 1-10 Mb distance bin (green bar). It sheds light on variables that might bias interpretations. Next, the authors explored the impact of the choice of the underlying distribution to model the observed count's data. Although not directly applicable to the MK library, which can be appropriately modeled with the negative binomial distribution (Fig 1), the authors examined the impact on CHiCANE's interaction calling when using the Poisson and truncated Poisson distributions. The resulting interaction peaks showed substantial differences compared to the default negative binomial model, with both the Poisson and truncated Poisson models having identified an approximately tenfold higher number of interaction peaks across all distance bins (Fig 5b). These numbers are likely to contain a more significant proportion of false positives, given that the negative binomial model's fits were reasonable (Fig 1). However, these data might be helpful for a discovery-focused study where a low type II error rate is desired at the cost of a high type I error rate, provided the model fits for these models explain the data adequately well. These alternative model fits might also be appropriate when the default negative binomial distribution does not explain the data very well. During the model fitting process, if the count data exhibit a lack of overdispersion, CHiCANE automatically switches to the Poisson distribution. Quantitative assessment of the quality of model fits between the default negative binomial and Poisson distributions can be done by comparing the fitted log-likelihood estimates of the matched distance bins, with the higher estimate indicating a better fit. The likelihood ratio test statistic can be used to identify the statistical significance of the differential log-likelihood estimates for a given distance bin.



Assessment of interaction peaks (q -value < 0.05) by distance bins using a range of statistical models applied through CHiCANE to the Javierre MK library. **a**, Bar plots showing the number of interaction peaks (q -value < 0.05) across distance bins when data was modeled using CHiCANE's default negative binomial distribution (blue), which allows for overdispersion, further adjusted for % GC content of bait and target fragments (orange), lengths of bait and target fragments (green) and TADs specified as whether both bait and target fragments were within a TAD boundary (brown). **b**, Bar plots showing the number of interaction peaks (q -value < 0.05) across distance bins when data was modeled using CHiCANE's default negative binomial distribution (blue), which allows for overdispersion. Interaction peaks identified using additional models based on Poisson distribution (orange) and truncated Poisson distribution (green) fitted using CHiCANE are also shown. Although CHiCANE supports truncated negative binomial distribution, these data were not generated due to memory requirements for this relatively large MK library.

Figure 5: Testing different models (Holgerson et al., 2021)

Conclusions

Comprehensive and quantitative interrogation of cancer immunometabolism requires molecular and cellular tools and sophisticated computational methods to analyze complex and large data sets. Given the maturity and robustness of NGS-based technologies and the availability of the associated computational tools reviewed here, we expect that enormous amounts of data will be generated in the upcoming years. This will pose considerable challenges to make these data available and, second, extract information for immuno-oncology from the data sets. Unfortunately, no centralized database hosts genomic/immunogenomic/immunometabolomic data from published clinical studies with immune checkpoint blockers. In many cases, researchers have to request access from individual laboratories.

Similarly, a centralized database that enables queries across scRNA-seq data sets would be beneficial. Both challenges could be solved, but they require community efforts to address and overcome ethical issues (for example, access to controlled data including human sequence information) and technical issues (for example, the size of scRNA-seq data sets). Thus, existing and future computational tools will be instrumental for interrogating cancer immunometabolism in individual patients and will ultimately enable precision immuno-oncology.

References

- AEVERMANN, B. D., NOVOTNY, M., BAKKEN, T., MILLER, J. A., DIEHL, A. D., OSUMI-SUTHERLAND, D., LASKEN, R. S., LEIN, E. S. & SCHEUERMANN, R. H. 2018. Cell type discovery using single-cell transcriptomics: implications for ontological representation. *Hum Mol Genet*, 27, R40-r47.
- AFIK, S., YATES, K. B., BI, K., DARKO, S., GODEC, J., GERDEMANN, U., SWADLING, L., DOUEK, D. C., KLENERMAN, P., BARNES, E. J., SHARPE, A. H., HAINING, W. N. & YOSEF, N. 2017. Targeted reconstruction of T cell receptor sequence from single cell RNA-seq links CDR3 length to T cell differentiation state. *Nucleic Acids Res*, 45, e148.
- AHL, P. J., HOPKINS, R. A., XIANG, W. W., AU, B., KALIAPERUMAL, N., FAIRHURST, A.-M. & CONNOLLY, J. E. 2020. Met-Flow, a strategy for single-cell metabolic analysis highlights dynamic changes in immune subpopulations. *Communications Biology*, 3, 305.
- AL-KHAMI, A. A., ZHENG, L., DEL VALLE, L., HOSSAIN, F., WYCZECOWSKA, D., ZABALETA, J., SANCHEZ, M. D., DEAN, M. J., RODRIGUEZ, P. C. & OCHOA, A. C. 2017. Exogenous lipid uptake induces metabolic and functional reprogramming of tumor-associated myeloid-derived suppressor cells. *Oncoimmunology*, 6, e1344804.
- ALLARD, B., POMMEY, S., SMYTH, M. J. & STAGG, J. 2013. Targeting CD73 enhances the antitumor activity of anti-PD-1 and anti-CTLA-4 mAbs. *Clin Cancer Res*, 19, 5626-35.
- ALVES, TIAGO C., PONGRATZ, REBECCA L., ZHAO, X., YARBOROUGH, O., SEREDA, S., SHIRIHAI, O., CLINE, GARY W., MASON, G. & KIBBEY, RICHARD G. 2015. Integrated, Step-Wise, Mass-Isotopomeric Flux Analysis of the TCA Cycle. *Cell Metabolism*, 22, 936-947.
- AMIR, E.-A. D., DAVIS, K. L., TADMOR, M. D., SIMONDS, E. F., LEVINE, J. H., BENDALL, S. C., SHENFELD, D. K., KRISHNASWAMY, S., NOLAN, G. P. & PE'ER, D. 2013. viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. *Nature Biotechnology*, 31, 545-552.
- ANDERS, S. & HUBER, W. 2010. Differential expression analysis for sequence count data. *Genome Biol*, 11, R106.
- ANDERS, S., MCCARTHY, D. J., CHEN, Y., OKONIEWSKI, M., SMYTH, G. K., HUBER, W. & ROBINSON, M. D. 2013. Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat Protoc*, 8, 1765-86.
- ANDREATTA, M. & NIELSEN, M. 2016. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics*, 32, 511-7.
- ARAKI, K., TURNER, A. P., SHAFFER, V. O., GANGAPPA, S., KELLER, S. A., BACHMANN, M. F., LARSEN, C. P. & AHMED, R. 2009. mTOR regulates memory CD8 T-cell differentiation. *Nature*, 460, 108-12.

- ARAN, D., HU, Z. & BUTTE, A. J. 2017. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol*, 18, 220.
- ARAN, D., LOONEY, A. P., LIU, L., WU, E., FONG, V., HSU, A., CHAK, S., NAIKAWADI, R. P., WOLTERS, P. J., ABATE, A. R., BUTTE, A. J. & BHATTACHARYA, M. 2019. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol*, 20, 163-172.
- ARISS, M. M., ISLAM, A. B. M. M. K., CRITCHER, M., ZAPPIA, M. P. & FROLOV, M. V. 2018. Single cell RNA-sequencing identifies a metabolic aspect of apoptosis in Rbf mutant. *Nature Communications*, 9, 5024.
- ARTS, R. J., NOVAKOVIC, B., TER HORST, R., CARVALHO, A., BEKKERING, S., LACHMANDAS, E., RODRIGUES, F., SILVESTRE, R., CHENG, S. C., WANG, S. Y., HABIBI, E., GONÇALVES, L. G., MESQUITA, I., CUNHA, C., VAN LAARHOVEN, A., VAN DE VEERDONK, F. L., WILLIAMS, D. L., VAN DER MEER, J. W., LOGIE, C., O'NEILL, L. A., DINARELLO, C. A., RIKSEN, N. P., VAN CREVEL, R., CLISH, C., NOTEBAART, R. A., JOOSTEN, L. A., STUNNENBERG, H. G., XAVIER, R. J. & NETEA, M. G. 2016. Glutaminolysis and Fumarate Accumulation Integrate Immunometabolic and Epigenetic Programs in Trained Immunity. *Cell Metab*, 24, 807-819.
- ASSENOV, Y., MÜLLER, F., LUTSIK, P., WALTER, J., LENGAUER, T. & BOCK, C. 2014. Comprehensive analysis of DNA methylation data with RnBeads. *Nat Methods*, 11, 1138-1140.
- AU - VAN DEN BOSSCHE, J., AU - BAARDMAN, J. & AU - DE WINTHER, M. P. J. 2015. Metabolic Characterization of Polarized M1 and M2 Bone Marrow-derived Macrophages Using Real-time Extracellular Flux Analysis. *JoVE*, e53424.
- AUSLANDER, N., ZHANG, G., LEE, J. S., FREDERICK, D. T., MIAO, B., MOLL, T., TIAN, T., WEI, Z., MADAN, S., SULLIVAN, R. J., BOLAND, G., FLAHERTY, K., HERLYN, M. & RUPPIN, E. 2018. Robust prediction of response to immune checkpoint blockade therapy in metastatic melanoma. *Nat Med*, 24, 1545-1549.
- AZIZI, E., CARR, A. J., PLITAS, G., CORNISH, A. E., KONOPACKI, C., PRABHAKARAN, S., NAINYS, J., WU, K., KISELIOVAS, V., SETTY, M., CHOI, K., FROMME, R. M., DAO, P., MCKENNEY, P. T., WASTI, R. C., KADAVERU, K., MAZUTIS, L., RUDENSKY, A. Y. & PE'ER, D. 2018. Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. *Cell*, 174, 1293-1308.e36.
- BAARDMAN, J., VERBERK, S. G. S., PRANGE, K. H. M., VAN WEEGHEL, M., VAN DER VELDEN, S., RYAN, D. G., WÜST, R. C. I., NEELE, A. E., SPEIJER, D., DENIS, S. W., WITTE, M. E., HOUTKOOPER, R. H., O'NEILL, L. A., KNATKO, E. V., DINKOVA-KOSTOVA, A. T., LUTGENS, E., DE WINTHER, M. P. J. & VAN DEN BOSSCHE, J. 2018. A Defective Pentose Phosphate Pathway Reduces Inflammatory Macrophage Responses during Hypercholesterolemia. *Cell Reports*, 25, 2044-2052.e5.
- BAMBOUSKOVA, M., GORVEL, L., LAMPROPOULOU, V., SERGUSHICHEV, A., LOGINICHEVA, E., JOHNSON, K., KORENFELD, D., MATHYER, M. E., KIM, H., HUANG, L. H., DUNCAN, D., BREGMAN, H., KESKIN, A., SANTEFORD, A., APTE, R. S., SEHGAL, R., JOHNSON, B., AMARASINGHE, G. K., SOARES, M. P., SATOH, T., AKIRA, S., HAI, T., DE GUZMAN STRONG, C., AUCLAIR, K., RODDY, T. P., BILLER, S. A., JOVANOVIĆ, M., KLECHEVSKY, E., STEWART, K. M., RANDOLPH, G. J. & ARTYOMOV, M. N. 2018. Electrophilic properties of itaconate and derivatives regulate the IκBζ-ATF3 inflammatory axis. *Nature*, 556, 501-504.
- BANTUG, G. R., FISCHER, M., GRÄHLERT, J., BALMER, M. L., UNTERSTAB, G., DEVELIOGLU, L., STEINER, R., ZHANG, L., COSTA, A. S. H., GUBSER, P. M., BURGNER, A. V., SAUDER, U., LÖLIGER, J., BELLE, R., DIMELOE, S., LÖTSCHER, J., JAUCH, A., RECHER, M., HÖNGER, G., HALL, M. N., ROMERO, P., FREZZA, C. & HESS, C. 2018. Mitochondria-Endoplasmic Reticulum Contact Sites Function as

- Immunometabolic Hubs that Orchestrate the Rapid Recall Response of Memory CD8(+) T Cells. *Immunity*, 48, 542-555.e6.
- BARATELLI, F., LIN, Y., ZHU, L., YANG, S. C., HEUZÉ-VOURC'H, N., ZENG, G., RECKAMP, K., DOHADWALA, M., SHARMA, S. & DUBINETT, S. M. 2005. Prostaglandin E2 induces FOXP3 gene expression and T regulatory cell function in human CD4+ T cells. *J Immunol*, 175, 1483-90.
- BASSANI-STERBERG, M., CHONG, C., GUILLAUME, P., SOLLEDER, M., PAK, H., GANNON, P. O., KANDALAFT, L. E., COUKOS, G. & GFELLER, D. 2017. Deciphering HLA-I motifs across HLA peptidomes improves neo-antigen predictions and identifies allosteric regulating HLA specificity. *PLoS Comput Biol*, 13, e1005725.
- BEAVIS, P. A., MILENKOVSKI, N., HENDERSON, M. A., JOHN, L. B., ALLARD, B., LOI, S., KERSHAW, M. H., STAGG, J. & DARCY, P. K. 2015. Adenosine Receptor 2A Blockade Increases the Efficacy of Anti-PD-1 through Enhanced Antitumor T-cell Responses. *Cancer Immunol Res*, 3, 506-17.
- BECHT, E., GIRALDO, N. A., LACROIX, L., BUTTARD, B., ELAROUCI, N., PETITPREZ, F., SELVES, J., LAURENT-PUIG, P., SAUTÈS-FRIDMAN, C., FRIDMAN, W. H. & DE REYNIÈS, A. 2016. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol*, 17, 218.
- BECHT, E., MCINNES, L., HEALY, J., DUTERTRE, C.-A., KWOK, I. W. H., NG, L. G., GINHOUX, F. & NEWELL, E. W. 2019. Dimensionality reduction for visualizing single-cell data using UMAP. *Nature Biotechnology*, 37, 38-44.
- BOCHET, L., LEHUÉDÉ, C., DAUVILLIER, S., WANG, Y. Y., DIRAT, B., LAURENT, V., DRAY, C., GUIET, R., MARIDONNEAU-PARINI, I., LE GONIDEC, S., COUDERC, B., ESCOURROU, G., VALET, P. & MULLER, C. 2013. Adipocyte-derived fibroblasts promote tumor progression and contribute to the desmoplastic reaction in breast cancer. *Cancer Res*, 73, 5657-68.
- BOEGEL, S., LÖWER, M., SCHÄFER, M., BUKUR, T., DE GRAAF, J., BOISGUÉRIN, V., TÜRECI, O., DIKEN, M., CASTLE, J. C. & SAHIN, U. 2012. HLA typing from RNA-Seq sequence reads. *Genome Med*, 4, 102.
- BOEHM, K. M., BHINDER, B., RAJA, V. J., DEPHOURE, N. & ELEMENTO, O. 2019. Predicting peptide presentation by major histocompatibility complex class I: an improved machine learning approach to the immunopeptidome. *BMC Bioinformatics*, 20, 7.
- BOLOTIN, D. A., POSLAVSKY, S., DAVYDOV, A. N., FRENKEL, F. E., FANCHI, L., ZOLOTAREVA, O. I., HEMMERS, S., PUTINTSEVA, E. V., OBRAZTSOVA, A. S., SHUGAY, M., ATAULLAKHANOV, R. I., RUDENSKY, A. Y., SCHUMACHER, T. N. & CHUDAKOV, D. M. 2017. Antigen receptor repertoire profiling from RNA-seq data. *Nat Biotechnol*, 35, 908-911.
- BOLOTIN, D. A., POSLAVSKY, S., MITROPHANOV, I., SHUGAY, M., MAMEDOV, I. Z., PUTINTSEVA, E. V. & CHUDAKOV, D. M. 2015. MiXCR: software for comprehensive adaptive immunity profiling. *Nat Methods*, 12, 380-1.
- BONSACK, M., HOPPE, S., WINTER, J., TICHY, D., ZELLER, C., KÜPPER, M. D., SCHITTER, E. C., BLATNIK, R. & RIEMER, A. B. 2019. Performance Evaluation of MHC Class-I Binding Prediction Tools Based on an Experimentally Validated MHC-Peptide Binding Data Set. *Cancer Immunol Res*, 7, 719-736.
- BORDBAR, A., FEIST, A. M., USAITE-BLACK, R., WOODCOCK, J., PALSSON, B. O. & FAMILI, I. 2011. A multi-tissue type genome-scale metabolic network for analysis of whole-body systems physiology. *BMC Systems Biology*, 5, 180.
- BORDBAR, A., MO, M. L., NAKAYASU, E. S., SCHRIMPE-RUTLEDGE, A. C., KIM, Y.-M., METZ, T. O., JONES, M. B., FRANK, B. C., SMITH, R. D., PETERSON, S. N., HYDUKE, D. R., ADKINS, J. N. & PALSSON, B. O.

2012. Model-driven multi-omic data analysis elucidates metabolic immunomodulators of macrophage activation. *Molecular Systems Biology*, 8, 558.
- BÖTTCHER, J. P., BONAVIDA, E., CHAKRAVARTY, P., BLEES, H., CABEZA-CABRERIZO, M., SAMMICHELI, S., ROGERS, N. C., SAHAI, E., ZELENAY, S. & REIS E SOUSA, C. 2018. NK Cells Stimulate Recruitment of cDC1 into the Tumor Microenvironment Promoting Cancer Immune Control. *Cell*, 172, 1022-1037.e14.
- BRAND, A., SINGER, K., GUDRUN, KOLITZUS, M., SCHOENHAMMER, G., THIEL, A., MATOS, C., BRUSS, C., KLOBUCH, S., PETER, K., KASTENBERGER, M., BOGDAN, C., SCHLEICHER, U., MACKENSEN, A., ULLRICH, E., FICHTNER-FEIGL, S., KESSELRING, R., MACK, M., RITTER, U., SCHMID, M., BLANK, C., DETTMER, K., PETER, HOFFMANN, P., WALENTA, S., EDWARD, POUYSSEGUR, J., VILLUNGER, A., STEVEN, A., SELIGER, B., SCHREML, S., HAFERKAMP, S., KOHL, E., KARRER, S., BERNEBURG, M., HERR, W., MUELLER-KLIESER, W., RENNER, K. & KREUTZ, M. 2016. LDHA-Associated Lactic Acid Production Blunts Tumor Immunosurveillance by T and NK Cells. *Cell Metabolism*, 24, 657-671.
- BRIN, E., WU, K., LU, H. T., HE, Y., DAI, Z. & HE, W. 2017. PEGylated arginine deiminase can modulate tumor immune microenvironment by affecting immune checkpoint expression, decreasing regulatory T cell accumulation and inducing tumor T cell infiltration. *Oncotarget*, 8, 58948-58963.
- BRUGGNER, R. V., BODENMILLER, B., DILL, D. L., TIBSHIRANI, R. J. & NOLAN, G. P. 2014. Automated identification of stratifying signatures in cellular subpopulations. *Proc Natl Acad Sci U S A*, 111, E2770-7.
- BUCHANAN, F. G., GORDEN, D. L., MATTA, P., SHI, Q., MATRISIAN, L. M. & DUBOIS, R. N. 2006. Role of beta-arrestin 1 in the metastatic progression of colorectal cancer. *Proc Natl Acad Sci U S A*, 103, 1492-7.
- BUCHKOVICH, M. L., BROWN, C. C., ROBASKY, K., CHAI, S., WESTFALL, S., VINCENT, B. G., WEIMER, E. T. & POWERS, J. G. 2017. HLAProfiler utilizes k-mer profiles to improve HLA calling accuracy for rare and common alleles in RNA-seq data. *Genome Med*, 9, 86.
- BUCK, M. D., O'SULLIVAN, D., KLEIN GELTINK, R. I., CURTIS, J. D., CHANG, C. H., SANIN, D. E., QIU, J., KRETZ, O., BRAAS, D., VAN DER WINDT, G. J., CHEN, Q., HUANG, S. C., O'NEILL, C. M., EDELSON, B. T., PEARCE, E. J., SESAKI, H., HUBER, T. B., RAMBOLD, A. S. & PEARCE, E. L. 2016. Mitochondrial Dynamics Controls T Cell Fate through Metabolic Programming. *Cell*, 166, 63-76.
- BUFFEN, K., OOSTING, M., QUINTIN, J., NG, A., KLEINNIJENHUIS, J., KUMAR, V., VAN DE VOSSE, E., WIJMEGA, C., VAN CREVEL, R., OOSTERWIJK, E., GROTENHUIS, A. J., VERMEULEN, S. H., KIEMENEY, L. A., VAN DE VEERDONK, F. L., CHAMILOS, G., XAVIER, R. J., VAN DER MEER, J. W., NETEA, M. G. & JOOSTEN, L. A. 2014. Autophagy controls BCG-induced trained immunity and the response to intravesical BCG therapy for bladder cancer. *PLoS Pathog*, 10, e1004485.
- BULIK-SULLIVAN, B., BUSBY, J., PALMER, C. D., DAVIS, M. J., MURPHY, T., CLARK, A., BUSBY, M., DUKE, F., YANG, A., YOUNG, L., OJO, N. C., CALDWELL, K., ABHYANKAR, J., BOUCHER, T., HART, M. G., MAKAROV, V., DE MONTPREVILLE, V. T., MERCIER, O., CHAN, T. A., SCAGLIOTTI, G., BIRONZO, P., NOVELLO, S., KARACHALIOU, N., ROSELL, R., ANDERSON, I., GABRAIL, N., HROM, J., LIMVARAPUSS, C., CHOQUETTE, K., SPIRA, A., ROUSSEAU, R., VOONG, C., RIZVI, N. A., FADEL, E., FRATTINI, M., JOOSS, K., SKOBERNE, M., FRANCIS, J. & YELENSKY, R. 2019. Deep learning using tumor HLA peptide mass spectrometry datasets improves neoantigen identification. *Nature Biotechnology*, 37, 55-63.
- BULLARD, J. H., PURDOM, E., HANSEN, K. D. & DUDOIT, S. 2010. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics*, 11, 94.

- CANZAR, S., NEU, K. E., TANG, Q., WILSON, P. C. & KHAN, A. A. 2017. BASIC: BCR assembly from single cells. *Bioinformatics*, 33, 425-427.
- CARBONNEAU, M., L. M. G., LALONDE, M. E., GERMAIN, M. A., MOTORINA, A., GUIOT, M. C., SECCO, B., VINCENT, E. E., TUMBER, A., HULEA, L., BERGEMAN, J., OPPERMANN, U., JONES, R. G., LAPLANTE, M., TOPISIROVIC, I., PETRECCA, K., HUOT, M. & MALLETT, F. A. 2016. The oncometabolite 2-hydroxyglutarate activates the mTOR signalling pathway. *Nat Commun*, 7, 12700.
- CARO-MALDONADO, A., WANG, R., NICHOLS, A. G., KURAOKA, M., MILASTA, S., SUN, L. D., GAVIN, A. L., ABEL, E. D., KELSOE, G., GREEN, D. R. & RATHMELL, J. C. 2014. Metabolic Reprogramming Is Required for Antibody Production That Is Suppressed in Anergic but Exaggerated in Chronically BAFF-Exposed B Cells. *The Journal of Immunology*, 192, 3626-3636.
- CARPENTER, A. E., JONES, T. R., LAMPRECHT, M. R., CLARKE, C., KANG, I. H., FRIMAN, O., GUERTIN, D. A., CHANG, J. H., LINDQUIST, R. A., MOFFAT, J., GOLLAND, P. & SABATINI, D. M. 2006. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol*, 7, R100.
- CARR, E. L., KELMAN, A., WU, G. S., GOPAUL, R., SENKEVITCH, E., AGHVANYAN, A., TURAY, A. M. & FRAUWIRTH, K. A. 2010. Glutamine Uptake and Metabolism Are Coordinately Regulated by ERK/MAPK during T Lymphocyte Activation. *The Journal of Immunology*, 185, 1037-1044.
- CASCONE, T., MCKENZIE, J. A., MBOFUNG, R. M., PUNT, S., WANG, Z., XU, C., WILLIAMS, L. J., WANG, Z., BRISTOW, C. A., CARUGO, A., PEOPLES, M. D., LI, L., KARPINETS, T., HUANG, L., MALU, S., CREASY, C., LEAHEY, S. E., CHEN, J., CHEN, Y., PELICANO, H., BERNATCHEZ, C., GOPAL, Y. N. V., HEFFERNAN, T. P., HU, J., WANG, J., AMARIA, R. N., GARRAWAY, L. A., HUANG, P., YANG, P., WISTUBA, I. I., WOODMAN, S. E., ROSZIK, J., DAVIS, R. E., DAVIES, M. A., HEYMACH, J. V., HWU, P. & PENG, W. 2018. Increased Tumor Glycolysis Characterizes Immune Resistance to Adoptive T Cell Therapy. *Cell Metabolism*, 27, 977-987.e4.
- CERAMI, E., GAO, J., DOGRUSOZ, U., GROSS, B. E., SUMER, S. O., AKSOY, B. A., JACOBSEN, A., BYRNE, C. J., HEUER, M. L., LARSSON, E., ANTIPIN, Y., REVA, B., GOLDBERG, A. P., SANDER, C. & SCHULTZ, N. 2012. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*, 2, 401-4.
- CHANG, C.-H., CURTIS, JONATHAN D., MAGGI, LEONARD B., FAUBERT, B., VILLARINO, ALEJANDRO V., O'SULLIVAN, D., HUANG, STANLEY C.-C., VAN DER WINDT, GERRITJE J. W., BLAGIH, J., QIU, J., WEBER, JASON D., PEARCE, EDWARD J., JONES, RUSSELL G. & PEARCE, ERIKA L. 2013. Posttranscriptional Control of T Cell Effector Function by Aerobic Glycolysis. *Cell*, 153, 1239-1251.
- CHANG, C.-H., QIU, J., O'SULLIVAN, D., MICHAEL, NOGUCHI, T., JONATHAN, CHEN, Q., GINDIN, M., MATTHEW, GERRITJE, TONC, E., ROBERT, EDWARD & ERIKA 2015. Metabolic Competition in the Tumor Microenvironment Is a Driver of Cancer Progression. *Cell*, 162, 1229-1241.
- CHELL, S. D., WITHERDEN, I. R., DOBSON, R. R., MOORGHEN, M., HERMAN, A. A., QUALTROUGH, D., WILLIAMS, A. C. & PARASKEVA, C. 2006. Increased EP4 receptor expression in colorectal cancer progression promotes cell growth and anchorage independence. *Cancer Res*, 66, 3106-13.
- CHEN, J., BARDES, E. E., ARONOW, B. J. & JEGGA, A. G. 2009. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res*, 37, W305-11.
- CHEN, J. H., PERRY, C. J., TSUI, Y. C., STARON, M. M., PARISH, I. A., DOMINGUEZ, C. X., ROSENBERG, D. W. & KAECH, S. M. 2015. Prostaglandin E2 and programmed cell death 1 signaling coordinately impair CTL function and survival during chronic viral infection. *Nat Med*, 21, 327-34.
- CHENG, S. C., QUINTIN, J., CRAMER, R. A., SHEPARDSON, K. M., SAEED, S., KUMAR, V., GIAMARELLOS-BOURBOULIS, E. J., MARTENS, J. H., RAO, N. A., AGHAJANIREFAH, A., MANJERI, G. R., LI, Y., IFRIM,

- D. C., ARTS, R. J., VAN DER VEER, B. M., DEEN, P. M., LOGIE, C., O'NEILL, L. A., WILLEMS, P., VAN DE VEERDONK, F. L., VAN DER MEER, J. W., NG, A., JOOSTEN, L. A., WIJMENGA, C., STUNNENBERG, H. G., XAVIER, R. J. & NETEA, M. G. 2014. mTOR- and HIF-1 α -mediated aerobic glycolysis as metabolic basis for trained immunity. *Science*, 345, 1250684.
- CHESNEY, J. A., TELANG, S., YADDANAPUDI, K. & GREWAL, J. S. 2016. Targeting 6-phosphofructo-2-kinase (PFKFB3) as an immunotherapeutic strategy. *Journal of Clinical Oncology*, 34, e14548-e14548.
- CHOI, B. K., LEE, D. Y., LEE, D. G., KIM, Y. H., KIM, S. H., OH, H. S., HAN, C. & KWON, B. S. 2017. 4-1BB signaling activates glucose and fatty acid metabolism to enhance CD8(+) T cell proliferation. *Cell Mol Immunol*, 14, 748-757.
- CHOWDHURY, P. S., CHAMOTO, K., KUMAR, A. & HONJO, T. 2018. PPAR-Induced Fatty Acid Oxidation in T Cells Increases the Number of Tumor-Reactive CD8(+) T Cells and Facilitates Anti-PD-1 Therapy. *Cancer Immunol Res*, 6, 1375-1387.
- CHRISTOFK, H. R., VANDER HEIDEN, M. G., HARRIS, M. H., RAMANATHAN, A., GERSZTEN, R. E., WEI, R., FLEMING, M. D., SCHREIBER, S. L. & CANTLEY, L. C. 2008. The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. *Nature*, 452, 230-233.
- COLEGIO, O. R., CHU, N.-Q., SZABO, A. L., CHU, T., RHEBERGEN, A. M., JAIRAM, V., CYRUS, N., BROKOWSKI, C. E., EISENBARTH, S. C., PHILLIPS, G. M., CLINE, G. W., PHILLIPS, A. J. & MEDZHITOV, R. 2014. Functional polarization of tumour-associated macrophages by tumour-derived lactic acid. *Nature*, 513, 559-563.
- CROMPTON, J. G., SUKUMAR, M., ROYCHOUDHURI, R., CLEVER, D., GROS, A., EIL, R. L., TRAN, E., HANADA, K., YU, Z., PALMER, D. C., KERKAR, S. P., MICHALEK, R. D., UPHAM, T., LEONARDI, A., ACQUAVELLA, N., WANG, E., MARINCOLA, F. M., GATTINONI, L., MURANSKI, P., SUNDRUD, M. S., KLEBANOFF, C. A., ROSENBERG, S. A., FEARON, D. T. & RESTIFO, N. P. 2015. Akt inhibition enhances expansion of potent tumor-specific lymphocytes with memory cell characteristics. *Cancer Res*, 75, 296-305.
- CRONIN, S. J. F., SEEHUS, C., WEIDINGER, A., TALBOT, S., REISSIG, S., SEIFERT, M., PIERSON, Y., MCNEILL, E., LONGHI, M. S., TURNES, B. L., KRESLAVSKY, T., KOGLER, M., HOFFMANN, D., TICEVIC, M., DA LUZ SCHEFFER, D., TORTOLA, L., CIKES, D., JAIS, A., RANGACHARI, M., RAO, S., PAOLINO, M., NOVATCHKOVA, M., AICHINGER, M., BARRETT, L., LATREMOLIERE, A., WIRNSBERGER, G., LAMETSCHWANDTNER, G., BUSSLINGER, M., ZICHA, S., LATINI, A., ROBSON, S. C., WAISMAN, A., ANDREWS, N., COSTIGAN, M., CHANNON, K. M., WEISS, G., KOZLOV, A. V., TEBBE, M., JOHNSON, K., WOOLF, C. J. & PENNINGER, J. M. 2018. The metabolite BH4 controls T cell proliferation in autoimmunity and cancer. *Nature*, 563, 564-568.
- CSÓKA, B., SELMECZY, Z., KOSCSÓ, B., NÉMETH, Z. H., PACHER, P., MURRAY, P. J., KEPKA-LENHART, D., MORRIS, S. M., JR., GAUSE, W. C., LEBOVICH, S. J. & HASKÓ, G. 2012. Adenosine promotes alternative macrophage activation via A2A and A2B receptors. *Faseb j*, 26, 376-86.
- CUBILLOS-RUIZ, J. R., SILBERMAN, P. C., RUTKOWSKI, M. R., CHOPRA, S., PERALES-PUCHALT, A., SONG, M., ZHANG, S., BETTIGOLE, S. E., GUPTA, D., HOLCOMB, K., ELLENSON, L. H., CAPUTO, T., LEE, A. H., CONEJO-GARCIA, J. R. & GLIMCHER, L. H. 2015. ER Stress Sensor XBP1 Controls Anti-tumor Immunity by Disrupting Dendritic Cell Homeostasis. *Cell*, 161, 1527-38.
- DAO, D., FRASER, A. N., HUNG, J., LJOSA, V., SINGH, S. & CARPENTER, A. E. 2016. CellProfiler Analyst: interactive data exploration, analysis and classification of large biological image sets. *Bioinformatics*, 32, 3210-3212.
- DE ROSA, V., GALGANI, M., PORCELLINI, A., COLAMATTEO, A., SANTOPAULO, M., ZUCHEGNA, C., ROMANO, A., DE SIMONE, S., PROCACCINI, C., LA ROCCA, C., CARRIERI, P. B., MANISCALCO, G. T., SALVETTI,

- M., BUSCARINU, M. C., FRANZESE, A., MOZZILLO, E., LA CAVA, A. & MATARESE, G. 2015. Glycolysis controls the induction of human regulatory T cells by modulating the expression of FOXP3 exon 2 splicing variants. *Nature Immunology*, 16, 1174-1184.
- DEMEURE, C. E., YANG, L. P., DESJARDINS, C., RAYNAULD, P. & DELESPESE, G. 1997. Prostaglandin E2 primes naive T cells for the production of anti-inflammatory cytokines. *Eur J Immunol*, 27, 3526-31.
- DENNIS, E. A., DEEMS, R. A., HARKEWICZ, R., QUEHENBERGER, O., BROWN, H. A., MILNE, S. B., MYERS, D. S., GLASS, C. K., HARDIMAN, G., REICHAUT, D., MERRILL, A. H., SULLARDS, M. C., WANG, E., MURPHY, R. C., RAETZ, C. R. H., GARRETT, T. A., GUAN, Z., RYAN, A. C., RUSSELL, D. W., MCDONALD, J. G., THOMPSON, B. M., SHAW, W. A., SUD, M., ZHAO, Y., GUPTA, S., MAURYA, M. R., FAHY, E. & SUBRAMANIAM, S. 2010. A Mouse Macrophage Lipidome*. *Journal of Biological Chemistry*, 285, 39976-39985.
- DIETL, K., RENNER, K., DETTMER, K., TIMISCHL, B., EBERHART, K., DORN, C., HELLERBRAND, C., KASTENBERGER, M., KUNZ-SCHUGHART, L. A., OEFNER, P. J., ANDRESEN, R., GOTTFRIED, E. & KREUTZ, M. P. 2010. Lactic Acid and Acidification Inhibit TNF Secretion and Glycolysis of Human Monocytes. *The Journal of Immunology*, 184, 1200-1209.
- DILTNEY, A. T., MENTZER, A. J., CARAPITO, R., CUTLAND, C., CEREB, N., MADHI, S. A., RHIE, A., KOREN, S., BAHRAM, S., MCVEAN, G. & PHILLIPPY, A. M. 2019. HLA*LA—HLA typing from linearly projected graph alignments. *Bioinformatics*, 35, 4394-4396.
- DING, J., ADICONIS, X., SIMMONS, S. K., KOWALCZYK, M. S., HESSION, C. C., MARJANOVIC, N. D., HUGHES, T. K., WADSWORTH, M. H., BURKS, T., NGUYEN, L. T., KWON, J. Y. H., BARAK, B., GE, W., KEDAIGLE, A. J., CARROLL, S., LI, S., HACOEN, N., ROZENBLATT-ROSEN, O., SHALEK, A. K., VILLANI, A.-C., REGEV, A. & LEVIN, J. Z. 2020. Systematic comparison of single-cell and single-nucleus RNA-sequencing methods. *Nature Biotechnology*, 38, 737-746.
- DING, J., CONDON, A. & SHAH, S. P. 2018. Interpretable dimensionality reduction of single cell transcriptome data with deep generative models. *Nat Commun*, 9, 2002.
- DIVAKARUNI, A. S., HSIEH, W. Y., MINARRIETA, L., DUONG, T. N., KIM, K. K. O., DESOUSA, B. R., ANDREYEV, A. Y., BOWMAN, C. E., CARADONNA, K., DRANKA, B. P., FERRICK, D. A., LIESA, M., STILES, L., ROGERS, G. W., BRAAS, D., CIARALDI, T. P., WOLFGANG, M. J., SPARWASSER, T., BEROD, L., BENSINGER, S. J. & MURPHY, A. N. 2018. Etomoxir Inhibits Macrophage Polarization by Disrupting CoA Homeostasis. *Cell Metabolism*, 28, 490-503.e7.
- EIL, R., VODNALA, S. K., CLEVER, D., KLEBANOFF, C. A., SUKUMAR, M., PAN, J. H., PALMER, D. C., GROS, A., YAMAMOTO, T. N., PATEL, S. J., GUITTARD, G. C., YU, Z., CARONARO, V., OKKENHAUG, K., SCHRUMP, D. S., LINEHAN, W. M., ROYCHOUDHURI, R. & RESTIFO, N. P. 2016. Ionic immune suppression within the tumour microenvironment limits T cell effector function. *Nature*, 537, 539-543.
- ELF, S. E. & CHEN, J. 2014. Targeting glucose metabolism in patients with cancer. *Cancer*, 120, 774-780.
- ELTAHLA, A. A., RIZZETTO, S., PIROZYAN, M. R., BETZ-STABLEIN, B. D., VENTURI, V., KEDZIERKA, K., LLOYD, A. R., BULL, R. A. & LUCIANI, F. 2016. Linking the T cell receptor to the single cell transcriptome in antigen-specific human T cells. *Immunol Cell Biol*, 94, 604-11.
- EVERTS, B., AMIEL, E., HUANG, S. C.-C., SMITH, A. M., CHANG, C.-H., LAM, W. Y., REDMANN, V., FREITAS, T. C., BLAGIH, J., VAN DER WINDT, G. J. W., ARTYOMOV, M. N., JONES, R. G., PEARCE, E. L. & PEARCE, E. J. 2014. TLR-driven early glycolytic reprogramming via the kinases TBK1-IKKe supports the anabolic demands of dendritic cell activation. *Nature Immunology*, 15, 323-332.

- FABREGAT, A., JUPE, S., MATTHEWS, L., SIDIROPOULOS, K., GILLESPIE, M., GARAPATI, P., HAW, R., JASSAL, B., KORNINGER, F., MAY, B., MILACIC, M., ROCA, C. D., ROTHFELS, K., SEVILLA, C., SHAMOVSKY, V., SHORSER, S., VARUSAI, T., VITERI, G., WEISER, J., WU, G., STEIN, L., HERMJAKOB, H. & D'EUSTACHIO, P. 2018. The Reactome Pathway Knowledgebase. *Nucleic Acids Res*, 46, D649-d655.
- FAUBERT, B., LI, K. Y., CAI, L., HENSLEY, C. T., KIM, J., ZACHARIAS, L. G., YANG, C., DO, Q. N., DOUCETTE, S., BURGUETE, D., LI, H., HUET, G., YUAN, Q., WIGAL, T., BUTT, Y., NI, M., TORREALBA, J., OLIVER, D., LENKINSKI, R. E., MALLOY, C. R., WACHSMANN, J. W., YOUNG, J. D., KERNSTINE, K. & DEBERARDINIS, R. J. 2017. Lactate Metabolism in Human Lung Tumors. *Cell*, 171, 358-371.e9.
- FERRER-FONT, L., PELLEFIGUES, C., MAYER, J. U., SMALL, S. J., JAIMES, M. C. & PRICE, K. M. 2020. Panel Design and Optimization for High-Dimensional Immunophenotyping Assays Using Spectral Flow Cytometry. *Curr Protoc Cytom*, 92, e70.
- FINOTELLO, F., MAYER, C., PLATTNER, C., LASCHOB, G., RIEDER, D., HACKL, H., KROGSDAM, A., LONCOVA, Z., POSCH, W., WILFLINGSIEDER, D., SOPPER, S., IJSSELSTEIJN, M., BROUWER, T. P., JOHNSON, D., XU, Y., WANG, Y., SANDERS, M. E., ESTRADA, M. V., ERICSSON-GONZALEZ, P., CHAROENTONG, P., BALKO, J., DE MIRANDA, N. & TRAJANOSKI, Z. 2019. Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. *Genome Med*, 11, 34.
- FRAUWIRTH, K. A., RILEY, J. L., HARRIS, M. H., PARRY, R. V., RATHMELL, J. C., PLAS, D. R., ELSTROM, R. L., JUNE, C. H. & THOMPSON, C. B. 2002. The CD28 signaling pathway regulates glucose metabolism. *Immunity*, 16, 769-77.
- FREEMERMAN, A. J., JOHNSON, A. R., SACKS, G. N., MILNER, J. J., KIRK, E. L., TROESTER, M. A., MACINTYRE, A. N., GORAKSHA-HICKS, P., RATHMELL, J. C. & MAKOWSKI, L. 2014. Metabolic reprogramming of macrophages: glucose transporter 1 (GLUT1)-mediated glucose metabolism drives a proinflammatory phenotype. *J Biol Chem*, 289, 7884-96.
- FU, X., LI, J., TIAN, Y., QUAN, W., ZHANG, S., LIU, Q., LIANG, F., ZHU, X., ZHANG, L., WANG, D. & HU, J. 2017. Long-read sequence assembly of the firefly *Pyrocoelia pectoralis* genome. *Gigascience*, 6, 1-7.
- GALVÁN-PEÑA, S., CARROLL, R. G., NEWMAN, C., HINCHY, E. C., PALSSON-MCDERMOTT, E., ROBINSON, E. K., COVARRUBIAS, S., NADIN, A., JAMES, A. M., HANEKLAUS, M., CARPENTER, S., KELLY, V. P., MURPHY, M. P., MODIS, L. K. & O'NEILL, L. A. 2019. Malonylation of GAPDH is an inflammatory signal in macrophages. *Nature Communications*, 10, 338.
- GARDEUX, V., DAVID, F. P. A., SHAJKOFCI, A., SCHWALIE, P. C. & DEPLANCKE, B. 2017. ASAP: a web-based platform for the analysis and interactive visualization of single-cell RNA-seq data. *Bioinformatics*, 33, 3123-3125.
- GEIGER, R., RIECKMANN, J. C., WOLF, T., BASSO, C., FENG, Y., FUHRER, T., KOGADEEVA, M., PICOTTI, P., MEISSNER, F., MANN, M., ZAMBONI, N., SALLUSTO, F. & LANZAVECCHIA, A. 2016. L-Arginine Modulates T Cell Metabolism and Enhances Survival and Anti-tumor Activity. *Cell*, 167, 829-842.e13.
- GFELLER, D., GUILLAUME, P., MICHAUX, J., PAK, H. S., DANIEL, R. T., RACLE, J., COUKOS, G. & BASSANI-STERNBERG, M. 2018. The Length Distribution and Multiple Specificity of Naturally Presented HLA-I Ligands. *J Immunol*, 201, 3705-3716.
- GIESEN, C., WANG, H. A., SCHAPIRO, D., ZIVANOVIC, N., JACOBS, A., HATTENDORF, B., SCHÜFFLER, P. J., GROLMUND, D., BUHMANN, J. M., BRANDT, S., VARGA, Z., WILD, P. J., GÜNTHER, D. & BODENMILLER, B. 2014. Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat Methods*, 11, 417-22.

- GIGOUX, M., LOVATO, A., LECONTE, J., LEUNG, J., SONENBERG, N. & SUH, W. K. 2014. Inducible costimulator facilitates T-dependent B cell activation by augmenting IL-4 translation. *Mol Immunol*, 59, 46-54.
- GONZALEZ-PEREZ, A., MUSTONEN, V., REVA, B., RITCHIE, G. R., CREIXELL, P., KARCHIN, R., VAZQUEZ, M., FINK, J. L., KASSAHN, K. S., PEARSON, J. V., BADER, G. D., BOUTROS, P. C., MUTHUSWAMY, L., OUELLETTE, B. F., REIMAND, J., LINDING, R., SHIBATA, T., VALENCIA, A., BUTLER, A., DRONOV, S., FLICEK, P., SHANNON, N. B., CARTER, H., DING, L., SANDER, C., STUART, J. M., STEIN, L. D. & LOPEZ-BIGAS, N. 2013. Computational approaches to identify functional genetic variants in cancer genomes. *Nat Methods*, 10, 723-9.
- GOTTFRIED, E., KUNZ-SCHUGHART, L. A., EBNER, S., MUELLER-KLIESER, W., HOVES, S., ANDREESSEN, R., MACKENSEN, A. & KREUTZ, M. 2006. Tumor-derived lactic acid modulates dendritic cell activation and antigen expression. *Blood*, 107, 2013-2021.
- GROSS, M. I., DEMO, S. D., DENNISON, J. B., CHEN, L., CHERNOV-ROGAN, T., GOYAL, B., JANES, J. R., LAIDIG, G. J., LEWIS, E. R., LI, J., MACKINNON, A. L., PARLATI, F., RODRIGUEZ, M. L. M., SHWONEK, P. J., SJOGREN, E. B., STANTON, T. F., WANG, T., YANG, J., ZHAO, F. & BENNETT, M. K. 2014. Antitumor Activity of the Glutaminase Inhibitor CB-839 in Triple-Negative Breast Cancer. *Molecular Cancer Therapeutics*, 13, 890-901.
- GUBIN, M. M., ESAULOVA, E., WARD, J. P., MALKOVA, O. N., RUNCIL, D., WONG, P., NOGUCHI, T., ARTHUR, C. D., MENG, W., ALSPACH, E., MEDRANO, R. F. V., FRONICK, C., FEHLINGS, M., NEWELL, E. W., FULTON, R. S., SHEEHAN, K. C. F., OH, S. T., SCHREIBER, R. D. & ARTYOMOV, M. N. 2018. High-Dimensional Analysis Delineates Myeloid and Lymphoid Compartment Remodeling during Successful Immune-Checkpoint Cancer Therapy. *Cell*, 175, 1014-1030.e19.
- GUBSER, P. M., BANTUG, G. R., RAZIK, L., FISCHER, M., DIMELOE, S., HOENGER, G., DUROVIC, B., JAUCH, A. & HESS, C. 2013. Rapid effector function of memory CD8⁺ T cells requires an immediate-early glycolytic switch. *Nature Immunology*, 14, 1064-1072.
- GUO, M., WANG, H., POTTER, S. S., WHITSETT, J. A. & XU, Y. 2015. SINCERA: A Pipeline for Single-Cell RNA-Seq Profiling Analysis. *PLoS Comput Biol*, 11, e1004575.
- HAGHIKIA, A., JÖRG, S., DUSCHA, A., BERG, J., MANZEL, A., WASCHBISCH, A., HAMMER, A., LEE, D. H., MAY, C., WILCK, N., BALOGH, A., OSTERMANN, A. I., SCHEBB, N. H., AKKAD, D. A., GROHME, D. A., KLEINWETTFELD, M., KEMPA, S., THÖNE, J., DEMIR, S., MÜLLER, D. N., GOLD, R. & LINKER, R. A. 2015. Dietary Fatty Acids Directly Impact Central Nervous System Autoimmunity via the Small Intestine. *Immunity*, 43, 817-29.
- HAN, Y. & KIM, D. 2017. Deep convolutional neural networks for pan-specific peptide-MHC class I binding prediction. *BMC Bioinformatics*, 18, 585.
- HARTMANN, F. J., MRDJEN, D., MCCAFFREY, E., GLASS, D. R., GREENWALD, N. F., BHARADWAJ, A., KHAIR, Z., VERBERK, S. G. S., BARANSKI, A., BASKAR, R., GRAF, W., VAN VALEN, D., VAN DEN BOSSCHE, J., ANGELO, M. & BENDALL, S. C. 2021. Single-cell metabolic profiling of human cytotoxic T cells. *Nature Biotechnology*, 39, 186-197.
- HASCHEMI, A., KOSMA, P., GILLE, L., EVANS, CHARLES R., BURANT, CHARLES F., STARKL, P., KNAPP, B., HAAS, R., SCHMID, JOHANNES A., JANDL, C., AMIR, S., LUBEC, G., PARK, J., ESTERBAUER, H., BILBAN, M., BRIZUELA, L., POSPISILIK, J. A., OTTERBEIN, LEO E. & WAGNER, O. 2012. The Sedoheptulose Kinase CARKL Directs Macrophage Polarization through Control of Glucose Metabolism. *Cell Metabolism*, 15, 813-826.

- HAYASHI, K., JUTABHA, P., ENDOU, H., SAGARA, H. & ANZAI, N. 2013. LAT1 Is a Critical Transporter of Essential Amino Acids for Immune Reactions in Activated Human T Cells. *The Journal of Immunology*, 191, 4080-4085.
- HE, X., LIN, H., YUAN, L. & LI, B. 2017. Combination therapy with L-arginine and α -PD-L1 antibody boosts immune response against osteosarcoma in immunocompetent mice. *Cancer Biol Ther*, 18, 94-100.
- HEATHER, J. M., ISMAIL, M., OAKES, T. & CHAIN, B. 2018. High-throughput sequencing of the T-cell receptor repertoire: pitfalls and opportunities. *Brief Bioinform*, 19, 554-565.
- HERBER, D. L., CAO, W., NEFEDOVA, Y., NOVITSKIY, S. V., NAGARAJ, S., TYURIN, V. A., CORZO, A., CHO, H. I., CELIS, E., LENNOX, B., KNIGHT, S. C., PADHYA, T., MCCAFFREY, T. V., MCCAFFREY, J. C., ANTONIA, S., FISHMAN, M., FERRIS, R. L., KAGAN, V. E. & GABRILOVICH, D. I. 2010. Lipid accumulation and dendritic cell dysfunction in cancer. *Nat Med*, 16, 880-6.
- HERMAN, J. S., SAGAR & GRÜN, D. 2018. FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. *Nat Methods*, 15, 379-386.
- HO, P.-C., JESSICA, ANDREW, STARON, M., LIU, X., AMEZQUITA, R., TSUI, Y.-C., CUI, G., MICEVIC, G., JOSE, STEVEN, E, KARL, FESKE, S., JASON, MARCUS, JEFFREY & SUSAN 2015a. Phosphoenolpyruvate Is a Metabolic Checkpoint of Anti-tumor T Cell Responses. *Cell*, 162, 1217-1228.
- HO, P. C., BIHUNIAK, J. D., MACINTYRE, A. N., STARON, M., LIU, X., AMEZQUITA, R., TSUI, Y. C., CUI, G., MICEVIC, G., PERALES, J. C., KLEINSTEIN, S. H., ABEL, E. D., INSOGNA, K. L., FESKE, S., LOCASALE, J. W., BOSENBERG, M. W., RATHMELL, J. C. & KAECH, S. M. 2015b. Phosphoenolpyruvate Is a Metabolic Checkpoint of Anti-tumor T Cell Responses. *Cell*, 162, 1217-28.
- HSIEH, W.-Y., ZHOU, Q. D., YORK, A. G., WILLIAMS, K. J., SCUMPIA, P. O., KRONENBERGER, E. B., HOI, X. P., SU, B., CHI, X., BUI, V. L., KHIALEEVA, E., KAPLAN, A., SON, Y. M., DIVAKARUNI, A. S., SUN, J., SMALE, S. T., FLAVELL, R. A. & BENSINGER, S. J. 2020. Toll-Like Receptors Induce Signal-Specific Reprogramming of the Macrophage Lipidome. *Cell Metabolism*, 32, 128-143.e5.
- HU, X., ZHANG, J., WANG, J., FU, J., LI, T., ZHENG, X., WANG, B., GU, S., JIANG, P., FAN, J., YING, X., ZHANG, J., CARROLL, M. C., WUCHERPFENNIG, K. W., HACHEN, N., ZHANG, F., ZHANG, P., LIU, J. S., LI, B. & LIU, X. S. 2019. Landscape of B cell immunity and related immune evasion in human cancers. *Nat Genet*, 51, 560-567.
- HUANG, D. W., SHERMAN, B. T. & LEMPICKI, R. A. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4, 44-57.
- HUANG, STANLEY C.-C., SMITH, AMBER M., EVERTS, B., COLONNA, M., PEARCE, ERIKA L., SCHILLING, JOEL D. & PEARCE, EDWARD J. 2016. Metabolic Reprogramming Mediated by the mTORC2-IRF4 Signaling Axis Is Essential for Macrophage Alternative Activation. *Immunity*, 45, 817-830.
- HUANG, S. C., EVERTS, B., IVANOVA, Y., O'SULLIVAN, D., NASCIMENTO, M., SMITH, A. M., BEATTY, W., LOVE-GREGORY, L., LAM, W. Y., O'NEILL, C. M., YAN, C., DU, H., ABUMRAD, N. A., URBAN, J. F., JR., ARTYOMOV, M. N., PEARCE, E. L. & PEARCE, E. J. 2014. Cell-intrinsic lysosomal lipolysis is essential for alternative activation of macrophages. *Nat Immunol*, 15, 846-55.
- HUANG, T. T., YEN, M. C., LIN, C. C., WENG, T. Y., CHEN, Y. L., LIN, C. M. & LAI, M. D. 2011. Skin delivery of short hairpin RNA of indoleamine 2,3 dioxygenase induces antitumor immunity against orthotopic and metastatic liver cancer. *Cancer Sci*, 102, 2214-20.
- HUI, S., GHERGUROVICH, J. M., MORSCHER, R. J., JANG, C., TENG, X., LU, W., ESPARZA, L. A., REYA, T., LE, Z., YANXIANG GUO, J., WHITE, E. & RABINOWITZ, J. D. 2017. Glucose feeds the TCA cycle via circulating lactate. *Nature*, 551, 115-118.

- HUSAIN, Z., HUANG, Y., SETH, P. & SUKHATME, V. P. 2013. Tumor-Derived Lactate Modifies Antitumor Immune Response: Effect on Myeloid-Derived Suppressor Cells and NK Cells. *The Journal of Immunology*, 191, 1486-1495.
- IANNONE, R., MIELE, L., MAIOLINO, P., PINTO, A. & MORELLO, S. 2014. Adenosine limits the therapeutic effectiveness of anti-CTLA4 mAb in a mouse melanoma model. *Am J Cancer Res*, 4, 172-81.
- INOUE, Y., YOSHIMURA, K., KURABE, N., KAHYO, T., KAWASE, A., TANAHASHI, M., OGAWA, H., INUI, N., FUNAI, K., SHINMURA, K., NIWA, H., SUDA, T. & SUGIMURA, H. 2017. Prognostic impact of CD73 and A2A adenosine receptor expression in non-small-cell lung cancer. *Oncotarget*, 8, 8738-8751.
- JACOBS, S. R., HERMAN, C. E., MACIVER, N. J., WOFFORD, J. A., WIEMAN, H. L., HAMMEN, J. J. & RATHMELL, J. C. 2008. Glucose uptake is limiting in T cell activation and requires CD28-mediated Akt-dependent and independent pathways. *J Immunol*, 180, 4476-86.
- JAN, BAARDMAN, J., NATASJA, SASKIA, ANNETTE, SUSAN, LUQUE-MARTIN, R., CHEN, H.-J., MARIEKE, AHMED, M., MARTEN, ALEX & MENNO 2016. Mitochondrial Dysfunction Prevents Repolarization of Inflammatory Macrophages. *Cell Reports*, 17, 684-696.
- JHA, ABHISHEK K., HUANG, STANLEY C.-C., SERGUSHICHEV, A., LAMPROPOULOU, V., IVANOVA, Y., LOGINICHEVA, E., CHMIELEWSKI, K., STEWART, KELLY M., ASHALL, J., EVERTS, B., PEARCE, EDWARD J., DRIGGERS, EDWARD M. & ARTYOMOV, MAXIM N. 2015. Network Integration of Parallel Metabolic and Transcriptional Data Reveals Metabolic Modules that Regulate Macrophage Polarization. *Immunity*, 42, 419-430.
- JIANG, L., CHEN, H., PINELLO, L. & YUAN, G. C. 2016. GiniClust: detecting rare cell types from single-cell gene expression data with Gini index. *Genome Biol*, 17, 144.
- JIANG, P., GU, S., PAN, D., FU, J., SAHU, A., HU, X., LI, Z., TRAUGH, N., BU, X., LI, B., LIU, J., FREEMAN, G. J., BROWN, M. A., WUCHERPFENNIG, K. W. & LIU, X. S. 2018. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med*, 24, 1550-1558.
- JOHNSON, M. O., WOLF, M. M., MADDEN, M. Z., ANDREJEVA, G., SUGIURA, A., CONTRERAS, D. C., MASEDA, D., LIBERTI, M. V., PAZ, K., KISHTON, R. J., JOHNSON, M. E., DE CUBAS, A. A., WU, P., LI, G., ZHANG, Y., NEWCOMB, D. C., WELLS, A. D., RESTIFO, N. P., RATHMELL, W. K., LOCASALE, J. W., DAVILA, M. L., BLAZAR, B. R. & RATHMELL, J. C. 2018. Distinct Regulation of Th17 and Th1 Cell Differentiation by Glutaminase-Dependent Metabolism. *Cell*, 175, 1780-1795.e19.
- JURTZ, V., PAUL, S., ANDREATTA, M., MARCATILI, P., PETERS, B. & NIELSEN, M. 2017. NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. *J Immunol*, 199, 3360-3368.
- KANDOTH, C., MCLELLAN, M. D., VANDIN, F., YE, K., NIU, B., LU, C., XIE, M., ZHANG, Q., MCMICHAEL, J. F., WYCZALKOWSKI, M. A., LEISERSON, M. D. M., MILLER, C. A., WELCH, J. S., WALTER, M. J., WENDL, M. C., LEY, T. J., WILSON, R. K., RAPHAEL, B. J. & DING, L. 2013. Mutational landscape and significance across 12 major cancer types. *Nature*, 502, 333-339.
- KANEHISA, M., GOTO, S., SATO, Y., FURUMICHI, M. & TANABE, M. 2012. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res*, 40, D109-14.
- KANNAN, Y., PEREZ-LLORET, J., LI, Y., ENTWISTLE, L. J., KHOURY, H., PAPOUTSOPOULOU, S., MAHMOOD, R., MANSOUR, N. R., CHING-CHENG HUANG, S., PEARCE, E. J., PEDRO, S. D. C. L., LEY, S. C. & WILSON, M. S. 2016. TPL-2 Regulates Macrophage Lipid Metabolism and M2 Differentiation to Control TH2-Mediated Immunopathology. *PLoS Pathog*, 12, e1005783.
- KAWAGUCHI, S., HIGASA, K., SHIMIZU, M., YAMADA, R. & MATSUDA, F. 2017. HLA-HD: An accurate HLA typing algorithm for next-generation sequencing data. *Hum Mutat*, 38, 788-797.

- KAWALEKAR, O. U., RS, O. C., FRAIETTA, J. A., GUO, L., MCGETTIGAN, S. E., POSEY, A. D., JR., PATEL, P. R., GUEDAN, S., SCHOLLER, J., KEITH, B., SNYDER, N. W., BLAIR, I. A., MILONE, M. C. & JUNE, C. H. 2016. Distinct Signaling of Coreceptors Regulates Specific Metabolism Pathways and Impacts Memory Development in CAR T Cells. *Immunity*, 44, 712.
- KELDER, T., VAN IERSEL, M. P., HANSPERS, K., KUTMON, M., CONKLIN, B. R., EVELO, C. T. & PICO, A. R. 2012. WikiPathways: building research communities on biological pathways. *Nucleic Acids Res*, 40, D1301-7.
- KELLY, M. P., JUNGBLUTH, A. A., WU, B. W., BOMALASKI, J., OLD, L. J. & RITTER, G. 2012. Arginine deiminase PEG20 inhibits growth of small cell lung cancers lacking expression of argininosuccinate synthetase. *British Journal of Cancer*, 106, 324-332.
- KEREN, L., BOSSE, M., MARQUEZ, D., ANGOSHTARI, R., JAIN, S., VARMA, S., YANG, S. R., KURIAN, A., VAN VALEN, D., WEST, R., BENDALL, S. C. & ANGELO, M. 2018. A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging. *Cell*, 174, 1373-1387.e19.
- KEREN, L., BOSSE, M., THOMPSON, S., RISOM, T., VIJAYARAGAVAN, K., MCCAFFREY, E., MARQUEZ, D., ANGOSHTARI, R., GREENWALD, N. F., FIENBERG, H., WANG, J., KAMBHAM, N., KIRKWOOD, D., NOLAN, G., MONTINE, T. J., GALLI, S. J., WEST, R., BENDALL, S. C. & ANGELO, M. 2019. MIBI-TOF: A multiplexed imaging platform relates cellular phenotypes and tissue structure. *Science Advances*, 5, eaax5851.
- KIMMEY, S. C., BORGES, L., BASKAR, R. & BENDALL, S. C. 2019. Parallel analysis of tri-molecular biosynthesis with cell identity and function in single cells. *Nature Communications*, 10, 1185.
- KISELEV, V. Y., YIU, A. & HEMBERG, M. 2018. scmap: projection of single-cell RNA-seq data across data sets. *Nat Methods*, 15, 359-362.
- KLEBANOFF, C. A., CROMPTON, J. G., LEONARDI, A. J., YAMAMOTO, T. N., CHANDRAN, S. S., EIL, R. L., SUKUMAR, M., VODNALA, S. K., HU, J., JI, Y., CLEVER, D., BLACK, M. A., GURUSAMY, D., KRUHLAK, M. J., JIN, P., STRONCEK, D. F., GATTINONI, L., FELDMAN, S. A. & RESTIFO, N. P. 2017. Inhibition of AKT signaling uncouples T cell differentiation from expansion for receptor-engineered adoptive immunotherapy. *JCI Insight*, 2.
- KLEBANOFF, C. A., FINKELSTEIN, S. E., SURMAN, D. R., LICHTMAN, M. K., GATTINONI, L., THEORET, M. R., GREWAL, N., SPIESS, P. J., ANTONY, P. A., PALMER, D. C., TAGAYA, Y., ROSENBERG, S. A., WALDMANN, T. A. & RESTIFO, N. P. 2004. IL-15 enhances the in vivo antitumor activity of tumor-reactive CD8+ T cells. *Proc Natl Acad Sci U S A*, 101, 1969-74.
- KLEBANOFF, C. A., GATTINONI, L., PALMER, D. C., MURANSKI, P., JI, Y., HINRICHS, C. S., BORMAN, Z. A., KERKAR, S. P., SCOTT, C. D., FINKELSTEIN, S. E., ROSENBERG, S. A. & RESTIFO, N. P. 2011. Determinants of successful CD8+ T-cell adoptive immunotherapy for large established tumors in mice. *Clin Cancer Res*, 17, 5343-52.
- KLEIN GELTINK, R. I., O'SULLIVAN, D., CORRADO, M., BREMSER, A., BUCK, M. D., BUESCHER, J. M., FIRAT, E., ZHU, X., NIEDERMANN, G., CAPUTA, G., KELLY, B., WARTHORST, U., RENSING-EHL, A., KYLE, R. L., VANDERSARREN, L., CURTIS, J. D., PATTERSON, A. E., LAWLESS, S., GRZES, K., QIU, J., SANIN, D. E., KRETZ, O., HUBER, T. B., JANSSENS, S., LAMBRECHT, B. N., RAMBOLD, A. S., PEARCE, E. J. & PEARCE, E. L. 2017. Mitochondrial Priming by CD28. *Cell*, 171, 385-397.e11.
- KLYSZ, D., TAI, X., ROBERT, P. A., CRAVEIRO, M., CRETENET, G., OBUROGLU, L., MONGELLAZ, C., FLOESS, S., FRITZ, V., MATIAS, M. I., YONG, C., SURH, N., MARIE, J. C., HUEHN, J., ZIMMERMANN, V., KINET, S., DARDALHON, V. & TAYLOR, N. 2015. Glutamine-dependent α -ketoglutarate production

- regulates the balance between T helper 1 cell and regulatory T cell generation. *Science Signaling*, 8, ra97-ra97.
- KO, Y.-H., LIN, Z., FLOMENBERG, N., PESTELL, R. G., HOWELL, A., SOTGIA, F., LISANTI, M. P. & MARTINEZ-OUTSCHOORN, U. E. 2011. Glutamine fuels a vicious cycle of autophagy in the tumor stroma and oxidative mitochondrial metabolism in epithelial cancer cells. *Cancer Biology & Therapy*, 12, 1085-1097.
- KRAWCZYK, C. M., HOLOWKA, T., SUN, J., BLAGIH, J., AMIEL, E., DEBERARDINIS, R. J., CROSS, J. R., JUNG, E., THOMPSON, C. B., JONES, R. G. & PEARCE, E. J. 2010. Toll-like receptor-induced changes in glycolytic metabolism regulate dendritic cell activation. *Blood*, 115, 4742-4749.
- KULESHOV, M. V., JONES, M. R., ROUILLARD, A. D., FERNANDEZ, N. F., DUAN, Q., WANG, Z., KOPLEV, S., JENKINS, S. L., JAGODNIK, K. M., LACHMANN, A., MCDERMOTT, M. G., MONTEIRO, C. D., GUNDERSEN, G. W. & MA'AYAN, A. 2016. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*, 44, W90-7.
- KUNG, C., HIXON, J., CHOE, S., MARKS, K., GROSS, S., MURPHY, E., DELABARRE, B., CIANCHETTA, G., SETHUMADHAVAN, S., WANG, X., YAN, S., GAO, Y., FANG, C., WEI, W., JIANG, F., WANG, S., QIAN, K., SAUNDERS, J., DRIGGERS, E., HIN, KUNII, K., MURRAY, S., YANG, H., YEN, K., LIU, W., LEWIS, MATTHEW, SHINSAN, JIN, S., FRANCESCO & DANG, L. 2012. Small Molecule Activation of PKM2 in Cancer Cells Induces Serine Auxotrophy. *Chemistry & Biology*, 19, 1187-1198.
- KUTMON, M., VAN IERSEL, M. P., BOHLER, A., KELDER, T., NUNES, N., PICO, A. R. & EVELO, C. T. 2015. PathVisio 3: an extendable pathway analysis toolbox. *PLoS Comput Biol*, 11, e1004085.
- LA MANNO, G., SOLDATOV, R., ZEISEL, A., BRAUN, E., HOCHGERNER, H., PETUKHOV, V., LIDSCHREIBER, K., KASTRITI, M. E., LÖNNERBERG, P., FURLAN, A., FAN, J., BORM, L. E., LIU, Z., VAN BRUGGEN, D., GUO, J., HE, X., BARKER, R., SUNDSTRÖM, E., CASTELO-BRANCO, G., CRAMER, P., ADAMEYKO, I., LINNARSSON, S. & KHARCHENKO, P. V. 2018. RNA velocity of single cells. *Nature*, 560, 494-498.
- LAM, WING Y., BECKER, AMY M., KENNERLY, KRISTA M., WONG, R., CURTIS, JONATHAN D., LLUFRIO, ELIZABETH M., MCCOMMIS, KYLE S., FAHRMANN, J., PIZZATO, HANNAH A., NUNLEY, RYAN M., LEE, J., WOLFGANG, MICHAEL J., PATTI, GARY J., FINCK, BRIAN N., PEARCE, ERIKA L. & BHATTACHARYA, D. 2016. Mitochondrial Pyruvate Import Promotes Long-Term Survival of Antibody-Secreting Plasma Cells. *Immunity*, 45, 60-73.
- LAMBRECHTS, D., WAUTERS, E., BOECKX, B., AIBAR, S., NITTNER, D., BURTON, O., BASSEZ, A., DECALUWÉ, H., PIRCHER, A., VAN DEN EYNDE, K., WEYNAND, B., VERBEKEN, E., DE LEYN, P., LISTON, A., VANSTEENKISTE, J., CARMELIET, P., AERTS, S. & THIENPONT, B. 2018. Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat Med*, 24, 1277-1289.
- LAMPROPOULOU, V., SERGUSHICHEV, A., BAMBOUSKOVA, M., NAIR, S., VINCENT, E. E., LOGINICHEVA, E., CERVANTES-BARRAGAN, L., MA, X., HUANG, S. C., GRISS, T., WEINHEIMER, C. J., KHADER, S., RANDOLPH, G. J., PEARCE, E. J., JONES, R. G., DIWAN, A., DIAMOND, M. S. & ARTYOMOV, M. N. 2016. Itaconate Links Inhibition of Succinate Dehydrogenase with Macrophage Metabolic Remodeling and Regulation of Inflammation. *Cell Metab*, 24, 158-66.
- LARSSON, K., KOCK, A., IDBORG, H., ARSENIAN HENRIKSSON, M., MARTINSSON, T., JOHNSEN, J. I., KOROTKOVA, M., KOGNER, P. & JAKOBSSON, P. J. 2015. COX/mPGES-1/PGE2 pathway depicts an inflammatory-dependent high-risk neuroblastoma subset. *Proc Natl Acad Sci U S A*, 112, 8070-5.
- LAVIN, Y., KOBAYASHI, S., LEADER, A., AMIR, E. D., ELEFANT, N., BIGENWALD, C., REMARK, R., SWEENEY, R., BECKER, C. D., LEVINE, J. H., MEINHOF, K., CHOW, A., KIM-SHULZE, S., WOLF, A., MEDAGLIA, C., LI, H., RYTLEWSKI, J. A., EMERSON, R. O., SOLOVYOV, A., GREENBAUM, B. D., SANDERS, C., VIGNALI,

- M., BEASLEY, M. B., FLORES, R., GNJATIC, S., PE'ER, D., RAHMAN, A., AMIT, I. & MERAD, M. 2017. Innate Immune Landscape in Early Lung Adenocarcinoma by Paired Single-Cell Analyses. *Cell*, 169, 750-765.e17.
- LAW, C. W., CHEN, Y., SHI, W. & SMYTH, G. K. 2014. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*, 15, R29.
- LEE, H. & KINGSFORD, C. 2018. Kourami: graph-guided assembly for novel human leukocyte antigen allele discovery. *Genome Biol*, 19, 16.
- LEVINE, J. H., SIMONDS, E. F., BENDALL, S. C., DAVIS, K. L., AMIR EL, A. D., TADMOR, M. D., LITVIN, O., FIENBERG, H. G., JAGER, A., ZUNDER, E. R., FINCK, R., GEDMAN, A. L., RADTKE, I., DOWNING, J. R., PE'ER, D. & NOLAN, G. P. 2015. Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell*, 162, 184-97.
- LI, B., LI, T., PIGNON, J. C., WANG, B., WANG, J., SHUKLA, S. A., DOU, R., CHEN, Q., HODI, F. S., CHOUERI, T. K., WU, C., HACOEN, N., SIGNORETTI, S., LIU, J. S. & LIU, X. S. 2016a. Landscape of tumor-infiltrating T cell repertoire of human cancers. *Nat Genet*, 48, 725-32.
- LI, B., SEVERSON, E., PIGNON, J. C., ZHAO, H., LI, T., NOVAK, J., JIANG, P., SHEN, H., ASTER, J. C., RODIG, S., SIGNORETTI, S., LIU, J. S. & LIU, X. S. 2016b. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. *Genome Biol*, 17, 174.
- LI, F.-L., LIU, J.-P., BAO, R.-X., YAN, G., FENG, X., XU, Y.-P., SUN, Y.-P., YAN, W., LING, Z.-Q., XIONG, Y., GUAN, K.-L. & YUAN, H.-X. 2018a. Acetylation accumulates PFKFB3 in cytoplasm to promote glycolysis and protects cells from cisplatin-induced apoptosis. *Nature Communications*, 9.
- LI, H.-M., YANG, J.-G., LIU, Z.-J., WANG, W.-M., YU, Z.-L., REN, J.-G., CHEN, G., ZHANG, W. & JIA, J. 2017a. Blockage of glycolysis by targeting PFKFB3 suppresses tumor growth and metastasis in head and neck squamous cell carcinoma. *Journal of Experimental & Clinical Cancer Research*, 36.
- LI, H., COURTOIS, E. T., SENGUPTA, D., TAN, Y., CHEN, K. H., GOH, J. J. L., KONG, S. L., CHUA, C., HON, L. K., TAN, W. S., WONG, M., CHOI, P. J., WEE, L. J. K., HILLMER, A. M., TAN, I. B., ROBSON, P. & PRABHAKAR, S. 2017b. Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat Genet*, 49, 708-718.
- LI, L., HUANG, L., YE, H., SONG, S. P., BAJWA, A., LEE, S. J., MOSER, E. K., JAWORSKA, K., KINSEY, G. R., DAY, Y. J., LINDEN, J., LOBO, P. I., ROSIN, D. L. & OKUSA, M. D. 2012. Dendritic cells tolerized with adenosine A₂AR agonist attenuate acute kidney injury. *J Clin Invest*, 122, 3931-42.
- LI, R., WEI, F., YU, J., LI, H., REN, X. & HAO, X. 2009. IDO inhibits T-cell function through suppressing Vav1 expression and activation. *Cancer Biol Ther*, 8, 1402-8.
- LI, W., TANIKAWA, T., KRYCZEK, I., XIA, H., LI, G., WU, K., WEI, S., ZHAO, L., VATAN, L., WEN, B., SHU, P., SUN, D., KLEER, C., WICHA, M., SABEL, M., TAO, K., WANG, G. & ZOU, W. 2018b. Aerobic Glycolysis Controls Myeloid-Derived Suppressor Cells and Tumor Immunity via a Specific CEBPB Isoform in Triple-Negative Breast Cancer. *Cell Metabolism*, 28, 87-103.e6.
- LI, Y., FANG, M., ZHANG, J., WANG, J., SONG, Y., SHI, J., LI, W., WU, G., REN, J., WANG, Z., ZOU, W. & WANG, L. 2016c. Hydrogel dual delivered celecoxib and anti-PD-1 synergistically improve antitumor immunity. *Oncoimmunology*, 5, e1074374.
- LIM, S., LIU, H., MADEIRA DA SILVA, L., ARORA, R., LIU, Z., PHILLIPS, J. B., SCHMITT, D. C., VU, T., MCCLELLAN, S., LIN, Y., LIN, W., PIAZZA, G. A., FODSTAD, O. & TAN, M. 2016. Immunoregulatory Protein B7-H3 Reprograms Glucose Metabolism in Cancer Cells by ROS-Mediated Stabilization of HIF1 α . *Cancer Res*, 76, 2231-42.

- LINDEMAN, I., EMERTON, G., MAMANOVA, L., SNIR, O., POLANSKI, K., QIAO, S. W., SOLLID, L. M., TEICHMANN, S. A. & STUBBINGTON, M. J. T. 2018. BraCeR: B-cell-receptor reconstruction and clonality inference from single-cell RNA-seq. *Nat Methods*, 15, 563-565.
- LIU, H., SHEN, Z., WANG, Z., WANG, X., ZHANG, H., QIN, J., QIN, X., XU, J. & SUN, Y. 2016. Increased expression of IDO associates with poor postoperative clinical outcome of patients with gastric adenocarcinoma. *Sci Rep*, 6, 21319.
- LIU, P.-S., WANG, H., LI, X., CHAO, T., TEAV, T., CHRISTEN, S., DI CONZA, G., CHENG, W.-C., CHOU, C.-H., VAVAKOVA, M., MURET, C., DEBACKERE, K., MAZZONE, M., HUANG, H.-D., FENDT, S.-M., IVANISEVIC, J. & HO, P.-C. 2017. α -ketoglutarate orchestrates macrophage activation through metabolic and epigenetic reprogramming. *Nature Immunology*, 18, 985-994.
- LIU, Z., CUI, Y., XIONG, Z., NASIRI, A., ZHANG, A. & HU, J. 2019. DeepSeqPan, a novel deep convolutional neural network model for pan-specific class I HLA-peptide binding affinity prediction. *Sci Rep*, 9, 794.
- LLUFRIO, E. M., WANG, L., NASER, F. J. & PATTI, G. J. 2018. Sorting cells alters their redox state and cellular metabolome. *Redox Biology*, 16, 381-387.
- LUN, A. T., MCCARTHY, D. J. & MARIONI, J. C. 2016. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Res*, 5, 2122.
- MA, E. H., BANTUG, G., GRISS, T., CONDOTTA, S., JOHNSON, R. M., SAMBORSKA, B., MAINOLFI, N., SURI, V., GUAK, H., BALMER, M. L., VERWAY, M. J., RAISSI, T. C., TSUI, H., BOUKHALED, G., HENRIQUES DA COSTA, S., FREZZA, C., KRAWCZYK, C. M., FRIEDMAN, A., MANFREDI, M., RICHER, M. J., HESS, C. & JONES, R. G. 2017. Serine Is an Essential Metabolite for Effector T Cell Expansion. *Cell Metabolism*, 25, 345-357.
- MA, E. H., VERWAY, M. J., JOHNSON, R. M., ROY, D. G., STEADMAN, M., HAYES, S., WILLIAMS, K. S., SHELDON, R. D., SAMBORSKA, B., KOSINSKI, P. A., KIM, H., GRISS, T., FAUBERT, B., CONDOTTA, S. A., KRAWCZYK, C. M., DEBERARDINIS, R. J., STEWART, K. M., RICHER, M. J., CHUBUKOV, V., RODDY, T. P. & JONES, R. G. 2019. Metabolic Profiling Using Stable Isotope Tracing Reveals Distinct Patterns of Glucose Utilization by Physiologically Activated CD8⁺ T Cells. *Immunity*, 51, 856-870.e5.
- MACARTHUR, D. G., MANOLIO, T. A., DIMMOCK, D. P., REHM, H. L., SHENDURE, J., ABECASIS, G. R., ADAMS, D. R., ALTMAN, R. B., ANTONARAKIS, S. E., ASHLEY, E. A., BARRETT, J. C., BIESECKER, L. G., CONRAD, D. F., COOPER, G. M., COX, N. J., DALY, M. J., GERSTEIN, M. B., GOLDSTEIN, D. B., HIRSCHHORN, J. N., LEAL, S. M., PENNACCHIO, L. A., STAMATOYANNOPOULOS, J. A., SUNYAEV, S. R., VALLE, D., VOIGHT, B. F., WINCKLER, W. & GUNTER, C. 2014. Guidelines for investigating causality of sequence variants in human disease. *Nature*, 508, 469-76.
- MAHIC, M., Yaqub, S., JOHANSSON, C. C., TASKÉN, K. & AANDAHL, E. M. 2006. FOXP3⁺CD4⁺CD25⁺ adaptive regulatory T cells express cyclooxygenase-2 and suppress effector T cells by a prostaglandin E2-dependent mechanism. *J Immunol*, 177, 246-54.
- MAJ, T., WANG, W., CRESPO, J., ZHANG, H., WANG, W., WEI, S., ZHAO, L., VATAN, L., SHAO, I., SZELIGA, W., LYSSOTIS, C., LIU, J. R., KRYCZEK, I. & ZOU, W. 2017. Oxidative stress controls regulatory T cell apoptosis and suppressor activity and PD-L1-blockade resistance in tumor. *Nat Immunol*, 18, 1332-1341.
- MAZUMDAR, C., DRIGGERS, E. M. & TURKA, L. A. 2020. The Untapped Opportunity and Challenge of Immunometabolism: A New Paradigm for Drug Discovery. *Cell Metabolism*, 31, 26-34.
- MCCARTHY, D. J., CAMPBELL, K. R., LUN, A. T. & WILLS, Q. F. 2017. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics*, 33, 1179-1186.

- MENK, A. V., SCHARPING, N. E., MORECI, R. S., ZENG, X., GUY, C., SALVATORE, S., BAE, H., XIE, J., YOUNG, H. A., WENDELL, S. G. & DELGOFFE, G. M. 2018. Early TCR Signaling Induces Rapid Aerobic Glycolysis Enabling Distinct Acute T Cell Effector Functions. *Cell Rep*, 22, 1509-1521.
- MEREU, E., LAFZI, A., MOUTINHO, C., ZIEGENHAIN, C., MCCARTHY, D. J., ÁLVAREZ-VARELA, A., BATLLE, E., SAGAR, GRÜN, D., LAU, J. K., BOUTET, S. C., SANADA, C., OOI, A., JONES, R. C., KAIHARA, K., BRAMPTON, C., TALAGA, Y., SASAGAWA, Y., TANAKA, K., HAYASHI, T., BRAEUNING, C., FISCHER, C., SAUER, S., TREFZER, T., CONRAD, C., ADICONIS, X., NGUYEN, L. T., REGEV, A., LEVIN, J. Z., PAREKH, S., JANJIC, A., WANGE, L. E., BAGNOLI, J. W., ENARD, W., GUT, M., SANDBERG, R., NIKAIDO, I., GUT, I., STEGLE, O. & HEYN, H. 2020. Benchmarking single-cell RNA-sequencing protocols for cell atlas projects. *Nature Biotechnology*, 38, 747-755.
- MERIGHI, S., MIRANDOLA, P., MILANI, D., VARANI, K., GESSI, S., KLOTZ, K. N., LEUNG, E., BARALDI, P. G. & BOREA, P. A. 2002. Adenosine receptors as mediators of both cell proliferation and cell death of cultured human melanoma cells. *J Invest Dermatol*, 119, 923-33.
- MI, H., MURUGANUJAN, A. & THOMAS, P. D. 2013. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res*, 41, D377-86.
- MICHALEK, R. D., GERRIETS, V. A., JACOBS, S. R., MACINTYRE, A. N., MACIVER, N. J., MASON, E. F., SULLIVAN, S. A., NICHOLS, A. G. & RATHMELL, J. C. 2011. Cutting Edge: Distinct Glycolytic and Lipid Oxidative Metabolic Programs Are Essential for Effector and Regulatory CD4⁺ T Cell Subsets. *The Journal of Immunology*, 186, 3299-3303.
- MILLER, A., NAGY, C., KNAPP, B., LAENGLE, J., PONWEISER, E., GROEGER, M., STARKL, P., BERGMANN, M., WAGNER, O. & HASCHEMI, A. 2017. Exploring Metabolic Configurations of Single Cells within Complex Tissue Microenvironments. *Cell Metabolism*, 26, 788-800.e6.
- MILLS, E. L., RYAN, D. G., PRAG, H. A., DIKOVSKAYA, D., MENON, D., ZASLONA, Z., JEDRYCHOWSKI, M. P., COSTA, A. S. H., HIGGINS, M., HAMS, E., SZPYT, J., RUNTSCH, M. C., KING, M. S., MCGOURAN, J. F., FISCHER, R., KESSLER, B. M., MCGETTRICK, A. F., HUGHES, M. M., CARROLL, R. G., BOOTY, L. M., KNATKO, E. V., MEAKIN, P. J., ASHFORD, M. L. J., MODIS, L. K., BRUNORI, G., SÉVIN, D. C., FALLON, P. G., CALDWELL, S. T., KUNJI, E. R. S., CHOUCANI, E. T., FREZZA, C., DINKOVA-KOSTOVA, A. T., HARTLEY, R. C., MURPHY, M. P. & O'NEILL, L. A. 2018. Itaconate is an anti-inflammatory metabolite that activates Nrf2 via alkylation of KEAP1. *Nature*, 556, 113-117.
- MINHAS, P. S., LIU, L., MOON, P. K., JOSHI, A. U., DOVE, C., MHATRE, S., CONTREPOIS, K., WANG, Q., LEE, B. A., CORONADO, M., BERNSTEIN, D., SNYDER, M. P., MIGAUD, M., MAJETI, R., MOCHLY-ROSEN, D., RABINOWITZ, J. D. & ANDREASSON, K. I. 2019. Macrophage de novo NAD(+) synthesis specifies immune function in aging and inflammation. *Nat Immunol*, 20, 50-63.
- MIRAGAIA, R. J., GOMES, T., CHOMKA, A., JARDINE, L., RIEDEL, A., HEGAZY, A. N., WHIBLEY, N., TUCCI, A., CHEN, X., LINDEMAN, I., EMERTON, G., KRAUSGRUBER, T., SHIELDS, J., HANIFFA, M., POWRIE, F. & TEICHMANN, S. A. 2019. Single-Cell Transcriptomics of Regulatory T Cells Reveals Trajectories of Tissue Adaptation. *Immunity*, 50, 493-504.e7.
- MITTAL, D., YOUNG, A., STANNARD, K., YONG, M., TENG, M. W., ALLARD, B., STAGG, J. & SMYTH, M. J. 2014. Antimetastatic effects of blocking PD-1 and the adenosine A2A receptor. *Cancer Res*, 74, 3652-8.
- MOON, J.-S., HISATA, S., PARK, M.-A., DENICOLA, GINA M., RYTER, STEFAN W., NAKAHIRA, K. & CHOI, AUGUSTINE M. K. 2015. mTORC1-Induced HK1-Dependent Glycolysis Regulates NLRP3 Inflammasome Activation. *Cell Reports*, 12, 102-115.

- MOSE, L. E., SELITSKY, S. R., BIXBY, L. M., MARRON, D. L., IGLESIA, M. D., SERODY, J. S., PEROU, C. M., VINCENT, B. G. & PARKER, J. S. 2016. Assembly-based inference of B-cell receptor repertoires from short read RNA sequencing data with V'DJer. *Bioinformatics*, 32, 3729-3734.
- MUNN, D. H., SHARMA, M. D., BABAN, B., HARDING, H. P., ZHANG, Y., RON, D. & MELLOR, A. L. 2005. GCN2 kinase in T cells mediates proliferative arrest and anergy induction in response to indoleamine 2,3-dioxygenase. *Immunity*, 22, 633-42.
- MURERA, D., ARBOGAST, F., ARNOLD, J., BOUIS, D., MULLER, S. & GROS, F. 2018. CD4 T cell autophagy is integral to memory maintenance. *Sci Rep*, 8, 5951.
- MURRAY, C. M., HUTCHINSON, R., BANTICK, J. R., BELFIELD, G. P., BENJAMIN, A. D., BRAZMA, D., BUNDICK, R. V., COOK, I. D., CRAGGS, R. I., EDWARDS, S., EVANS, L. R., HARRISON, R., HOLNESS, E., JACKSON, A. P., JACKSON, C. G., KINGSTON, L. P., PERRY, M. W. D., ROSS, A. R. J., RUGMAN, P. A., SIDHU, S. S., SULLIVAN, M., TAYLOR-FISHWICK, D. A., WALKER, P. C., WHITEHEAD, Y. M., WILKINSON, D. J., WRIGHT, A. & DONALD, D. K. 2005. Monocarboxylate transporter MCT1 is a target for immunosuppression. *Nature Chemical Biology*, 1, 371-376.
- NABE, S., YAMADA, T., SUZUKI, J., TORIYAMA, K., YASUOKA, T., KUWAHARA, M., SHIRAISHI, A., TAKENAKA, K., YASUKAWA, M. & YAMASHITA, M. 2018. Reinforce the antitumor activity of CD 8 + T cells via glutamine restriction. *Cancer Science*, 109, 3737-3750.
- NAKAYA, M., XIAO, Y., ZHOU, X., CHANG, J.-H., CHANG, M., CHENG, X., BLONSKA, M., LIN, X. & SUN, S.-C. 2014. Inflammatory T Cell Responses Rely on Amino Acid Transporter ASCT2 Facilitation of Glutamine Uptake and mTORC1 Kinase Activation. *Immunity*, 40, 692-705.
- NEWMAN, A. M., LIU, C. L., GREEN, M. R., GENTLES, A. J., FENG, W., XU, Y., HOANG, C. D., DIEHN, M. & ALIZADEH, A. A. 2015. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*, 12, 453-7.
- NEWMAN, A. M., STEEN, C. B., LIU, C. L., GENTLES, A. J., CHAUDHURI, A. A., SCHERER, F., KHODADOUST, M. S., ESFAHANI, M. S., LUCA, B. A., STEINER, D., DIEHN, M. & ALIZADEH, A. A. 2019. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol*, 37, 773-782.
- NIELSEN, M., LUNDEGAARD, C., WORNING, P., LAUEMØLLER, S. L., LAMBERTH, K., BUUS, S., BRUNAK, S. & LUND, O. 2003. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci*, 12, 1007-17.
- NIU, Z., SHI, Q., ZHANG, W., SHU, Y., YANG, N., CHEN, B., WANG, Q., ZHAO, X., CHEN, J., CHENG, N., FENG, X., HUA, Z., JI, J. & SHEN, P. 2017. Caspase-1 cleaves PPAR γ for potentiating the pro-tumor action of TAMs. *Nat Commun*, 8, 766.
- NOMURA, M., LIU, J., ROVIRA, I. I., GONZALEZ-HURTADO, E., LEE, J., WOLFGANG, M. J. & FINKEL, T. 2016. Fatty acid oxidation in macrophage polarization. *Nature Immunology*, 17, 216-217.
- O'DONNELL, T. J., RUBINSTEYN, A., BONSACK, M., RIEMER, A. B., LASERSON, U. & HAMMERBACHER, J. 2018. MHCflurry: Open-Source Class I MHC Binding Affinity Prediction. *Cell Syst*, 7, 129-132.e4.
- O'SULLIVAN, T. E., JOHNSON, L. R., KANG, H. H. & SUN, J. C. 2015. BNIP3- and BNIP3L-Mediated Mitophagy Promotes the Generation of Natural Killer Cell Memory. *Immunity*, 43, 331-42.
- OHTA, A., KINI, R., OHTA, A., SUBRAMANIAN, M., MADASU, M. & SITKOVSKY, M. 2012. The development and immunosuppressive functions of CD4⁺ CD25⁺ FoxP3⁺ regulatory T cells are under influence of the adenosine-A2A adenosine receptor pathway. *Frontiers in Immunology*, 3.
- ORENBUCH, R., FILIP, I., COMITO, D., SHAMAN, J., PE'ER, I. & RABADAN, R. 2019. arcasHLA: high-resolution HLA typing from RNAseq. *Bioinformatics*, 36, 33-40.

- OTT, P. A., HU, Z., KESKIN, D. B., SHUKLA, S. A., SUN, J., BOZYM, D. J., ZHANG, W., LUOMA, A., GIOBBIE-HURDER, A., PETER, L., CHEN, C., OLIVE, O., CARTER, T. A., LI, S., LIEB, D. J., EISENHAURE, T., GJINI, E., STEVENS, J., LANE, W. J., JAVERI, I., NELLAIPPAN, K., SALAZAR, A. M., DALEY, H., SEAMAN, M., BUCHBINDER, E. I., YOON, C. H., HARDEN, M., LENNON, N., GABRIEL, S., RODIG, S. J., BAROUCH, D. H., ASTER, J. C., GETZ, G., WUCHERPFENNIG, K., NEUBERG, D., RITZ, J., LANDER, E. S., FRITSCH, E. F., HACHOEN, N. & WU, C. J. 2017. An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature*, 547, 217-221.
- PALSSON-MCDERMOTT, EVA M., CURTIS, ANNE M., GOEL, G., LAUTERBACH, MARIO A. R., SHEEDY, FREDERICK J., GLEESON, LAURA E., VAN DEN BOSCH, MIRIAM W. M., QUINN, SUSAN R., DOMINGO-FERNANDEZ, R., JOHNSTON, DANIEL G. W., JIANG, J.-K., ISRAELSEN, WILLIAM J., KEANE, J., THOMAS, C., CLISH, C., VANDER HEIDEN, M., XAVIER, RAMNIK J. & O'NEILL, LUKE A. J. 2015. Pyruvate Kinase M2 Regulates Hif-1 α Activity and IL-1 β Induction and Is a Critical Determinant of the Warburg Effect in LPS-Activated Macrophages. *Cell Metabolism*, 21, 65-80.
- PALSSON-MCDERMOTT, E. M., DYCK, L., ZASŁONA, Z., MENON, D., MCGETTRICK, A. F., MILLS, K. H. G. & O'NEILL, L. A. 2017. Pyruvate Kinase M2 Is Required for the Expression of the Immune Checkpoint PD-L1 in Immune Cells and Tumors. *Frontiers in Immunology*, 8.
- PATSOUKIS, N., BARDHAN, K., CHATTERJEE, P., SARI, D., LIU, B., BELL, L. N., KAROLY, E. D., FREEMAN, G. J., PETKOVA, V., SETH, P., LI, L. & BOUSSIOTIS, V. A. 2015. PD-1 alters T-cell metabolic reprogramming by inhibiting glycolysis and promoting lipolysis and fatty acid oxidation. *Nat Commun*, 6, 6692.
- PELGROM, L. R., VAN DER HAM, A. J. & EVERTS, B. 2016. Analysis of TLR-Induced Metabolic Changes in Dendritic Cells Using the Seahorse XFe96 Extracellular Flux Analyzer. In: MCCOY, C. E. (ed.) *Toll-Like Receptors: Practice and Methods*. New York, NY: Springer New York.
- PÉREZ-ESCUREDO, J., DADHICH, R. K., DHUP, S., CACACE, A., VAN HÉE, V. F., DE SAEDELEER, C. J., SBOARINA, M., RODRIGUEZ, F., FONTENILLE, M.-J., BRISSON, L., PORPORATO, P. E. & SONVEAUX, P. 2016. Lactate promotes glutamine uptake and metabolism in oxidative cancer cells. *Cell Cycle*, 15, 72-83.
- PIERSON, E. & YAU, C. 2015. ZIFA: Dimensionality reduction for zero-inflated single-cell gene expression analysis. *Genome Biol*, 16, 241.
- PILON-THOMAS, S., KODUMUDI, K. N., EL-KENAWI, A. E., RUSSELL, S., WEBER, A. M., LUDDY, K., DAMAGHI, M., WOJTKOWIAK, J. W., MULÉ, J. J., IBRAHIM-HASHIM, A. & GILLIES, R. J. 2016. Neutralization of Tumor Acidity Improves Antitumor Responses to Immunotherapy. *Cancer Research*, 76, 1381-1390.
- POULOPOULOU, C., MARKAKIS, I., DAVAKI, P., NIKOLAOU, C., POULOPOULOS, A., RAPTIS, E. & VASSILOPOULOS, D. 2005. Modulation of voltage-gated potassium channels in human T lymphocytes by extracellular glutamate. *Mol Pharmacol*, 67, 856-67.
- PREVITE, D. M., MENK, A. V., WRIGHT, R. H., COUDRIET, G. M., DELGOFFE, G. M. & PIGANELLI, J. D. 2017. Lymphocyte activation gene-3 regulates mitochondrial biogenesis and metabolism of naive CD4⁺ T cells. *The Journal of Immunology*, 198, 150.1-150.1.
- PURAM, S. V., TIROSH, I., PARIKH, A. S., PATEL, A. P., YIZHAK, K., GILLESPIE, S., RODMAN, C., LUO, C. L., MROZ, E. A., EMERICK, K. S., DESCHLER, D. G., VARVARES, M. A., MYLVAGANAM, R., ROZENBLATT-ROSEN, O., ROCCO, J. W., FAQUIN, W. C., LIN, D. T., REGEV, A. & BERNSTEIN, B. E. 2017. Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell*, 171, 1611-1624.e24.

- QIU, F., CHEN, Y. R., LIU, X., CHU, C. Y., SHEN, L. J., XU, J., GAUR, S., FORMAN, H. J., ZHANG, H., ZHENG, S., YEN, Y., HUANG, J., KUNG, H. J. & ANN, D. K. 2014. Arginine starvation impairs mitochondrial respiratory function in ASS1-deficient breast cancer cells. *Sci Signal*, 7, ra31.
- QIU, P., SIMONDS, E. F., BENDALL, S. C., GIBBS, K. D., JR., BRUGGNER, R. V., LINDERMAN, M. D., SACHS, K., NOLAN, G. P. & PLEVITIS, S. K. 2011. Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. *Nat Biotechnol*, 29, 886-91.
- QIU, X., MAO, Q., TANG, Y., WANG, L., CHAWLA, R., PLINER, H. A. & TRAPNELL, C. 2017. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods*, 14, 979-982.
- RACLE, J., DE JONGE, K., BAUMGAERTNER, P., SPEISER, D. E. & GFELLER, D. 2017. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *Elife*, 6.
- RAPAPORT, F., KHANIN, R., LIANG, Y., PIRUN, M., KREK, A., ZUMBO, P., MASON, C. E., SOCCI, N. D. & BETEL, D. 2013. Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. *Genome Biol*, 14, R95.
- RAUD, B., ROY, D. G., DIVAKARUNI, A. S., TARASENKO, T. N., FRANKE, R., MA, E. H., SAMBORSKA, B., HSIEH, W. Y., WONG, A. H., STÜVE, P., ARNOLD-SCHRAUF, C., GUDERIAN, M., LOCHNER, M., RAMPERTAAP, S., ROMITO, K., MONSALE, J., BRÖNSTRUP, M., BENSINGER, S. J., MURPHY, A. N., MCGUIRE, P. J., JONES, R. G., SPARWASSER, T. & BEROD, L. 2018. Etomoxir Actions on Regulatory and Memory T Cells Are Independent of Cpt1a-Mediated Fatty Acid Oxidation. *Cell Metabolism*, 28, 504-515.e7.
- REDMOND, D., PORAN, A. & ELEMENTO, O. 2016. Single-cell TCRseq: paired recovery of entire T-cell alpha and beta chain transcripts in T-cell receptors from single-cell RNAseq. *Genome Med*, 8, 80.
- REIMAND, J., ISSERLIN, R., VOISIN, V., KUCERA, M., TANNUS-LOPES, C., ROSTAMIANFAR, A., WADI, L., MEYER, M., WONG, J., XU, C., MERICO, D. & BADER, G. D. 2019. Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nature Protocols*, 14, 482-517.
- REIMAND, J., KULL, M., PETERSON, H., HANSEN, J. & VILO, J. 2007. g:Profiler--a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Res*, 35, W193-200.
- RITCHIE, M. E., PHIPSON, B., WU, D., HU, Y., LAW, C. W., SHI, W. & SMYTH, G. K. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*, 43, e47.
- RIZZETTO, S., KOPPSTEIN, D. N. P., SAMIR, J., SINGH, M., REED, J. H., CAI, C. H., LLOYD, A. R., ELTAHLA, A. A., GOODNOW, C. C. & LUCIANI, F. 2018. B-cell receptor reconstruction from single-cell RNA-seq with VDJ Puzzle. *Bioinformatics*, 34, 2846-2847.
- ROBINSON, M. D., MCCARTHY, D. J. & SMYTH, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26, 139-40.
- RODRIGUEZ, A. E., DUCKER, G. S., BILLINGHAM, L. K., MARTINEZ, C. A., MAINOLFI, N., SURI, V., FRIEDMAN, A., MANFREDI, M. G., WEINBERG, S. E., RABINOWITZ, J. D. & CHANDEL, N. S. 2019. Serine Metabolism Supports Macrophage IL-1 β Production. *Cell Metabolism*, 29, 1003-1011.e4.
- SABHARWAL, S. S., ROSEN, D. B., GREIN, J., TEDESCO, D., JOYCE-SHAIKH, B., UEDA, R., SEMANA, M., BAUER, M., BANG, K., STEVENSON, C., CUA, D. J. & ZÚÑIGA, L. A. 2018. GITR Agonism Enhances Cellular Metabolism to Support CD8(+) T-cell Proliferation and Effector Cytokine Production in a Mouse Tumor Model. *Cancer Immunol Res*, 6, 1199-1211.
- SAG, D., CEKIC, C., WU, R., LINDEN, J. & HEDRICK, C. C. 2015. The cholesterol transporter ABCG1 links cholesterol homeostasis and tumour immunity. *Nat Commun*, 6, 6354.

- SAHIN, U., DERHOVANESEAN, E., MILLER, M., KLOKE, B. P., SIMON, P., LÖWER, M., BUKUR, V., TADMOR, A. D., LUXEMBURGER, U., SCHRÖRS, B., OMOKOKO, T., VORMEHR, M., ALBRECHT, C., PARUZYSKI, A., KUHN, A. N., BUCK, J., HEESCH, S., SCHREEB, K. H., MÜLLER, F., ORTSEIFER, I., VOGLER, I., GODEHARDT, E., ATTIG, S., RAE, R., BREITKREUZ, A., TOLLIVER, C., SUCHAN, M., MARTIC, G., HOHBERGER, A., SORN, P., DIEKMANN, J., CIESLA, J., WAKSMANN, O., BRÜCK, A. K., WITT, M., ZILLGEN, M., ROTHERMEL, A., KASEMANN, B., LANGER, D., BOLTE, S., DIKEN, M., KREITER, S., NEMECEK, R., GEBHARDT, C., GRABBE, S., HÖLLER, C., UTIKAL, J., HUBER, C., LOQUAI, C. & TÜRECI, Ö. 2017. Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. *Nature*, 547, 222-226.
- SAMUSIK, N., GOOD, Z., SPITZER, M. H., DAVIS, K. L. & NOLAN, G. P. 2016. Automated mapping of phenotype space with single-cell data. *Nat Methods*, 13, 493-6.
- SCHAPIRO, D., JACKSON, H. W., RAGHURAMAN, S., FISCHER, J. R., ZANOTELLI, V. R. T., SCHULZ, D., GIESEN, C., CATENA, R., VARGA, Z. & BODENMILLER, B. 2017. histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nat Methods*, 14, 873-876.
- SCHARPING, N. E., MENK, A. V., MORECI, R. S., WHETSTONE, R. D., DADEY, R. E., WATKINS, S. C., FERRIS, R. L. & DELGOFFE, G. M. 2016. The Tumor Microenvironment Represses T Cell Mitochondrial Biogenesis to Drive Intratumoral T Cell Metabolic Insufficiency and Dysfunction. *Immunity*, 45, 374-88.
- SCHNEIDER, C. A., RASBAND, W. S. & ELICEIRI, K. W. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*, 9, 671-5.
- SCHUBERT, O. T., RÖST, H. L., COLLINS, B. C., ROSENBERGER, G. & AEBERSOLD, R. 2017. Quantitative proteomics: challenges and opportunities in basic and applied research. *Nat Protoc*, 12, 1289-1294.
- SENA, L. A., LI, S., JAIRAMAN, A., PRAKRIYA, M., EZPONDA, T., HILDEMAN, D. A., WANG, C. R., SCHUMACKER, P. T., LICHT, J. D., PERLMAN, H., BRYCE, P. J. & CHANDEL, N. S. 2013. Mitochondria are required for antigen-specific T cell activation through reactive oxygen species signaling. *Immunity*, 38, 225-36.
- SETTY, M., KISELIOVAS, V., LEVINE, J., GAYOSO, A., MAZUTIS, L. & PE'ER, D. 2019. Characterization of cell fate probabilities in single-cell data with Palantir. *Nature Biotechnology*, 37, 451-460.
- SHANKER, A., DE AQUINO, M. T. P., HODO, T. & UZHACHENKO, R. 2018. Glutamate receptors provide costimulatory signals to improve T cell immune response. *The Journal of Immunology*, 200, 47.24-47.24.
- SHANNON, P., MARKIEL, A., OZIER, O., BALIGA, N. S., WANG, J. T., RAMAGE, D., AMIN, N., SCHWIKOWSKI, B. & IDEKER, T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, 13, 2498-504.
- SHAO, W., PEDRIOLI, P. G. A., WOLSKI, W., SCURTESCU, C., SCHMID, E., VIZCAÍNO, J. A., COURCELLES, M., SCHUSTER, H., KOWALEWSKI, D., MARINO, F., ARLEHAMN, C. S. L., VAUGHAN, K., PETERS, B., SETTE, A., OTTENHOFF, T. H. M., MEIJGAARDEN, K. E., NIEUWENHUIZEN, N., KAUFMANN, S. H. E., SCHLAPBACH, R., CASTLE, J. C., NESVIZHSKII, A. I., NIELSEN, M., DEUTSCH, E. W., CAMPBELL, D. S., MORITZ, R. L., ZUBAREV, R. A., YTTERBERG, A. J., PURCELL, A. W., MARCILLA, M., PARADELA, A., WANG, Q., COSTELLO, C. E., TERNETTE, N., VAN VELEN, P. A., VAN ELS, C., HECK, A. J. R., DE SOUZA, G. A., SOLLID, L. M., ADMON, A., STEVANOVIC, S., RAMMENSEE, H. G., THIBAUT, P., PERREAULT, C., BASSANI-STERBERG, M., AEBERSOLD, R. & CARON, E. 2018. The SystemMHC Atlas project. *Nucleic Acids Res*, 46, D1237-d1247.

- SHARMA, S., YANG, S. C., ZHU, L., RECKAMP, K., GARDNER, B., BARATELLI, F., HUANG, M., BATRA, R. K. & DUBINETT, S. M. 2005. Tumor cyclooxygenase-2/prostaglandin E2-dependent promotion of FOXP3 expression and CD4⁺ CD25⁺ T regulatory cell activities in lung cancer. *Cancer Res*, 65, 5211-20.
- SHEKHAR, K., BRODIN, P., DAVIS, M. M. & CHAKRABORTY, A. K. 2014. Automatic Classification of Cellular Expression by Nonlinear Stochastic Embedding (ACCENSE). *Proc Natl Acad Sci U S A*, 111, 202-7.
- SHIME, H., YABU, M., AKAZAWA, T., KODAMA, K., MATSUMOTO, M., SEYA, T. & INOUE, N. 2008. Tumor-Secreted Lactic Acid Promotes IL-23/IL-17 Proinflammatory Pathway. *The Journal of Immunology*, 180, 7175-7183.
- SHUKLA, S. A., ROONEY, M. S., RAJASAGI, M., TIAO, G., DIXON, P. M., LAWRENCE, M. S., STEVENS, J., LANE, W. J., DELLAGATTA, J. L., STEELMAN, S., SOUGNEZ, C., CIBULSKIS, K., KIEZUN, A., HACOEN, N., BRUSIC, V., WU, C. J. & GETZ, G. 2015. Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. *Nat Biotechnol*, 33, 1152-8.
- SISKA, P. J., BECKERMANN, K. E., MASON, F. M., ANDREJEVA, G., GREENPLATE, A. R., SENDOR, A. B., CHIANG, Y. J., CORONA, A. L., GEMTA, L. F., VINCENT, B. G., WANG, R. C., KIM, B., HONG, J., CHEN, C. L., BULLOCK, T. N., IRISH, J. M., RATHMELL, W. K. & RATHMELL, J. C. 2017. Mitochondrial dysregulation and glycolytic insufficiency functionally impair CD8 T cells infiltrating human renal cell carcinoma. *JCI Insight*, 2.
- SMYTH, G. K. 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*, 3, Article3.
- SNIJDEWINT, F. G., KALIŃSKI, P., WIERENGA, E. A., BOS, J. D. & KAPSENBERG, M. L. 1993. Prostaglandin E2 differentially modulates cytokine secretion profiles of human T helper lymphocytes. *J Immunol*, 150, 5321-9.
- SOMMER, C., STRAEHLE, C., KÖTHE, U. & HAMPRECHT, F. A. Ilastik: Interactive learning and segmentation toolkit. 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 30 March-2 April 2011. 230-233.
- SORRENTINO, C., MIELE, L., PORTA, A., PINTO, A. & MORELLO, S. 2015. Myeloid-derived suppressor cells contribute to A2B adenosine receptor-induced VEGF production and angiogenesis in a mouse melanoma model. *Oncotarget*, 6, 27478-89.
- SPITZER, M. H., GHERARDINI, P. F., FRAGIADAKIS, G. K., BHATTACHARYA, N., YUAN, R. T., HOTSON, A. N., FINCK, R., CARMI, Y., ZUNDER, E. R., FANTL, W. J., BENDALL, S. C., ENGLEMAN, E. G. & NOLAN, G. P. 2015. IMMUNOLOGY. An interactive reference framework for modeling a dynamic immune system. *Science*, 349, 1259425.
- STEGGERDA, S. M., BENNETT, M. K., CHEN, J., EMBERLEY, E., HUANG, T., JANES, J. R., LI, W., MACKINNON, A. L., MAKKOUK, A., MARGUIER, G., MURRAY, P. J., NEOU, S., PAN, A., PARLATI, F., RODRIGUEZ, M. L. M., VAN DE VELDE, L. A., WANG, T., WORKS, M., ZHANG, J., ZHANG, W. & GROSS, M. I. 2017. Inhibition of arginase by CB-1158 blocks myeloid cell-mediated immune suppression in the tumor microenvironment. *J Immunother Cancer*, 5, 101.
- STUART, T., BUTLER, A., HOFFMAN, P., HAFEMEISTER, C., PAPALEXI, E., MAUCK, W. M., 3RD, HAO, Y., STOECKIUS, M., SMIBERT, P. & SATIJA, R. 2019. Comprehensive Integration of Single-Cell Data. *Cell*, 177, 1888-1902.e21.
- STUBBINGTON, M. J. T., LÖNNBERG, T., PROSERPIO, V., CLARE, S., SPEAK, A. O., DOUGAN, G. & TEICHMANN, S. A. 2016. T cell fate and clonality inference from single-cell transcriptomes. *Nat Methods*, 13, 329-332.

- STURM, G., FINOTELLO, F., PETITPREZ, F., ZHANG, J. D., BAUMBACH, J., FRIDMAN, W. H., LIST, M. & ANEICHYK, T. 2019. Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. *Bioinformatics*, 35, i436-i445.
- SUBRAMANIAN, A., TAMAYO, P., MOOTHA, V. K., MUKHERJEE, S., EBERT, B. L., GILLETTE, M. A., PAULOVICH, A., POMEROY, S. L., GOLUB, T. R., LANDER, E. S. & MESIROV, J. P. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*, 102, 15545-50.
- SUKUMAR, M., LIU, J., JI, Y., SUBRAMANIAN, M., CROMPTON, J. G., YU, Z., ROYCHOUDHURI, R., PALMER, D. C., MURANSKI, P., KAROLY, E. D., MOHNEY, R. P., KLEBANOFF, C. A., LAL, A., FINKEL, T., RESTIFO, N. P. & GATTINONI, L. 2013. Inhibiting glycolytic metabolism enhances CD8+ T cell memory and antitumor function. *J Clin Invest*, 123, 4479-88.
- SUN, X., WU, Y., GAO, W., ENJOJI, K., CSIZMADIA, E., MÜLLER, C. E., MURAKAMI, T. & ROBSON, S. C. 2010. CD39/ENTPD1 expression by CD4+Foxp3+ regulatory T cells promotes hepatic metastatic tumor growth in mice. *Gastroenterology*, 139, 1030-40.
- SWAIN, A., BAMBOUSKOVA, M., KIM, H., ANDHEY, P. S., DUNCAN, D., AUCLAIR, K., CHUBUKOV, V., SIMONS, D. M., RODDY, T. P., STEWART, K. M. & ARTYOMOV, M. N. 2020. Comparative evaluation of itaconate and its derivatives reveals divergent inflammasome and type I interferon regulation in macrophages. *Nature Metabolism*, 2, 594-602.
- SZKLARCZYK, D., FRANCESCHINI, A., WYDER, S., FORSLUND, K., HELLER, D., HUERTA-CEPAS, J., SIMONOVIC, M., ROTH, A., SANTOS, A., TSAFOU, K. P., KUHN, M., BORK, P., JENSEN, L. J. & VON MERING, C. 2015. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res*, 43, D447-52.
- SZOLEK, A., SCHUBERT, B., MOHR, C., STURM, M., FELDHAHN, M. & KOHLBACHER, O. 2014. OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics*, 30, 3310-6.
- TANNAHILL, G. M., CURTIS, A. M., ADAMIK, J., PALSSON-MCDERMOTT, E. M., MCGETTRICK, A. F., GOEL, G., FREZZA, C., BERNARD, N. J., KELLY, B., FOLEY, N. H., ZHENG, L., GARDET, A., TONG, Z., JANY, S. S., CORR, S. C., HANEKLAUS, M., CAFFREY, B. E., PIERCE, K., WALMSLEY, S., BEASLEY, F. C., CUMMINS, E., NIZET, V., WHYTE, M., TAYLOR, C. T., LIN, H., MASTERS, S. L., GOTTLIEB, E., KELLY, V. P., CLISH, C., AURON, P. E., XAVIER, R. J. & O'NEILL, L. A. J. 2013. Succinate is an inflammatory signal that induces IL-1 β through HIF-1 α . *Nature*, 496, 238-242.
- TANNIR, N. M., MOTZER, R. J., AGARWAL, N., LIU, P.-Y., WHITING, S. H., O'KEEFFE, B., TRAN, X., FIJI, G. P. & ESCUDIER, B. 2018. CANTATA: A randomized phase 2 study of CB-839 in combination with cabozantinib vs. placebo with cabozantinib in patients with advanced/metastatic renal cell carcinoma. *Journal of Clinical Oncology*, 36, TPS4601-TPS4601.
- TAPPEINER, E., FINOTELLO, F., CHAROENTONG, P., MAYER, C., RIEDER, D. & TRAJANOSKI, Z. 2017. TIminer: NGS data mining pipeline for cancer immunology and immunotherapy. *Bioinformatics*, 33, 3140-3141.
- THOMMEN, D. S., KOELZER, V. H., HERZIG, P., ROLLER, A., TREFNY, M., DIMELOE, S., KIIALAINEN, A., HANHART, J., SCHILL, C., HESS, C., SAVIC PRINCE, S., WIESE, M., LARDINOIS, D., HO, P. C., KLEIN, C., KARANIKAS, V., MERTZ, K. D., SCHUMACHER, T. N. & ZIPPELIUS, A. 2018. A transcriptionally and functionally distinct PD-1(+) CD8(+) T cell pool with predictive potential in non-small-cell lung cancer treated with PD-1 blockade. *Nat Med*, 24, 994-1004.

- TIAN, L., DONG, X., FREYTAG, S., KA, L. C., SU, S., JALALABADI, A., AMANN-ZALCENSTEIN, D., WEBER, T. S., SEIDI, A., JABBARI, J. S., NAIK, S. H. & RITCHIE, M. E. 2019. Benchmarking single cell RNA-sequencing analysis pipelines using mixture control experiments. *Nat Methods*, 16, 479-487.
- TIROSH, I., IZAR, B., PRAKADAN, S. M., WADSWORTH, M. H., 2ND, TREACY, D., TROMBETTA, J. J., ROTEM, A., RODMAN, C., LIAN, C., MURPHY, G., FALLAHI-SICHANI, M., DUTTON-REGESTER, K., LIN, J. R., COHEN, O., SHAH, P., LU, D., GENSHAFT, A. S., HUGHES, T. K., ZIEGLER, C. G., KAZER, S. W., GAILLARD, A., KOLB, K. E., VILLANI, A. C., JOHANNESSEN, C. M., ANDREEV, A. Y., VAN ALLEN, E. M., BERTAGNOLLI, M., SORGER, P. K., SULLIVAN, R. J., FLAHERTY, K. T., FREDERICK, D. T., JANÉ-VALBUENA, J., YOON, C. H., ROZENBLATT-ROSEN, O., SHALEK, A. K., REGEV, A. & GARRAWAY, L. A. 2016. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*, 352, 189-96.
- TRAPNELL, C., WILLIAMS, B. A., PERTEA, G., MORTAZAVI, A., KWAN, G., VAN BAREN, M. J., SALZBERG, S. L., WOLD, B. J. & PACHTER, L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*, 28, 511-5.
- TSAI, H. J., JIANG, S. S., HUNG, W. C., BORTHAKUR, G., LIN, S. F., PEMMARAJU, N., JABBOUR, E., BOMALASKI, J. S., CHEN, Y. P., HSIAO, H. H., WANG, M. C., KUO, C. Y., CHANG, H., YEH, S. P., CORTES, J., CHEN, L. T. & CHEN, T. Y. 2017. A Phase II Study of Arginine Deiminase (ADI-PEG20) in Relapsed/Refractory or Poor-Risk Acute Myeloid Leukemia Patients. *Sci Rep*, 7, 11253.
- TSURUTANI, N., MITTAL, P., ST ROSE, M. C., NGOI, S. M., SVEDOVA, J., MENORET, A., TREADWAY, F. B., LAUBENBACHER, R., SUÁREZ-RAMÍREZ, J. E., CAULEY, L. S., ADLER, A. J. & VELLA, A. T. 2016. Costimulation Endows Immunotherapeutic CD8 T Cells with IL-36 Responsiveness during Aerobic Glycolysis. *J Immunol*, 196, 124-34.
- TURCOTTE, M., SPRING, K., POMMEY, S., CHOUINARD, G., COUSINEAU, I., GEORGE, J., CHEN, G. M., GENDOO, D. M., HAIBE-KAINS, B., KARN, T., RAHIMI, K., LE PAGE, C., PROVENCHER, D., MESMASSON, A. M. & STAGG, J. 2015. CD73 is associated with poor prognosis in high-grade serous ovarian cancer. *Cancer Res*, 75, 4494-503.
- UPADHYAY, A. A., KAUFFMAN, R. C., WOLABAUGH, A. N., CHO, A., PATEL, N. B., REISS, S. M., HAVENAR-DAUGHTON, C., DAWOUD, R. A., THARP, G. K., SANZ, I., PULENDRAN, B., CROTTY, S., LEE, F. E., WRAMMERT, J. & BOSINGER, S. E. 2018. BALDR: a computational pipeline for paired heavy and light chain immunoglobulin reconstruction in single-cell RNA-seq data. *Genome Med*, 10, 20.
- VAN DEN BOSSCHE, J., LAMERS, W. H., KOEHLER, E. S., GEUNS, J. M. C., ALHONEN, L., UIMARI, A., PIRNES-KARHU, S., VAN OVERMEIRE, E., MORIAS, Y., BRYNS, L., VEREECKE, L., DE BAETSELIER, P. & VAN GINDERACHTER, J. A. 2012. Pivotal Advance: Arginase-1-independent polyamine production stimulates the expression of IL-4-induced alternatively activated macrophage markers while inhibiting LPS-induced expression of inflammatory genes. *Journal of Leukocyte Biology*, 91, 685-699.
- VAN DEN BOSSCHE, J. & SARABER, D. L. 2018. Metabolic regulation of macrophages in tissues. *Cellular Immunology*, 330, 54-59.
- VAN DEN BOSSCHE, J. & VAN DER WINDT, G. J. W. 2018. Fatty Acid Oxidation in Macrophages and T Cells: Time for Reassessment? *Cell Metabolism*, 28, 538-540.
- VAN DER WINDT, G. J., EVERTS, B., CHANG, C. H., CURTIS, J. D., FREITAS, T. C., AMIEL, E., PEARCE, E. J. & PEARCE, E. L. 2012. Mitochondrial respiratory capacity is a critical regulator of CD8+ T cell memory development. *Immunity*, 36, 68-78.

- VAN DER WINDT, G. J. W., CHANG, C.-H. & PEARCE, E. L. 2016. Measuring Bioenergetics in T Cells Using a Seahorse Extracellular Flux Analyzer. *Current Protocols in Immunology*, 113, 3.16B.1-3.16B.14.
- VAN DER WINDT, G. J. W., O'SULLIVAN, D., EVERTS, B., HUANG, S. C.-C., BUCK, M. D., CURTIS, J. D., CHANG, C.-H., SMITH, A. M., AI, T., FAUBERT, B., JONES, R. G., PEARCE, E. J. & PEARCE, E. L. 2013. CD8 memory T cells have a bioenergetic advantage that underlies their rapid recall ability. *Proceedings of the National Academy of Sciences*, 110, 14336-14341.
- VAN GASSEN, S., CALLEBAUT, B., VAN HELDEN, M. J., LAMBRECHT, B. N., DEMEESTER, P., DHAENE, T. & SAEYS, Y. 2015. FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry A*, 87, 636-45.
- VAN UNEN, V., HÖLLT, T., PEZZOTTI, N., LI, N., REINDERS, M. J. T., EISEMANN, E., KONING, F., VILANOVA, A. & LELIEVELDT, B. P. F. 2017. Visual analysis of mass cytometry data by hierarchical stochastic neighbour embedding reveals rare cell types. *Nat Commun*, 8, 1740.
- VAN VALEN, D. A., KUDO, T., LANE, K. M., MACKLIN, D. N., QUACH, N. T., DEFELICE, M. M., MAAYAN, I., TANOUCHI, Y., ASHLEY, E. A. & COVERT, M. W. 2016. Deep Learning Automates the Quantitative Analysis of Individual Cells in Live-Cell Imaging Experiments. *PLoS Comput Biol*, 12, e1005177.
- VATS, D., MUKUNDAN, L., ODEGAARD, J. I., ZHANG, L., SMITH, K. L., MOREL, C. R., GREAVES, D. R., MURRAY, P. J. & CHAWLA, A. 2006. Oxidative metabolism and PGC-1 β attenuate macrophage-mediated inflammation. *Cell Metabolism*, 4, 13-24.
- VILLANI, A.-C., SATIJA, R., REYNOLDS, G., SARKIZOVA, S., SHEKHAR, K., FLETCHER, J., GRIESBECK, M., BUTLER, A., ZHENG, S., LAZO, S., JARDINE, L., DIXON, D., STEPHENSON, E., NILSSON, E., GRUNDBERG, I., MCDONALD, D., FILBY, A., LI, W., DE JAGER, P. L., ROZENBLATT-ROSEN, O., LANE, A. A., HANIFFA, M., REGEV, A. & HACOEN, N. 2017. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science*, 356, eaah4573.
- VITA, R., OVERTON, J. A., GREENBAUM, J. A., PONOMARENKO, J., CLARK, J. D., CANTRELL, J. R., WHEELER, D. K., GABBARD, J. L., HIX, D., SETTE, A. & PETERS, B. 2015. The immune epitope database (IEDB) 3.0. *Nucleic Acids Res*, 43, D405-12.
- VIZCAÍNO, J. A., CSORDAS, A., DEL-TORO, N., DIANES, J. A., GRISS, J., LAVIDAS, I., MAYER, G., PEREZ-RIVEROL, Y., REISINGER, F., TERNENT, T., XU, Q. W., WANG, R. & HERMIAKOB, H. 2016. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res*, 44, 11033.
- WANG, B., ZHU, J., PIERSON, E., RAMAZZOTTI, D. & BATZOGLOU, S. 2017. Visualization and analysis of single-cell RNA-seq data by kernel-based similarity learning. *Nat Methods*, 14, 414-416.
- WANG, R., DILLON, CHRISTOPHER P., SHI, LEWIS Z., MILASTA, S., CARTER, R., FINKELSTEIN, D., MCCORMICK, LAURA L., FITZGERALD, P., CHI, H., MUNGER, J. & GREEN, DOUGLAS R. 2011. The Transcription Factor Myc Controls Metabolic Reprogramming upon T Lymphocyte Activation. *Immunity*, 35, 871-882.
- WARDE-FARLEY, D., DONALDSON, S. L., COMES, O., ZUBERI, K., BADRAWI, R., CHAO, P., FRANZ, M., GROUIOS, C., KAZI, F., LOPES, C. T., MAITLAND, A., MOSTAFAVI, S., MONTOJO, J., SHAO, Q., WRIGHT, G., BADER, G. D. & MORRIS, Q. 2010. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res*, 38, W214-20.
- WEIDE, B., MARTENS, A., HASSEL, J. C., BERKING, C., POSTOW, M. A., BISSCHOP, K., SIMEONE, E., MANGANA, J., SCHILLING, B., DI GIACOMO, A. M., BRENNER, N., KÄHLER, K., HEINZERLING, L., GUTZMER, R., BENDER, A., GEBHARDT, C., ROMANO, E., MEIER, F., MARTUS, P., MAIO, M., BLANK, C., SCHADENDORF, D., DUMMER, R., ASCIERTO, P. A., HOSPERS, G., GARBE, C. & WOLCHOK, J. D.

2016. Baseline Biomarkers for Outcome of Melanoma Patients Treated with Pembrolizumab. *Clinical Cancer Research*, 22, 5487-5496.
- WOLF, F. A., ANGERER, P. & THEIS, F. J. 2018. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol*, 19, 15.
- WOLF, F. A., HAMEY, F. K., PLASS, M., SOLANA, J., DAHLIN, J. S., GÖTTGENS, B., RAJEWSKY, N., SIMON, L. & THEIS, F. J. 2019. PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol*, 20, 59.
- XIANG, Y., STINE, Z. E., XIA, J., LU, Y., O'CONNOR, R. S., ALTMAN, B. J., HSIEH, A. L., GOUW, A. M., THOMAS, A. G., GAO, P., SUN, L., SONG, L., YAN, B., SLUSHER, B. S., ZHUO, J., OOI, L. L., LEE, C. G. L., MANCUSO, A., MCCALLION, A. S., LE, A., MILONE, M. C., RAYPORT, S., FELSHER, D. W. & DANG, C. V. 2015. Targeted inhibition of tumor-specific glutaminase diminishes cell-autonomous tumorigenesis. *Journal of Clinical Investigation*, 125, 2293-2306.
- XIE, C., YEO, Z. X., WONG, M., PIPER, J., LONG, T., KIRKNESS, E. F., BIGGS, W. H., BLOOM, K., SPELLMAN, S., VIERRA-GREEN, C., BRADY, C., SCHEUERMANN, R. H., TELENTI, A., HOWARD, S., BREWERTON, S., TURPAZ, Y. & VENTER, J. C. 2017. Fast and accurate HLA typing from short-read next-generation sequence data with xHLA. *Proc Natl Acad Sci U S A*, 114, 8059-8064.
- XUE, J., SCHMIDT, S. V., SANDER, J., DRAFFEHN, A., KREBS, W., QUESTER, I., DE NARDO, D., GOHEL, T. D., EMDE, M., SCHMIDLEITHNER, L., GANESAN, H., NINO-CASTRO, A., MALLMANN, M. R., LABZIN, L., THEIS, H., KRAUT, M., BEYER, M., LATZ, E., FREEMAN, T. C., ULAS, T. & SCHULTZE, J. L. 2014. Transcriptome-based network analysis reveals a spectrum model of human macrophage activation. *Immunity*, 40, 274-88.
- YAN, H., PARSONS, D. W., JIN, G., MCLENDON, R., RASHEED, B. A., YUAN, W., KOS, I., BATINIC-HABERLE, I., JONES, S., RIGGINS, G. J., FRIEDMAN, H., FRIEDMAN, A., REARDON, D., HERNDON, J., KINZLER, K. W., VELCULESCU, V. E., VOGELSTEIN, B. & BIGNER, D. D. 2009. IDH1 and IDH2 mutations in gliomas. *N Engl J Med*, 360, 765-73.
- YANG, H. & WANG, K. 2015. Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. *Nat Protoc*, 10, 1556-66.
- YANG, M., MA, C., LIU, S., SHAO, Q., GAO, W., SONG, B., SUN, J., XIE, Q., ZHANG, Y., FENG, A., LIU, Y., HU, W. & QU, X. 2010. HIF - dependent induction of adenosine receptor A2b skews human dendritic cells to a Th2 - stimulating phenotype under hypoxia. *Immunology & Cell Biology*, 88, 165-171.
- YANG, W., BAI, Y., XIONG, Y., ZHANG, J., CHEN, S., ZHENG, X., MENG, X., LI, L., WANG, J., XU, C., YAN, C., WANG, L., CHANG, C. C., CHANG, T. Y., ZHANG, T., ZHOU, P., SONG, B. L., LIU, W., SUN, S. C., LIU, X., LI, B. L. & XU, C. 2016. Potentiating the antitumour response of CD8(+) T cells by modulating cholesterol metabolism. *Nature*, 531, 651-5.
- YANG, W., LEE, K. W., SRIVASTAVA, R. M., KUO, F., KRISHNA, C., CHOWELL, D., MAKAROV, V., HOEN, D., DALIN, M. G., WEXLER, L., GHOSSEIN, R., KATABI, N., NADEEM, Z., COHEN, M. A., TIAN, S. K., ROBINE, N., ARORA, K., GEIGER, H., AGIUS, P., BOUVIER, N., HUBERMAN, K., VANNESS, K., HAVEL, J. J., SIMS, J. S., SAMSTEIN, R. M., MANDAL, R., TEPE, J., GANLY, I., HO, A. L., RIAZ, N., WONG, R. J., SHUKLA, N., CHAN, T. A. & MORRIS, L. G. T. 2019. Immunogenic neoantigens derived from gene fusions stimulate T cell responses. *Nat Med*, 25, 767-775.
- YEN, M. C., LIN, C. C., CHEN, Y. L., HUANG, S. S., YANG, H. J., CHANG, C. P., LEI, H. Y. & LAI, M. D. 2009. A novel cancer therapy by skin delivery of indoleamine 2,3-dioxygenase siRNA. *Clin Cancer Res*, 15, 641-9.

- YORK, A. G., WILLIAMS, K. J., ARGUS, J. P., ZHOU, Q. D., BRAR, G., VERGNES, L., GRAY, E. E., ZHEN, A., WU, N. C., YAMADA, D. H., CUNNINGHAM, C. R., TARLING, E. J., WILKS, M. Q., CASERO, D., GRAY, D. H., YU, A. K., WANG, E. S., BROOKS, D. G., SUN, R., KITCHEN, S. G., WU, T. T., REUE, K., STETSON, D. B. & BENSINGER, S. J. 2015. Limiting Cholesterol Biosynthetic Flux Spontaneously Engages Type I IFN Signaling. *Cell*, 163, 1716-29.
- ZAITSEV, K., BAMBOUSKOVA, M., SWAIN, A. & ARTYOMOV, M. N. 2019. Complete deconvolution of cellular mixtures based on linearity of transcriptional signatures. *Nat Commun*, 10, 2209.
- ZECH, T., EJSING, C. S., GAUS, K., DE WET, B., SHEVCHENKO, A., SIMONS, K. & HARDER, T. 2009. Accumulation of raft lipids in T-cell plasma membrane domains engaged in TCR signalling. *Embo j*, 28, 466-76.
- ZELENAY, S., VAN DER VEEN, A. G., BÖTTCHER, J. P., SNELGROVE, K. J., ROGERS, N., ACTON, S. E., CHAKRAVARTY, P., GIROTTI, M. R., MARAIS, R., QUEZADA, S. A., SAHAI, E. & REIS E SOUSA, C. 2015. Cyclooxygenase-Dependent Tumor Growth through Evasion of Immunity. *Cell*, 162, 1257-70.
- ZENG, B., GLICKSBERG, B. S., NEWBURY, P., CHEKALIN, E., XING, J., LIU, K., WEN, A., CHOW, C. & CHEN, B. 2021. OCTAD: an open workspace for virtually screening therapeutics targeting precise cancer patient groups using gene expression features. *Nature Protocols*, 16, 728-753.
- ZENG, H., COHEN, S., GUY, C., SHRESTHA, S., NEALE, G., BROWN, S. A., CLOER, C., KISHTON, R. J., GAO, X., YOUNGBLOOD, B., DO, M., LI, M. O., LOCASALE, J. W., RATHMELL, J. C. & CHI, H. 2016. mTORC1 and mTORC2 Kinase Signaling and Glucose Metabolism Drive Follicular Helper T Cell Differentiation. *Immunity*, 45, 540-554.
- ZENG, W. Z. D., GLICKSBERG, B. S., LI, Y. & CHEN, B. 2019. Selecting precise reference normal tissue samples for cancer research using a deep learning approach. *BMC Med Genomics*, 12, 21.
- ZHANG, L., TSCHUMI, B. O., LOPEZ-MEJIA, I. C., OBERLE, S. G., MEYER, M., SAMSON, G., RÜEGG, M. A., HALL, M. N., FAJAS, L., ZEHN, D., MACH, J. P., DONDA, A. & ROMERO, P. 2016. Mammalian Target of Rapamycin Complex 2 Controls CD8 T Cell Memory Differentiation in a Foxo1-Dependent Manner. *Cell Rep*, 14, 1206-1217.
- ZHANG, Y., DAQUINAG, A. C., AMAYA-MANZANARES, F., SIRIN, O., TSENG, C. & KOLONIN, M. G. 2012. Stromal progenitor cells from endogenous adipose tissue contribute to pericytes and adipocytes that populate the tumor microenvironment. *Cancer Res*, 72, 5198-208.
- ZHANG, Y., KURUPATI, R., LIU, L., ZHOU, X. Y., ZHANG, G., HUDAIHED, A., FILISIO, F., GILES-DAVIS, W., XU, X., KARAKOUSIS, G. C., SCHUCHTER, L. M., XU, W., AMARAVADI, R., XIAO, M., SADEK, N., KREPLER, C., HERLYN, M., FREEMAN, G. J., RABINOWITZ, J. D. & ERTL, H. C. J. 2017. Enhancing CD8(+) T Cell Fatty Acid Catabolism within a Metabolically Challenging Tumor Microenvironment Increases the Efficacy of Melanoma Immunotherapy. *Cancer Cell*, 32, 377-391.e9.
- ZHAO, W. & SHER, X. 2018. Systematically benchmarking peptide-MHC binding predictors: From synthetic to naturally processed epitopes. *PLoS Comput Biol*, 14, e1006457.
- ZHENG, C., ZHENG, L., YOO, J. K., GUO, H., ZHANG, Y., GUO, X., KANG, B., HU, R., HUANG, J. Y., ZHANG, Q., LIU, Z., DONG, M., HU, X., OUYANG, W., PENG, J. & ZHANG, Z. 2017. Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell*, 169, 1342-1356.e16.
- ZHENG, W., O'HEAR, C. E., ALLI, R., BASHAM, J. H., ABDELSAMED, H. A., PALMER, L. E., JONES, L. L., YOUNGBLOOD, B. & GEIGER, T. L. 2018. PI3K orchestration of the in vivo persistence of chimeric antigen receptor-modified T cells. *Leukemia*, 32, 1157-1167.

- ZHENG, X., KOROPATNICK, J., CHEN, D., VELENOSI, T., LING, H., ZHANG, X., JIANG, N., NAVARRO, B., ICHIM, T. E., URQUHART, B. & MIN, W. 2013. Silencing IDO in dendritic cells: a novel approach to enhance cancer immunotherapy in a murine breast cancer model. *Int J Cancer*, 132, 967-77.
- ZHI, X., WANG, Y., ZHOU, X., YU, J., JIAN, R., TANG, S., YIN, L. & ZHOU, P. 2010. RNAi-mediated CD73 suppression induces apoptosis and cell-cycle arrest in human breast cancer cells. *Cancer Sci*, 101, 2561-9.
- ZHU, X., WOLFGRUBER, T. K., TASATO, A., ARISDAKESSIAN, C., GARMIRE, D. G. & GARMIRE, L. X. 2017. Granatum: a graphical single-cell RNA-Seq analysis pipeline for genomics scientists. *Genome Med*, 9, 108.