A novel bioinformatics pipeline for the identification of immune inhibitory receptors as potential therapeutic targets

Akashdip Singh[1,2], Alberto Miranda Bedate*[1,3], Helen J. von Richthofen*[1,2], Saskia V. Vijver*[1,2], Michiel van der Vlist[1,2], Raphael Kuhn[4], Alexander Yermanos[1,4], Jurgen Kuball[1,3], Can Keşmir[5], M. Ines Pascoal Ramos[1,2], Linde Meyaard[1,2]

[1]Center for Translational Immunology, University Medical Centre Utrecht, Utrecht University, Utrecht, The Netherlands

[2]Oncode Institute, Utrecht, The Netherlands

[3]Department of Haematology, University Medical Centre Utrecht, Utrecht University, Utrecht, The Netherlands

[4]Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland

[5]Theoretical Biology and Bioinformatics, Department of Biology, Utrecht University, Utrecht, The Netherlands

Conflict of interests

JK is cofounder, shareholder and scientific advisor of GADETA. LM has received funding for investigator-initiated studies on inhibitory receptors from Boehringer Ingelheim, NextCure, NGM Biopharmaceuticals and Argenx. AS, AMB, HJR, MV, JK, CK, MIPR and LM are named as co-inventors on a patent application based on this work. The remaining authors declare no potential conflicts of interest.

Abstract

Blocking inhibitory receptors like PD-1 and CTLA-4 has revolutionized cancer treatment in recent years. However, despite major successes in melanoma and lung cancer, the majority of cancer types are not responsive to these immunotherapies. As such, there is an ongoing need for the identification of novel inhibitory receptors as drug targets. Most inhibitory receptors signal via immunoreceptor tyrosine-based inhibitory motifs (ITIMs) and previous studies have estimated that our genome contains over 1600 ITIM-bearing transmembrane proteins. However, further testing and development of this large number of candidates requires increased understanding of their expression patterns and likelihood to function as inhibitory receptor.

To assist in the selection of novel inhibitory receptor as therapeutic targets, we designed a novel bioinformatics pipeline integrating machine learning-guided structural predictions and sequence-based likelihood models to identify 51 known and 390 putative inhibitory receptors. Using publicly available transcriptomics data of immune cells, we determined the expression of these novel inhibitory receptors, and classified them into previously proposed functional categories.

Known and putative inhibitory receptors were expressed across a wide variety of immune cells, and we found cell type-specific patterns in expression of these receptors. We used our pipeline to study inhibitory receptor expression patterns in single cell transcriptomics data of tumour infiltrating T cells. We determined that putative immune inhibitory receptors were expressed differentially in $CD4^+$ and $CD8^+$ T cell subsets, including exhausted $CD8^+$ T cells and $CD4^+$ memory T cells, which could allow for subset-specific targeting.

In conclusion, we present an inhibitory receptor pipeline that identifies 51 known and 390 novel inhibitory receptors. This pipeline will support future drug target selection across diseases where therapeutic targeting of immune inhibitory receptors is warranted.

Introduction

In the past decade, checkpoint blockade therapy has revolutionized the treatment of cancer by releasing immune cells from inhibition in the tumour microenvironment. Approval of blocking antibodies targeting immune inhibitory receptors, i.e., cytotoxic T cell lymphocyte antigen 4 (CTLA-4) (ipilimumab), the programmed cell death protein 1 (PD-1)/PD-L1 axis (nivolumab, atezolizumab) and lymphocyte-activation gene 3 (LAG-3)[1], resulted in significantly increased responses in previously hard-to-treat cancers, such as metastatic melanoma or non-small-cell lung cancer (NSCLC)[2]. Despite these successes, most cancer patients do not respond to checkpoint therapy, or do not have a durable treatment response[3]. Furthermore, the central role for CTLA-4 and PD-1 in maintaining peripheral tolerance results in immune related adverse events for many patients[4]. Further understanding of tissue- or activation-specific expression of different inhibitory receptors can assist in limiting toxicities, in particular when receptors are also expressed on non-immune cells[5]. Additional benefit can be gained by targeting inhibitory receptors with more specialized functions or expression patterns, such as T cell immunoglobulin and mucin domain-containing receptor 3 (TIM-3) or T cell immunoreceptor with Ig and ITIM domains (TIGIT)[6].

We previously proposed that inhibitory receptors can be classified into distinct functional categories based on the regulation of their expression[7]. Using mathematical modelling, we defined four categories: 1) Negative feedback receptor expression is induced by cellular activation to ensure a timely resolution of the response and limit excessive responses; 2) Threshold receptors are expressed on immune cells in the resting state and control the threshold for immune cell activation to provide context and prevent unnecessary or futile responses; 3) Threshold-disinhibition receptors are threshold receptors of which the expression is downregulated after activation. This allows for a more potent response after initial activation; and 4) Threshold-negative feedback receptors are threshold receptors of which expression is further induced upon cellular activation.

It is reasonable to assume that therapeutic interference with inhibitory receptors of different functional categories would have different outcomes. In cancer, blocking threshold receptors will lower the activation threshold of immune cells with antitumour activities. Meanwhile, interference with negative feedback receptors will prolong the antitumour and cytolytic activities of immune cells.

The vast majority of currently characterized inhibitory receptors recruit phosphatases, such as SH2 domain-containing phosphatase (SHP-1), SHP-2, or SHIP-1 to their intracellular immunoreceptor via tyrosine-based inhibitory and/or switch motifs (ITIM/ITSMs)[8]. This allows these receptors to dephosphorylate and inactivate other signalling molecules, such as those recruited to the T cell receptor complex, although some receptors are also capable of relaying activating signals through an ITSM[9,10]. ITIM sequences consist of the consensus amino acid sequence (V|L|I|S)xYxx(I|L|V), where x can be any amino acid, in the intracellular domain of a protein. This domain can be extended to (V|L|I|S|T)xYxx(I|L|V) to also include ITSM sequences, as there are inhibitory receptors, such as PD-1, that rely partly on the ITSM for their inhibitory function[11]. So far, around 50 ITIM/ITSM containing inhibitory receptors have been described, although it has previously been estimated that over 1600 ITIM-containing molecules can be found in the human genome[12].

More specific selection of potential drug targets out of this large collection of predicted receptors is required for further development and subsequent targeting in disease. Here, we set out to predict potential targets by developing a computational framework to identify putative inhibitory receptors and integrate three-dimensional structure predictions of proteins to those with a high likelihood of encoding functional inhibitory receptors. Furthermore, as inhibitory receptor expression follows specific dynamics and can play different roles in orchestrating immune responses[7], we analysed changes in the expression of putative inhibitory receptors in different immune cells and activation states. Our analysis will aid in the

selection of inhibitory receptors as therapeutic targets for specific diseases, to minimize toxicity and maximize efficacy.

Results

Putative ITIM/ITSM-bearing immune inhibitory receptors can be found in the human genome

To identify putative novel inhibitory receptors, we first retrieved the protein sequences of all 96,457 protein-coding transcripts, corresponding to 19,353 individual human genes, annotated in the Ensembl database (release 105, December 2021)[13]. We retrieved all deposited isoforms, as certain proteins might only contain an ITIM in a specific isoform. We then screened the entire protein sequence for the presence of an ITIM or ITSM, based on the consensus sequence (V|L|I|S|T)xYxx(I|L|V), resulting in 40,063 protein sequences from the initial 96,457. We next used the TOPCONS server, which is based on a consensus of several prediction tools[14], to determine the membrane topology of all proteins with an ITIM/ITSM motif. TOPCONS predicted 12,576 out of the 40,056 input proteins to be integral membrane proteins, of which 7,445 contained multiple transmembrane domains and the remaining 5,131 are single-pass membrane proteins.

We next filtered proteins for the presence of the ITIM or ITSM sequence in an intracellular domain of the protein based on the predicted topology. This yielded with 4,613 protein sequences bearing one or more intracellular ITIMs or ITSMs, which correspond to 1,562 genes (Figure 1A, Supplementary Table 1). To assess how well our pipeline reflects previously documented inhibitory receptors, we confirmed that all the 52 proteins known for their ITIM-mediated immune inhibitory effects, such as PD-1 or leukocyte-associated immunoglobulin-like receptor 1 (LAIR-1) (Table 1)[7], were recalled successfully, with the exception of *MPIG6B,* encoding for the inhibitory receptor G6B on platelets and megakaryocytes[15], due to an incorrect prediction of its transmembrane domain.

The composition of an ITIM/ITSM sequence is highly variable, as only 3 out of 6 amino acids positions are limited to specific amino acids. As a result of this, large proteins or proteins containing many tyrosines and hydrophobic residues are more likely to contain an ITIM/ITSM by chance (Supplementary Figure 1A). We therefore determined the likelihood of encountering an ITIM/ITSM by chance in the intracellular region of the putative targets. To this end, we randomly permutated the amino acid sequences of the intracellular domains by shuffling their order using the Fisher-Yates algorithm. We did this 10,000 times for each receptor and assessed how often the permutated sequence contained more or an equal amount of

ITIM/ITSMs compared to the original sequence. For 41 out of 51 known receptors, these odds were low (< 5%). However, for 10 out of 51, we observed ITIM/ITSM sequences in the permutated sequence between 5% and ~25% of the time. Based on these findings, we decided to exclude novel targets when they were above this threshold of 25% (Figure 1B). Due to the nature of permutation testing, there is some variation in the individual likelihood values for each protein sequence. However, as they were generally below 0.25 in any given iteration, we decided to define this value as a threshold for inclusion. After this statistical filtering, we were left with 837 inhibitory receptor genes. As expected, the average length of the intracellular domains of the excluded proteins was much higher than that of the remaining targets (Supplementary Figure 1B).

AlphaFold structure predictions can assist in identifying likely functional ITIM/ITSMs

Protein interaction motifs are most commonly found in intrinsically disordered regions within proteins[16], as structured regions are less accessible to binding with interacting partners. We therefore hypothesized that ITIM/ITSM sequences locate in disordered regions to allow for interaction with phosphatases and kinases. We therefore leveraged AlphaFold to infer three-dimensional structure predictions of the remaining inhibitory receptor candidates to determine where the ITIM/ITSM sequences were located in three dimensional space[17].

For 832 of the 837 remaining inhibitory receptor genes, we were able to retrieve the predicted structures using AlphaFold. For every residue in the three-dimensional model, AlphaFold defines a confidence score (pLDDT) between 0 and 100 pLDDT. Low scores indicate reduced confidence, and regions with a low score are more likely to be intrinsically disordered[18]. We determined the average pLDDT of all six residues encompassing the ITIM/ITSM sequence, what resulted in an overall confidence score for each intracellular ITIM/ITSM (Figure 1C) in all targets.

We found that 99 out of 101 ITIM/ITSMs of the 51 known receptors had low confidence scores, i.e., less than 80 pLDDT, with an average confidence score of 49.3 pLDDT. 64 of the ITIMs even had pLDDT values below 50, which means they are particularly likely to be intrinsically disordered[18] (Supplementary Figure 1C). This result confirmed the hypothesis that ITIM/ITSMs are likely to be located in disordered regions in the protein and indicates that the AlphaFold confidence score can be a valuable asset to determine the potential functionality of an ITIM sequence in predicted targets.

Based on the known receptors, we determined a threshold of 80 pLDDT and were left with 390 out of 832 putative ITIM-bearing inhibitory receptor genes likely to contain a functional ITIM/ITSM. Among these

genes, 54.6% consisted of multi-spanning proteins, while the remaining 177 single transmembrane domain targets consisted of 145 (37.2%) type I and 32 (8.2%) type II receptors. In some cases, a single gene was predicted to have both single-spanning and multi-spanning isoforms. Most putative inhibitory receptors (275) were based on the canonical transcript, as annotated by Ensembl, while 115 were identified as a non-canonical isoform of the gene.

Putative inhibitory receptors are expressed across immune cell subsets

To better characterize these receptors, we studied the expression patterns of the retrieved known and putative inhibitory receptor genes in various immune cells. We collected publicly available bulk RNA sequencing data of various immune cell subsets, both at resting state and after *in vitro* stimulation[19,20] (Table 2). We included all putative inhibitory receptors, including those that were identified in a non-canonical transcript of the gene. For this analysis, we considered a receptor as expressed when the expression was above the median overall gene expression in the sample, in either resting or activated state. From our set of inhibitory receptor genes, 2/51 known and 41/390 predicted genes were not detected in these datasets.

Almost all functionally characterized immune inhibitory receptors are type I transmembrane proteins, while only six are type II single-spanners. So far, no multi-spanning receptors have been described to have an inhibitory function in the immune system. Also, many genes encoding multi-spanning proteins only showed very limited expression across immune cell subsets in the datasets we used for our analysis. Therefore, we analysed the multi-spanning targets (Supplementary Figure 2) separately from the single-spanning novel targets (Figure 2).

We investigated the expression of 215 single-spanning receptor genes in these immune cell subsets and found that neutrophils (137 genes) and monocytes (120 genes) expressed the highest number of inhibitory receptor genes. Lymphoid cells expressed fewer inhibitory receptor genes compared to myeloid cells, with B and NK cells expressing 111 genes, followed by $CD8^+$ (106 genes) and $CD4^+$ (99 genes) T cells (Figure 2B). The relative contribution of known and novel targets in each subset varied between different immune cells. We found more novel targets being expressed in $CD4^+$ (86/99 genes, 86.9%) and $CD8^+$ T cells (88/106 genes, 84.3%), compared to neutrophils (98/137 genes, 71.5%). Neutrophils and monocytes uniquely expressed 9 and 3 inhibitory receptor genes respectively, but the vast majority were shared between two or more immune cell subsets (Figure 2C).

Compared to the single-spanners, we found slightly fewer multi-spanning inhibitory receptor genes expressed in the different immune cell subsets. Out of 197 predicted genes, neutrophils expressed the most multi-spanning inhibitory receptor genes (118 genes), while the other subsets expressed a similar number: Monocytes expressed 100 inhibitory receptor genes, followed by B cells (95 genes), $CD4^+$ (90 genes) and $CD8^+$ (90 genes) T cells, and NK cells (89 genes) (Supplementary Figure 2A-B). Neutrophils (14 genes) expressed the most unique set of multi-spanning inhibitory receptors, while most other genes were shared between all subsets (Supplementary Figure 2C).

Known and novel immune inhibitory receptors of different functional categories can be identified in immune cells

We next assigned the putative inhibitory receptor into the previously described functional categories by comparing the expression in the resting state versus after activation in the available data from different cell types[7] (Table 3, Supplementary Table 2). We defined negative feedback receptors as below median expression in the resting state, but expressed after activation. Threshold receptors were defined as being expressed in the resting state and having less than 0.5 $\log_2$ fold change after activation, while threshold-disinhibition and threshold-negative feedback receptors were defined by more than 0.5 $\log_2$ fold decrease or increase in expression upon activation, respectively. Using this functional classification system, we found that all immune cell subsets, except B cells, expressed receptors of all four categories, although all subsets primarily expressed threshold receptors.

We did not observe many changes in inhibitory receptor gene expression after activation in neutrophils. Six receptor genes, such as *SLAMF1*, were absent on resting neutrophils, but expression was induced after stimulation, categorizing them as negative feedback receptors. We identified 124 receptors as threshold receptors, while 4 and 3 genes were identified as threshold-disinhibition and threshold-negative feedback receptors, respectively (Figure 2B).

In contrast, in monocytes 26 genes, including *SIGLEC8* or *SIGLEC11*, had the expression pattern of negative feedback receptors. We found 72 receptors, e.g., *SIGLEC10* or *PILRA*, that maintained their expression after activation, while 21 were downregulated, such as *VSTM1* or *CLEC12A*. Only a single receptor, *SLAMF7*, was categorized as a threshold-negative feedback receptor. Of note, a cluster of 14 receptors was shared exclusively between monocytes and neutrophils (Figure 2C) and thus could be indicative of a set of conserved myeloid-specific checkpoints.

We observed expression of 17 negative feedback receptors, e.g., *SIGLEC15* and *VSTM4*, and 81 threshold receptors, such as *CD244* and *CD300A* on NK cells. Additionally, 11 genes were classified as threshold-disinhibition receptors, e.g., *FCRL6*, while 2 were determined to be threshold-negative feedback receptors.

B cells expressed 111 predicted inhibitory receptor genes, of which 10, like *SIRPA* and *SIGLEC6*, were negative feedback receptors. We found 85 threshold receptors, such as *BTLA*, as well as 16 threshold-disinhibition receptors, e.g., *PILRA* and *LST1*. B cells were the only subset that are not predicted to express threshold-negative feedback receptors. B cells also shared expression of 6 receptors with monocytes and neutrophils, while 2 genes were uniquely expressed by B cells.

$CD4^+$ and $CD8^+$ T cells expressed 8 and 9 negative feedback receptors, respectively, such as *PDCD1,* in addition to 73 (on $CD4^+$ T cells) and 78 (on $CD8^+$ T cells) threshold receptors, e.g., *CEACAM1* and *SIT1*. In addition, we found 16 threshold-disinhibition receptors on $CD4^+$ T cells and 15 on $CD8^+$ T cells, such as *LAIR1* on $CD4^+$ T cells and *AXL* on both subsets, as well as 2 threshold-negative feedback receptors on $CD4^+$ and 4 on $CD8^+$ T cells.

All immune cell subsets expressed fewer multi-membrane spanning inhibitory receptor genes compared to single-spanners (Supplementary Figure 2A, Supplementary Table 2). When we assigned these multi-spanning targets to the different functional categories, a similar picture emerged as for the single-spanners, where most receptors categorized as threshold receptors in all immune cell types. Whereas monocytes, B cells, $CD4^+$ and $CD8^+$ T cells expressed more single-spanning threshold-disinhibition receptors than negative feedback receptors, the opposite was observed for multi-spanning inhibitory receptor genes. NK cells had less dynamic gene regulation of multi-spanning targets than of single-spanners, while categorization for neutrophils was similar between multi- and single-spanners (Supplementary Figure 2B). Interestingly, the multi-spanning targets were shared between three or more subsets more often than was the case for the single-spanners (Supplementary Figure 2C).

In summary, we identified 398 known and putative inhibitory receptors being expressed across a variety of immune cell subsets, with mostly overlapping expression patterns. We classified these receptors into different functional categories based on their expression in the resting and activated states. We found that all immune cells expressed mostly threshold receptors. Among the receptors that changed expression after activation, monocytes and NK cells mainly expressed negative feedback receptors, while the T cells and B cells expressed more threshold-disinhibition receptors. Neutrophils did not have very dynamic gene expression patterns.

Known and putative inhibitory receptors are expressed on tumour infiltrating T cells

To further translate our findings to a disease context, we explored the expression of putative inhibitory receptor genes in tumour infiltrating lymphocytes. We used publicly available single cell RNA sequencing data of $CD4^+$ and $CD8^+$ T cells from 21 types of cancer and determined inhibitory receptor gene expression in the different T cell subsets identified by Zheng *et al.* (2021)[21] (Figure 3A). After quality control and filtering, the expression of 133/215 single-spanners, and 111/197 of the multi-spanners could be assessed in this dataset.

Among T cell subsets, expression of single-spanning inhibitory receptor genes was variable (Figure 3A). We found 27 and 46 inhibitory receptor genes in naïve $CD4^+$ and $CD8^+$ T cells, respectively, while T cells with a more differentiated phenotype, e.g., regulatory $CD4^+$ T cells (68 genes) and exhausted $CD8^+$ T cells (83 genes) expressed a wider array of known and predicted inhibitory receptor genes (Figure 3B-C).

When we assessed expression of multi-spanning inhibitory receptor genes (Supplementary Figure 3A), we found similar variable expression across T cell subsets as for the single-spanners. Naïve $CD4^+$ T cells expressed 21 multi-spanning receptor genes, while many more were found in the other subsets (Supplementary Figure 3B). Resident memory $CD8^+$ T cells expressed 87 multi-spanning inhibitory receptor genes, while $CD8^+$ memory and effector memory expressed 9 and 26 genes, respectively (Supplementary Figure 3C).

Based on the previously proposed functional categorization of known and novel inhibitory receptors (Figure 2 and Supplementary Figure 2), we assessed the functional categories expressed across the tumour infiltrating T cell subsets (Supplementary Figure 4). Similarly to the immune subset analysis, we observed that the subsets primarily expressed threshold receptors.

To validate that we rediscovered known immunotherapeutic targets, we assessed expression of the known inhibitory receptor genes in tumour infiltrating lymphocytes of melanoma patients using the same dataset [21]. We found high expression of known immunotherapeutic targets, such as PD-1, in addition to other inhibitory receptors being targeted in clinical trials, such as TIGIT (Supplementary Figure 5).

Overall, this shows that the putative inhibitory immune receptors we identified are expressed on tumour infiltrating T cells and differ between cellular subsets. Further investigation into differences and similarities between these populations in different cancer types could assist in targeting these receptors for therapy.

In conclusion, we designed a novel, combined bioinformatics approach to predict around 400 immune inhibitory receptors in the human genome, including both single and multi-spanning targets. Our list of novel targets will assist drug target selection in diseases where inhibitory receptor targeting is warranted.

Discussion

This study introduces some major improvements over previous attempts at predicting novel inhibitory receptors in the human genome. In 2004, Staub *et al.*, performed a search limited to type I membrane receptors containing an annotated extracellular domain and a restricted ITIM sequence. They identified a total of 94 genes encoding inhibitory receptors, of which 32 had been previously described to rely on their ITIM to inhibit immune function[22]. A later study by Daëron *et al.* (2008), performed a more extensive analysis of ITIM-bearing receptors across multiple species, including chickens and *Drosophila*[12]. In humans, they identified 613 predicted type I and type II single-spanners and 992 multi-spanning transmembrane proteins, corresponding to around 1500 genes. This number agrees with what we retrieved in the initial steps of our bioinformatics pipeline, where we identified 1,562 genes. Our efforts to reduce this number to candidates with a high probability of functioning as inhibitory receptors resulted in 390 novel targets. Compared to the previous studies, our method has key advantages. Firstly, the TOPCONS algorithm we used is superior to TMHMM and Philius, which have been used previously, at predicting membrane topology[14,23,24]. Most importantly, we excluded many potential false positives by filtering out proteins that are likely to contain an ITIM by chance and used three-dimensional structure predictions to enrich the remaining targets for those containing ITIMs in unstructured regions. Furthermore, our expression analysis and functional categorization of the putative inhibitory receptors can assist in gaining a better understanding of their relevance in different contexts.

An important component of our pipeline is the use of AlphaFold[17]. Our approach was exceptionally good at filtering out many proteins in the olfactory receptor family, for which almost all ITIMs identified were found in highly structured regions. We determined a threshold of 80 pLDDT based on the calculated prediction scores of the ITIM/ITSMs in known inhibitory receptors, as this encompassed 99 out of 101 ITIM/ITSMs. However, this excluded ITIMs in *CD5* (Y402) and *BTLA* (Y282), with confidence scores of 85.2 and 81.5, respectively. Of note, studies have suggested that this particular ITIM in *CD5* might not be functional or essential for CD5 function[25]. It is important to highlight that the AlphaFold structure predictions are based on the proteins being in isolation. It is possible that ITIMs become available after

conformational changes following ligand binding, or that ITIMs in disordered regions are not functional when the proteins are found in complexes[26,27].

We acknowledge that our approach has limitations. We used known ITIM-bearing inhibitory receptors as validation for the various steps of the pipeline and retrieved all but one receptor, resulting in a retrieval of 98%. We did not retrieve *MPIG6B*, as we were not able to properly predict the transmembrane helix in the ITIM-containing isoforms of *MPIG6B*. With increasing computing power and further development of artificial intelligence, these prediction algorithms are likely to further improve[28,29]. The identification of novel ITIM/ITSM-bearing inhibitory receptors also relies on our current knowledge of known inhibitory receptors and their motifs. This influences the thresholds in our analyses, and might be skewed by larger protein families, e.g., KIR- and LILRB-family receptors. Additional filtering of the candidates could be achieved by assessing subcellular localization, as it is possible that the transmembrane proteins we identified are not localized on the outer membrane, but instead on intracellular membranes. Although they could still function as inhibitory receptors, this would complicate the ability to effectively target them with therapeutic antibodies, and might instead require other blocking agents, such as proteolysis-targeting chimeras (PROTACS) or small molecule inhibitors[30].

We focused our search on receptors bearing ITIM or ITSM sequences, while inhibitory receptors have been described that function independent of these short sequences. We previously reported that CD200R has a unique signalling motif that is not shared with any other protein[31]. Also, TIM-3 contains multiple tyrosine residues that are not part of any known signalling motif[32]. A small number (26 genes) of putative inhibitory receptors contained an ITIM embedded in an ITAM sequence, i.e., Yxx(I|L)x_{6-12}Yxx(I|L). Although some ITAM-bearing molecules are capable of mediating inhibitory signals in an ITIM-independent manner, these receptors may not all be inhibitory[33].

Until now, no known functional multimembrane spanning immune inhibitory receptors have been reported, although some ITIM-bearing multi-spanners, e.g., *SSTR2* and *HCRTR1*, were shown to recruit SHP-2[34,35]. It is possible that multi-spanning inhibitory receptors have not yet been fully characterized due to their size and limited targetability, making them difficult to study. Of interest, in our analysis multi-spanning receptors were often expressed by multiple immune cell subsets. This may suggest that multi-spanning inhibitory receptors would be capable of inhibiting a broad array of immune responses.

We previously argued how different functional categories of inhibitory receptors might serve distinct purposes[7]. Here, we grouped known and putative inhibitory receptors in different functional categories based on their expression in isolated immune cell subsets in resting and stimulated state and show that a

large number of them is expressed and regulated in a cell type-specific manner. However, as we categorized these receptors based on *in vitro* activation with limited stimuli, expression must be further validated in *in vivo* settings. Nonetheless, we tried to characterize the functional categories expressed by tumour infiltrating lymphocytes by extrapolating the *in vitro* defined functional categorization per gene. This showed that mainly threshold receptors and some (threshold-)negative feedback receptors are expressed by the different T cell subsets. This would open the possibility of targeting different functional categories for cancer immunotherapy, since targeting a threshold receptor to lower the threshold for activation and a negative feedback receptor to lengthen and strengthen the cellular response might therefore be more effective than targeting two receptors of a single functional category. However, we acknowledge that this will require further validation of expression patterns *in vivo* in different cancers and immune cell subsets.

We found that the highest number of target genes are expressed in neutrophils, which fits the need to control and precisely regulate their potentially damaging effector functions[36]. However, inhibitory receptor gene expression on these cells did not change much after activation. This does not exclude regulation of receptor expression at the protein level, which would fit with the more short-lived nature of these cells. For instance, LAIR-1 protein is stored in intracellular granules which results in rapid membrane expression on activated granulocytes[36]. On monocytes, we found many negative feedback receptors. Considering that monocytes (especially macrophages) play a central role in orchestrating the overall immune response, timely resolution of their activation is key to successful resolution of inflammation, which could be achieved by these receptors.

For the functional categorization of novel inhibitory receptor genes, we collected multiple publicly available datasets containing different isolated immune cell subsets in the resting state and after cell type-specific activation. For 115 genes, we predicted a non-canonical isoform, but not the canonical sequence, as a putative inhibitory receptor. This means that our expression analyses for these inhibitory receptor genes might not be fully reflective of the expression of individual isoforms of those genes, considering that isoform usage can vary widely between different cells and activation states, as has been well described for CEACAM1[37,38]. In the present study, we also did not consider the ligand expression of these known and putative receptors. Regulation of inhibitory receptors does not only occur at the level of the receptor itself, but the presence or absence of ligands, decoy receptors, and other interactions all add additional layers of regulation of control to the immune system[7].

To further prioritize inhibitory receptors in immune cell subsets or diseases of interest, gene co-expression networks of putative inhibitory receptors could be assessed. On the one hand, the co-occurrence of putative inhibitory receptors with known inhibitory receptors within a module could be one approach, while on the other hand the presence of putative inhibitory receptors in a different module could suggest novel regulation of different biological functions than the known receptors. The location of the putative inhibitory receptors in the network could also change depending on the cell type and the activation status of the cell. Additionally, one could look at the co-expression of candidates with other genes within a gene module to look at potential biological function, and at co-expression with signalling molecules known to interact with inhibitory receptors, such as Csk, SHP-1, SHP-2 and SHIP1, although their regulation might be more post-translationally regulated rather than at the transcriptional level.

As proof-of-concept, we assessed expression of both known and novel targets in various tumour infiltrating T cell subsets. This helps understanding how these potential drug targets are differentially expressed in a disease context. As many receptors we identified were shared across immune cell subsets, it would be particularly interesting to block those that are enriched in, e.g., effector $CD8^+$ T cell subsets. While some T cell subsets, such as exhausted $CD8^+$ T cells and $CD4^+$ regulatory T cells, appear to not differ much in their expression of either single- or multi-spanning receptors, we do observe that, for example, effector memory $CD4^+$ T cells or EMRA $CD8^+$ T cells express single-spanning inhibitory receptors to a higher extent than multi-spanning inhibitory receptors. It is possible that these differences and similarities reflect some of the roles multi-spanning inhibitory receptors could play in regulating immune cells, for example in response to chemokines, as many chemokine receptors are multi-spanning proteins. In contrast, receptors expressed in subsets that can have pro-tumorigenic properties, such as $CD4^+$ T_{reg} or Th17 cells[39,40], should not be activated; these targets could instead benefit from agonist therapy. Furthermore, these putative receptors might also be attractive in autoimmune diseases, where immune checkpoints are increasingly being explored as potential targets for therapy[41,42].

Taken together, we have designed a bioinformatic pipeline to identify a key list of 390 predicted inhibitory receptors to be used as a starting point for further research into novel targets for targeting in diseases.

Methods

Identification of putative inhibitory receptors in the human genome

To identify putative inhibitory receptors in the human genome, amino acid sequences for 96,457 protein-coding, as annotated by Ensembl, transcripts were retrieved from Ensembl (release 105, December 2021)[13] using the *biomaRt* package[43] in R. Amino acid sequences were then filtered based on the presence of an ITIM or ITSM sequence, (V|L|I|S|T)xYxx(I|L|V). Seven out of 40,063 ITIM-containing proteins were larger than 10,000 amino acids, and were excluded because they are not feasible for membrane topology prediction using the TOPCONS method[14]. Topology results for the remaining 40,056 proteins were then retrieved, and proteins were further filtered based on the presence of a predicted transmembrane domain (12,576 out of 40,056 proteins). For the remaining proteins, we determined whether the identified ITIM or ITSM sequence corresponded to an intracellular section of the protein, which resulted in 4,613 protein sequences remaining, corresponding to 1,562 individual genes (Figure 1A). A further 61 protein sequences were excluded as they could not be matched to a HUGO Gene Nomenclature Committee (HGNC) gene symbol[44].

Determining the likelihood of intracellular protein domain containing an ITIM by chance

To determine the odds of finding an intracellular ITIM or ITSM sequence in a specific protein, all intracellular domains of the proteins were retrieved based on the TOPCONS predictions. For each protein, the amino acids in the intracellular domain were then shuffled randomly, using the *stringi* package[45], and the number of identified ITIM or ITSM sequences in the permutated sequences was compared to the actual number of intracellular ITIM or ITSM sequences in the original protein. Each protein was permutated 10,000 times, and the number of occurrences of a permutated protein containing more or equal ITIM/ITSMs to the original sequence was counted and used as measure for the likelihood of finding an ITIM or ITSM by chance. A threshold was defined based on likelihood scores for the proteins corresponding to known inhibitory receptors, and was set at 0.25+ε, with ε being a random number determined error between 0 and 0.01.

Determining AlphaFold model confidence scores for individual ITIMs

Three-dimensional structure of protein sequences that passed the likelihood threshold were next determined using AlphaFold (v2.2.0) under default parameters with amber minimization (run_relax = True) using the pdb70, uniref90, mgnify (mgy_clusters_2018_12), uniclust30 (uniclust30_2018_08) and bfd (bfd_metaclust_clu_complete_id30_c90_final_seq) databases on a high-performance cluster[17,46]. AlphaFold was not able to determine the structure for 46 of these proteins, of which 6 were too long and 40 returned a run time error involving HHblits and the hh-suite. The three-dimensional prediction of 37 of these proteins was instead determined using ColabFold[47] (v1.5.2) using MMseqs2 and HHsearch under

default parameters without using template information (template_mode = none) and using mmseq2_uniref_env as MSA mode. The top ranked prediction files were retrieved for each protein, and loaded into R using the *bio3d* package[48]. Prediction confidence scores (pLDDT) were retrieved, and the average pLDDT for all six amino acids encompassing the ITIM or ITSM sequence was used as the AlphaFold confidence score for that ITIM or ITSM. Proteins that had no ITIM or ITSM sequences with an average confidence score below 80 were used in further analysis.

Classification of receptor functional categories

Based on the *in vitro* RNA sequencing dataset of immune cells in the resting state and after activation, functional categories were defined based on expression patterns as outlined previously[7]. Data was retrieved from Calderon *et al.* (2019) and West *et al.* (2016)[19,20] Gene symbols were matched to Ensembl Gene IDs using biomaRt and Ensembl (release 108, October 2022)[13]. For the neutrophil data[19], TPM values were retrieved and log2 scaled. For other immune cells, count matrix was normalized using the *DESeq2* package[49]. Culture conditions for all immune cells were simplified to major immune cell subsets, and to *resting* and *stimulated* (Table 2). Average expression for each inhibitory receptor gene was determined for each sample. An inhibitory receptor gene was considered expressed when it was above the median expression of all genes within the sample. Genes that were not expressed in the resting state, but were expressed in the stimulated state were considered negative feedback receptors. Genes expressed in the resting state were considered threshold, threshold-disinhibition, or threshold-negative feedback receptors if the log2 fold change in expression was between -0.5 and 0.5, below -0.5, or above 0.5, respectively. Genes expressed in the resting state, but not after activation were also considered threshold-disinhibition receptors.

Collection and normalization of immune cell RNAseq data

For the analysis of tumour infiltrating T cells, a normalized expression matrix was retrieved from Zheng *et al.* (2021)[21,50]. The average Z score of all cells in a T cell subset was determined, weighted by the relative contribution of different subclusters identified in the study. Inhibitory receptor genes were considered expressed in a T cell subset if the average Z score was above 0.

Data availability

Code and data used to generate the figures in this study can be retrieved from https://github.com/AkashdipSingh/IIR_pipeline

References

1 Chocarro L, Bocanegra A, Blanco E, *et al.* Cutting-Edge: Preclinical and Clinical Development of the First Approved Lag-3 Inhibitor. *Cells* 2022;**11**:2351. doi:10.3390/cells11152351

2 Andrews LP, Yano H, Vignali DAA. Inhibitory receptors and ligands beyond PD-1, PD-L1 and CTLA-4: breakthroughs or backups. *Nat Immunol* 2019;**20**:1425–34. doi:10.1038/s41590-019-0512-0

3 Sharma P, Hu-Lieskovan S, Wargo JA, *et al.* Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy. *Cell* 2017;**168**:707–23. doi:10.1016/j.cell.2017.01.017

4 Wolchok JD, Kluger H, Callahan MK, *et al.* Nivolumab plus Ipilimumab in Advanced Melanoma. *N Engl J Med* 2013;**369**:122–33. doi:10.1056/NEJMoa1302369

5 von Richthofen HJ, Meyaard L. Sensing context: Inhibitory receptors on non-hematopoietic cells. *Eur J Immunol* 2023;**NA**:2250306. doi:10.1002/eji.202250306

6 Schnell A, Bod L, Madi A, *et al.* The yin and yang of co-inhibitory receptors: toward anti-tumor immunity without autoimmunity. *Cell Res* 2020;**30**:285–99. doi:10.1038/s41422-020-0277-x

7 Rumpret M, Drylewicz J, Ackermans LJE, *et al.* Functional categories of immune inhibitory receptors. *Nat Rev Immunol* 2020;**20**:771–80. doi:10.1038/s41577-020-0352-z

8 Long EO. REGULATION OF IMMUNE RESPONSES THROUGH INHIBITORY RECEPTORS. *Annu Rev Immunol* 1999;**17**:875–904. doi:10.1146/annurev.immunol.17.1.875

9 Dietrich J, Cella M, Colonna M. Ig-Like Transcript 2 (ILT2)/Leukocyte Ig-Like Receptor 1 (LIR1) Inhibits TCR Signaling and Actin Cytoskeleton Reorganization. *J Immunol* 2001;**166**:2514–21. doi:10.4049/jimmunol.166.4.2514

10 Ostrakhovitch EA, Li SS-C. The role of SLAM family receptors in immune cell signaling. *Biochem Cell Biol* 2006;**84**:832–43. doi:10.1139/O06-191

11 Chemnitz JM, Parry RV, Nichols KE, *et al.* SHP-1 and SHP-2 Associate with Immunoreceptor Tyrosine-Based Switch Motif of Programmed Death 1 upon Primary Human T Cell Stimulation, but Only Receptor Ligation Prevents T Cell Activation. *J Immunol* 2004;**173**:945–54. doi:10.4049/jimmunol.173.2.945

12 Daëron M, Jaeger S, Du Pasquier L, *et al.* Immunoreceptor tyrosine-based inhibition motifs: a quest in the past and future. *Immunol Rev* 2008;**224**:11–43. doi:10.1111/j.1600-065X.2008.00666.x

13 Cunningham F, Allen JE, Allen J, *et al.* Ensembl 2022. *Nucleic Acids Res* 2022;**50**:D988–95. doi:10.1093/nar/gkab1049

14 Tsirigos KD, Peters C, Shu N, *et al.* The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res* 2015;**43**:W401–7. doi:10.1093/nar/gkv485

15 de Vet ECJM, Aguado B, Campbell RD. G6b, a Novel Immunoglobulin Superfamily Member Encoded in the Human Major Histocompatibility Complex, Interacts with SHP-1 and SHP-2. *J Biol Chem* 2001;**276**:42070–6. doi:10.1074/jbc.M103214200

16 Tompa P, Davey NE, Gibson TJ, *et al.* A Million Peptide Motifs for the Molecular Biologist. *Mol Cell* 2014;**55**:161–9. doi:10.1016/j.molcel.2014.05.032

17 Jumper J, Evans R, Pritzel A, *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* 2021;**596**:583–9. doi:10.1038/s41586-021-03819-2

18 Tunyasuvunakool K, Adler J, Wu Z, *et al.* Highly accurate protein structure prediction for the human proteome. *Nature* 2021;**596**:590–6. doi:10.1038/s41586-021-03828-1

19 West EE, Spolski R, Kazemian M, *et al.* A TSLP-complement axis mediates neutrophil killing of methicillin-resistant *Staphylococcus aureus*. *Sci Immunol* 2016;**1**:eaaf8471. doi:10.1126/sciimmunol.aaf8471

20 Calderon D, Nguyen MLT, Mezger A, *et al.* Landscape of stimulation-responsive chromatin across diverse human immune cells. *Nat Genet* 2019;**51**:1494–505. doi:10.1038/s41588-019-0505-9

21 Zheng L, Qin S, Si W, *et al.* Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science* 2021;**374**:abe6474. doi:10.1126/science.abe6474

22 Staub E, Rosenthal A, Hinzmann B. Systematic identification of immunoreceptor tyrosine-based inhibitory motifs in the human proteome. *Cell Signal* 2004;**16**:435–56. doi:10.1016/j.cellsig.2003.08.013

23 Möller S, Croning MDR, Apweiler R. Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* 2001;**17**:646–53. doi:10.1093/bioinformatics/17.7.646

24 Reynolds SM, Käll L, Riffle ME, *et al.* Transmembrane Topology and Signal Peptide Prediction Using Dynamic Bayesian Networks. *PLoS Comput Biol* 2008;**4**:e1000213. doi:10.1371/journal.pcbi.1000213

25 Voisinne G, Gonzalez de Peredo A, Roncagalli R. CD5, an Undercover Regulator of TCR Signaling. *Front Immunol* 2018;**9**:2900. doi:10.3389/fimmu.2018.02900

26 Jakob U, Kriwacki R, Uversky VN. Conditionally and Transiently Disordered Proteins: Awakening Cryptic Disorder To Regulate Protein Function. *Chem Rev* 2014;**114**:6779–805. doi:10.1021/cr400459c

27 Kiefhaber T, Bachmann A, Jensen KS. Dynamics and mechanisms of coupled protein folding and binding reactions. *Curr Opin Struct Biol* 2012;**22**:21–9. doi:10.1016/j.sbi.2011.09.010

28 Rives A, Meier J, Sercu T, *et al.* Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proc Natl Acad Sci* 2021;**118**:e2016239118. doi:10.1073/pnas.2016239118

29 . Tsirigos KD, Govindarajan S, Bassot C, *et al.* Topology of membrane proteins — predictions, limitations and variations. *Curr Opin Struct Biol* 2018;**50**:9–17. doi:10.1016/j.sbi.2017.10.003

30 Békés M, Langley DR, Crews CM. PROTAC targeted protein degraders: the past is prologue. *Nat Rev Drug Discov* 2022;**21**:181–200. doi:10.1038/s41573-021-00371-6

31 Timmerman LM, de Graaf JF, Satravelas N, *et al.* Identification of a novel conserved signaling motif in CD200 receptor required for its inhibitory function. *PLOS ONE* 2021;**16**:e0244770. doi:10.1371/journal.pone.0244770

32 Wolf Y, Anderson AC, Kuchroo VK. TIM3 comes of age as an inhibitory receptor. *Nat Rev Immunol* 2020;**20**:173–85. doi:10.1038/s41577-019-0224-6

33 Barrow AD, Trowsdale J. You say ITAM and I say ITIM, let's call the whole thing off: the ambiguity of immunoreceptor signalling. *Eur J Immunol* 2006;**36**:1646–53. doi:10.1002/eji.200636195

34 Ferjoux G, Lopez F, Esteve J-P, *et al.* Critical Role of Src and SHP-2 in sst2 Somatostatin Receptor-mediated Activation of SHP-1 and Inhibition of Cell Proliferation. *Mol Biol Cell* 2003;**14**:3911–28. doi:10.1091/mbc.e03-02-0069

35 Voisin T, Firar AE, Rouyer-Fessard C, *et al.* A hallmark of immunoreceptor, the tyrosine-based inhibitory motif ITIM, is present in the G protein-coupled receptor OX1R for orexins and drives apoptosis: a novel mechanism. *FASEB J* 2008;**22**:1993–2002. doi:10.1096/fj.07-098723

36 Geerdink RJ, Hennus MP, Westerlaken GHA, *et al.* LAIR-1 limits neutrophil extracellular trap formation in viral bronchiolitis. *J Allergy Clin Immunol* 2018;**141**:811–4. doi:10.1016/j.jaci.2017.08.031

37 Gaur S, Shively JE, Yen Y, *et al.* Altered splicing of CEACAM1 in breast cancer: Identification of regulatory sequences that control splicing of CEACAM1 into long or short cytoplasmic domain isoforms. *Mol Cancer* 2008;**7**:46. doi:10.1186/1476-4598-7-46

38 Helfrich I, Singer BB. Size Matters: The Functional Role of the CEACAM1 Isoform Signature and Its Impact for NK Cell-Mediated Killing in Melanoma. *Cancers* 2019;**11**:356. doi:10.3390/cancers11030356

39 Tay RE, Richardson EK, Toh HC. Revisiting the role of CD4+ T cells in cancer immunotherapy—new insights into old paradigms. *Cancer Gene Ther* 2021;**28**:5–17. doi:10.1038/s41417-020-0183-x

40 Karpisheh V, Ahmadi M, Abbaszadeh-Goudarzi K, *et al.* The role of Th17 cells in the pathogenesis and treatment of breast cancer. *Cancer Cell Int* 2022;**22**:108. doi:10.1186/s12935-022-02528-8

41 Grebinoski S, Vignali DA. Inhibitory receptor agonists: the future of autoimmune disease therapeutics? *Curr Opin Immunol* 2020;**67**:1–9. doi:10.1016/j.coi.2020.06.001

42 Zhai Y, Moosavi R, Chen M. Immune Checkpoints, a Novel Class of Therapeutic Targets for Autoimmune Diseases. *Front Immunol* 2021;**12**:645699. doi:10.3389/fimmu.2021.645699

43 Durinck S, Spellman PT, Birney E, *et al.* Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc* 2009;**4**:1184–91. doi:10.1038/nprot.2009.97

44 Seal RL, Braschi B, Gray K, *et al.* Genenames.org: the HGNC resources in 2023. *Nucleic Acids Res* 2023;**51**:D1003–9. doi:10.1093/nar/gkac888

45 Gagolewski M. stringi: Fast and Portable Character String Processing in R. *J Stat Softw* 2022;**103**. doi:10.18637/jss.v103.i02

46 Varadi M, Anyango S, Deshpande M, *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* 2022;**50**:D439–44. doi:10.1093/nar/gkab1061

47 Mirdita M, Schütze K, Moriwaki Y, *et al.* ColabFold: making protein folding accessible to all. *Nat Methods* 2022;**19**:679–82. doi:10.1038/s41592-022-01488-1

48 Grant BJ, Skjærven L, Yao X. The BIO3D packages for structural bioinformatics. *Protein Sci* 2021;**30**:20–30. doi:10.1002/pro.3923

49 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;**15**:550. doi:10.1186/s13059-014-0550-8

50 Zheng, Liangtao, Qin, Shishang. Codes for the paper 'Pan-cancer single-cell landscape of tumor-infiltrating T cells'. Published Online First: 17 December 2021. doi:10.5281/ZENODO.5461803

Acknowledgements

We would like to thank the IT team at ETHZ for their assistance with the setup of AlphaFold and the members of the Inhibitory Receptor Lab for their helpful comments during the preparation of the manuscript.

Figures

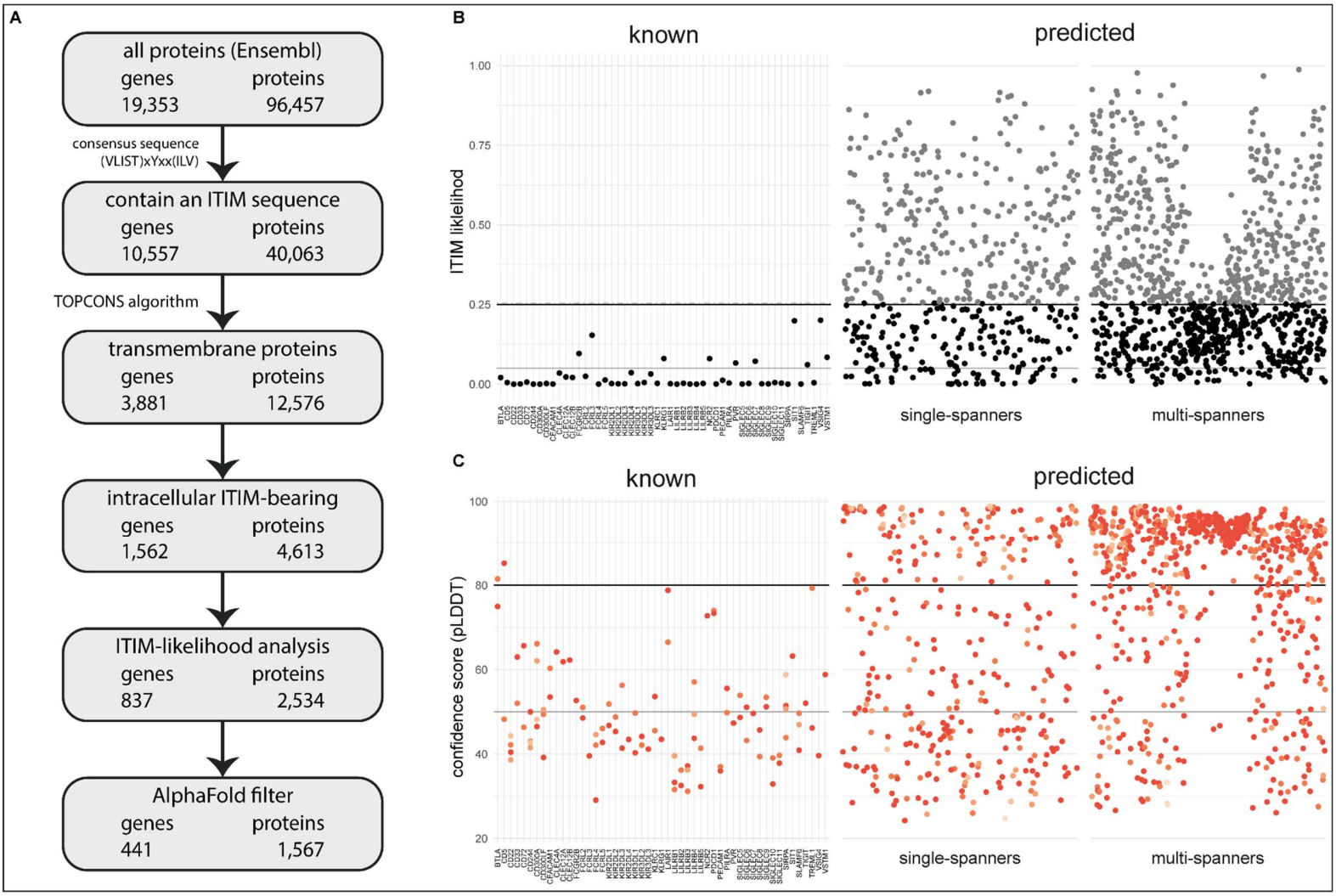
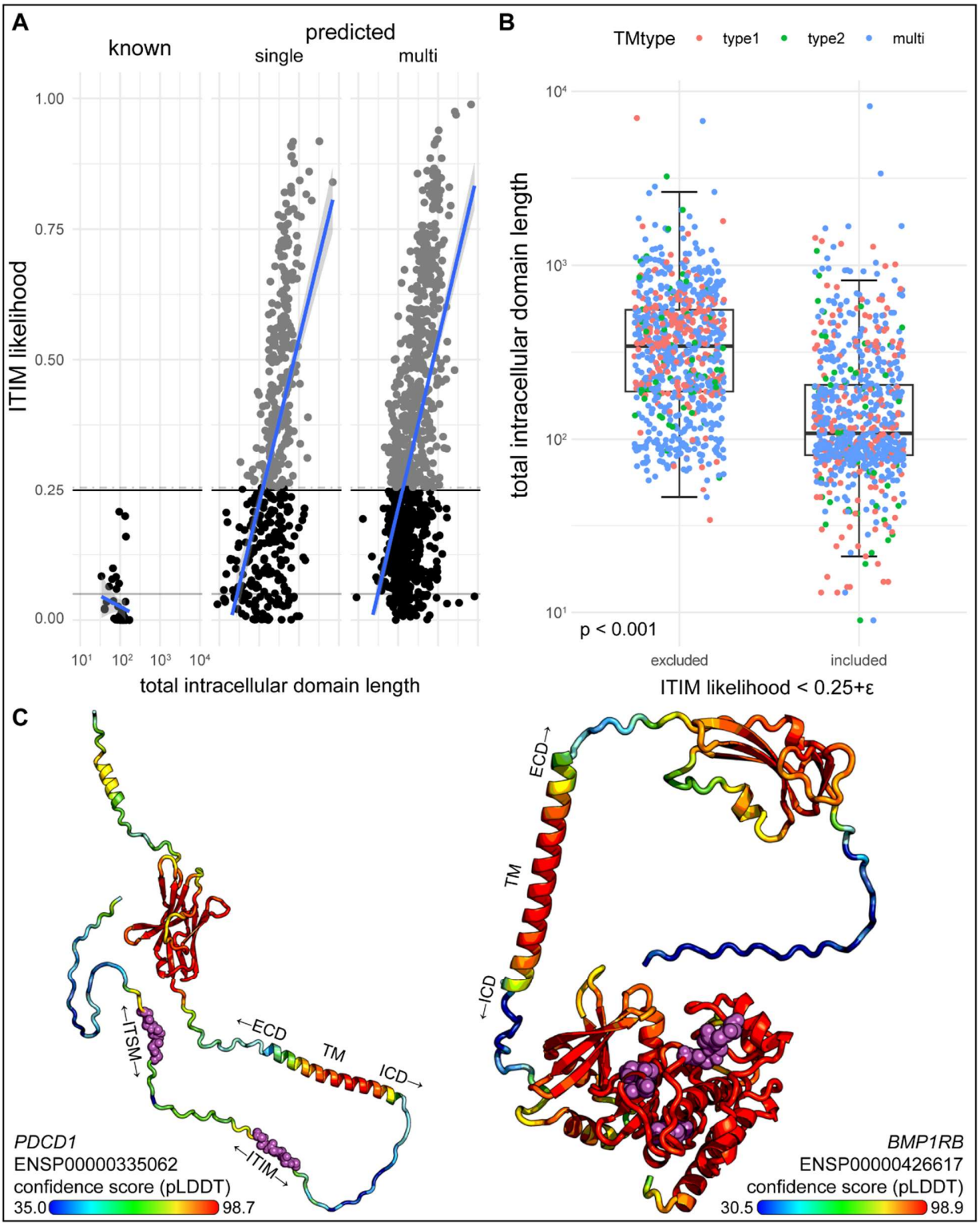


Figure 1. A novel bioinformatics approach for the identification of predicted inhibitory receptors. A. Schematic overview of the bioinformatics pipeline with the number of unique genes and corresponding proteins remaining at every step. All amino acid sequences corresponding to a protein-coding transcript were retrieved from Ensembl. **B.** Intracellular domains of identified proteins were permutated 10,000 times, and the number of ITIM or ITSM occurrences were compared to the number of ITIMs in the original sequence to determine the likelihood of a specific intracellular domain containing an ITIM or ITSM. Threshold was determined based on known inhibitory receptors, and set at 0.25+ε, with ε being a random number between 0 and 0.01 to better predict the borderline predictions. The black line indicates 0.25 likelihood, the grey dotted line indicates 0.25+ε and the solid grey line indicates 0.05 likelihood**. C.** Three-dimensional structure for all proteins was predicted using AlphaFold, and the average model prediction score (pLDDT) was determined for each individual ITIM or ITSM in the protein. Proteins with all ITIMs above 80 pLDDT were excluded. The black line indicates 80 pLDDT threshold and the solid grey line indicates 50 pLDDT. For **B** and **C**, one protein is plotted for every unique gene symbol for clarity.



Supplementary Figure 1. A novel bioinformatics approach for the identification of predicted inhibitory receptors. **A.** The overall likelihood of finding an ITIM in the intracellular domain based on the permutation method was plotted as function of total intracellular domain length for known and predicted single-spanner and multi-spanner proteins. **B.** The average size of the intracellular domain of proteins that were excluded by the likelihood filter was significantly higher than of those that were included. **C**. Examples of three-dimensional structure predictions by AlphaFold. Colouring indicates high (red) and low (blue) confidence residues in the protein, ITIMs for *PDCD1* (left) and *BMP1RB* (right) are depicted as purple spheres.

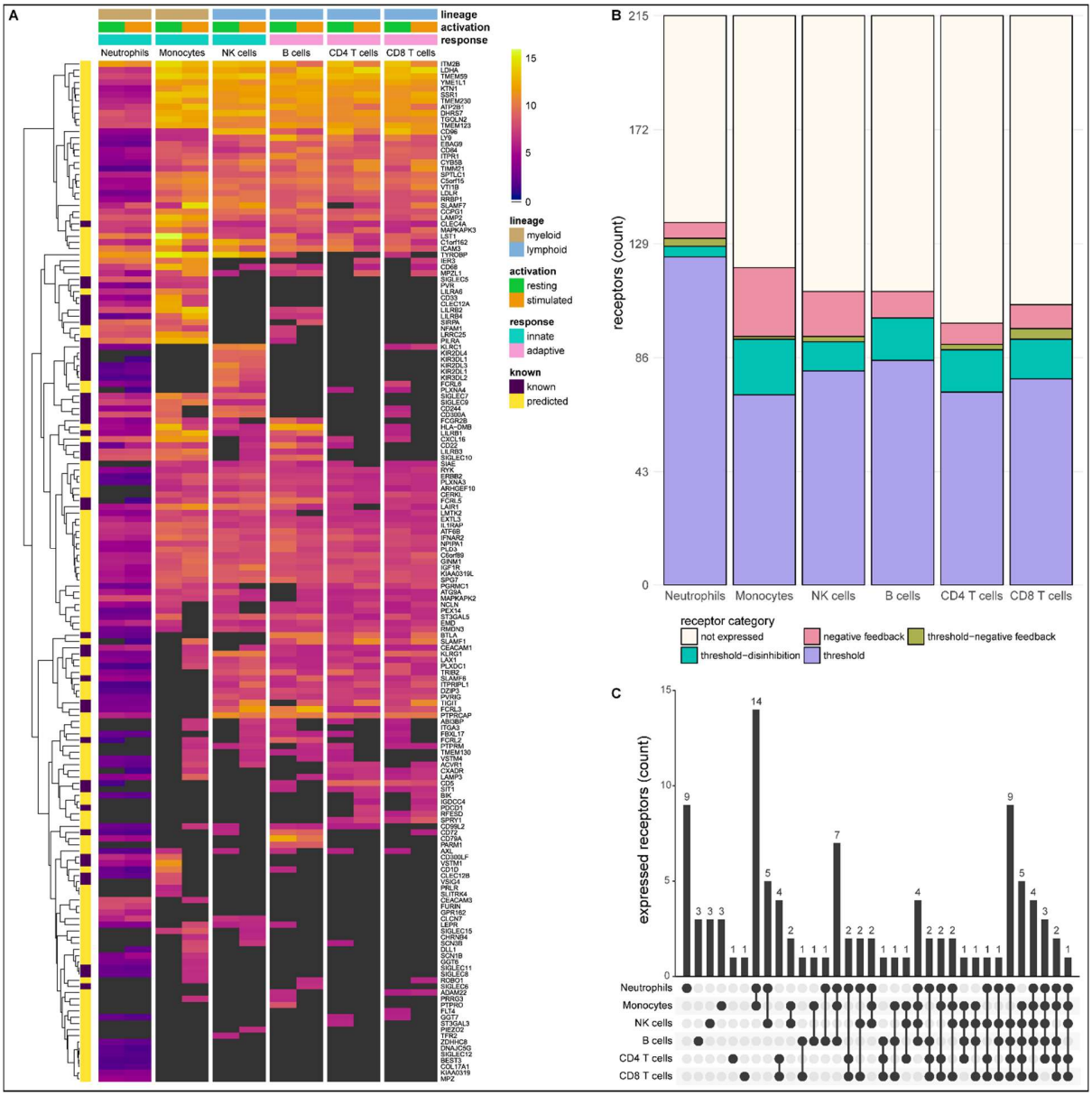
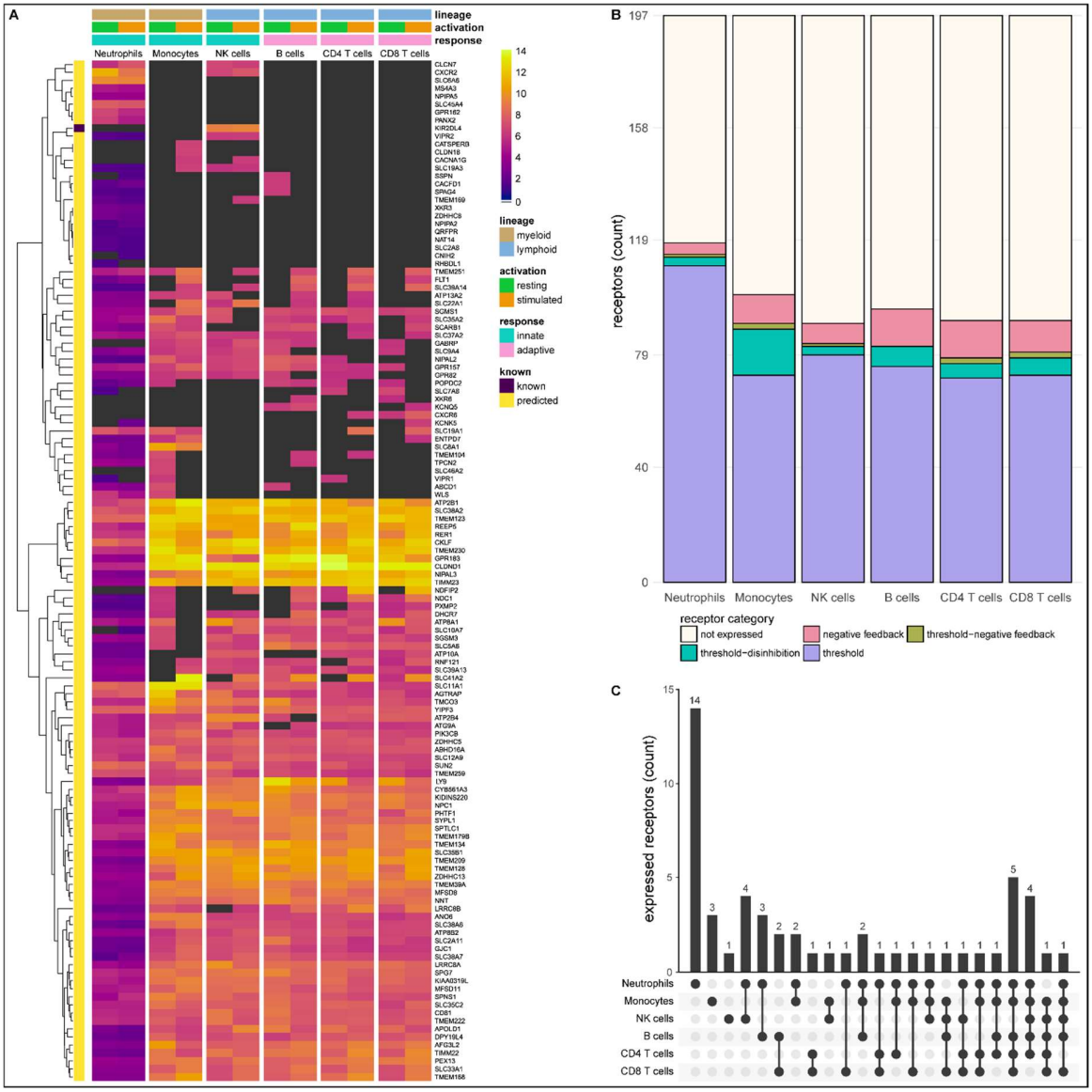


Figure 2. Known and predicted single-spanning inhibitory receptors are expressed in different cell types in the resting state and after activation. A. Heatmap with normalized expression data for known and predicted single-spanning inhibitory receptor genes in different cell types, in the resting state or after activation. Receptors were considered not expressed (black) when expression was below median overall gene expression in the sample. Data for neutrophils was retrieved from a different source than the other cell types. **B.** Novel and known receptors were classified into different functional categories based on changes, or lack thereof, in expression after stimulation. Threshold receptors were expressed at resting state, and did not change after activation (i.e., change in expression < 0.5 log2 fold change). Threshold-negative feedback and threshold-disinhibition receptors were defined by > 0.5 log2 fold change up or downregulation, respectively, in expression after activation. Negative feedback receptors were absent in the resting state, but were expressed after activation. **C.** Upset plot showing the number of single-spanning receptors that are expressed uniquely by individual immune cell subsets, or shared between subsets as indicated by connected circles. Sixty-four genes are expressed in all cell types (not depicted).



Supplementary Figure 2. Known and predicted multi-spanning inhibitory receptors are expressed in different cell types in the resting state and after activation. A. Heatmap with normalized expression data for known and predicted multi-spanning inhibitory receptor genes in different cell types, in the resting state or after activation. Receptors were considered not expressed (black) when expression was below median overall gene expression in the sample. Data for neutrophils was retrieved from a different source than the other cell types. **B.** Known and predicted multi-spanning receptors were classified into different functional categories based on changes, or lack thereof, in expression after stimulation. Threshold receptors were expressed at resting state, and did not change after activation (i.e., change in expression < 0.5 log2 fold change). Threshold-negative feedback and threshold-disinhibition receptors were defined by > 0.5 log2 fold change up or downregulation, respectively, in expression after activation. Negative feedback receptors were absent in the resting state, but expressed after activation. **C.** Upset plot showing the number of multi-spanning receptors that are expressed uniquely by individual immune cell subsets, or shared between subsets as indicated by connected circles. Seventy-five genes are expressed in all cell types (not depicted). All receptors shown are shared with at least one other immune cell subset or are uniquely expressed by a cell subset. Cell subsets without uniquely expressed putative inhibitory receptors, i.e., B cells and T cells, were excluded from the panel for clarity.

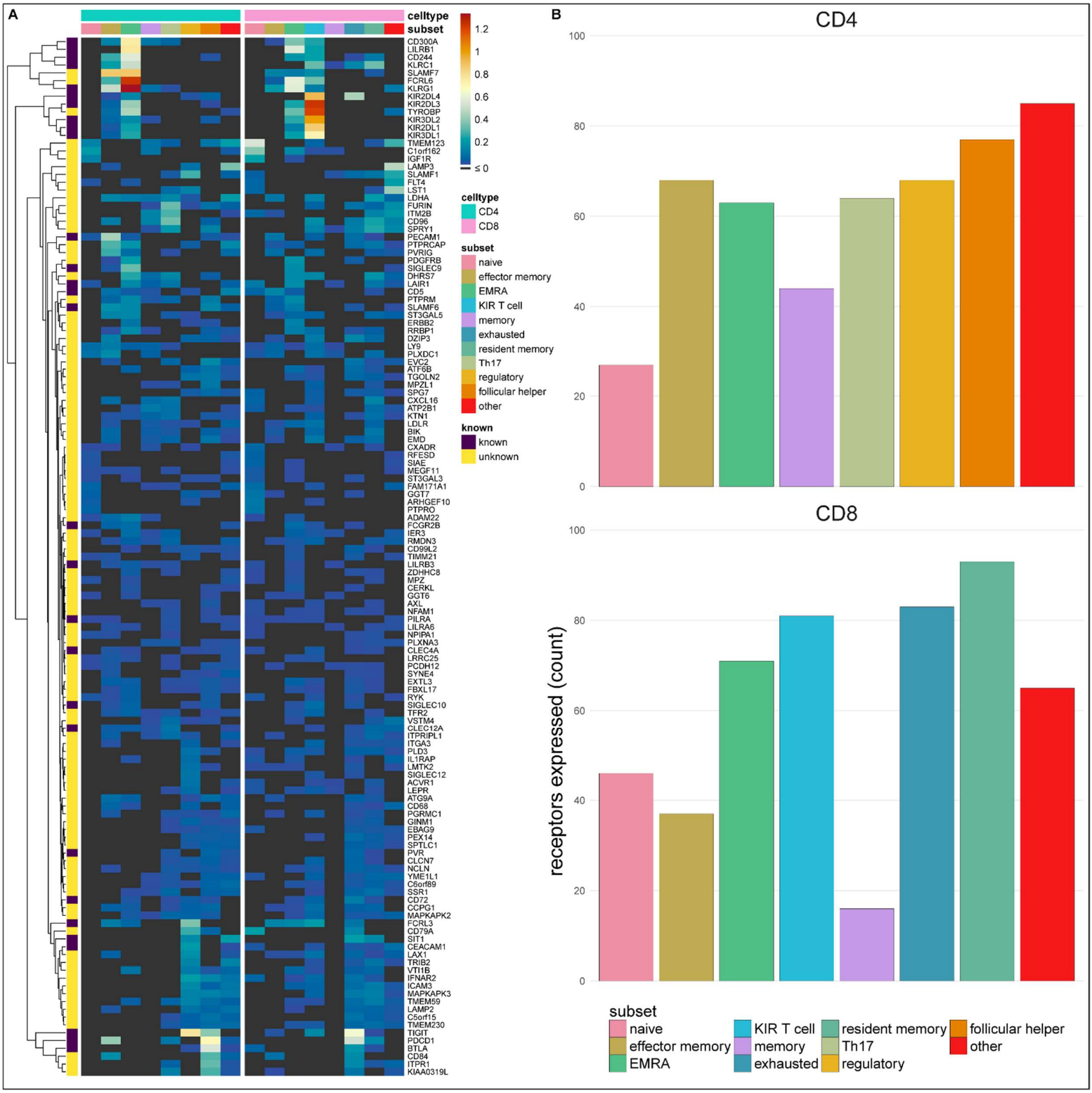
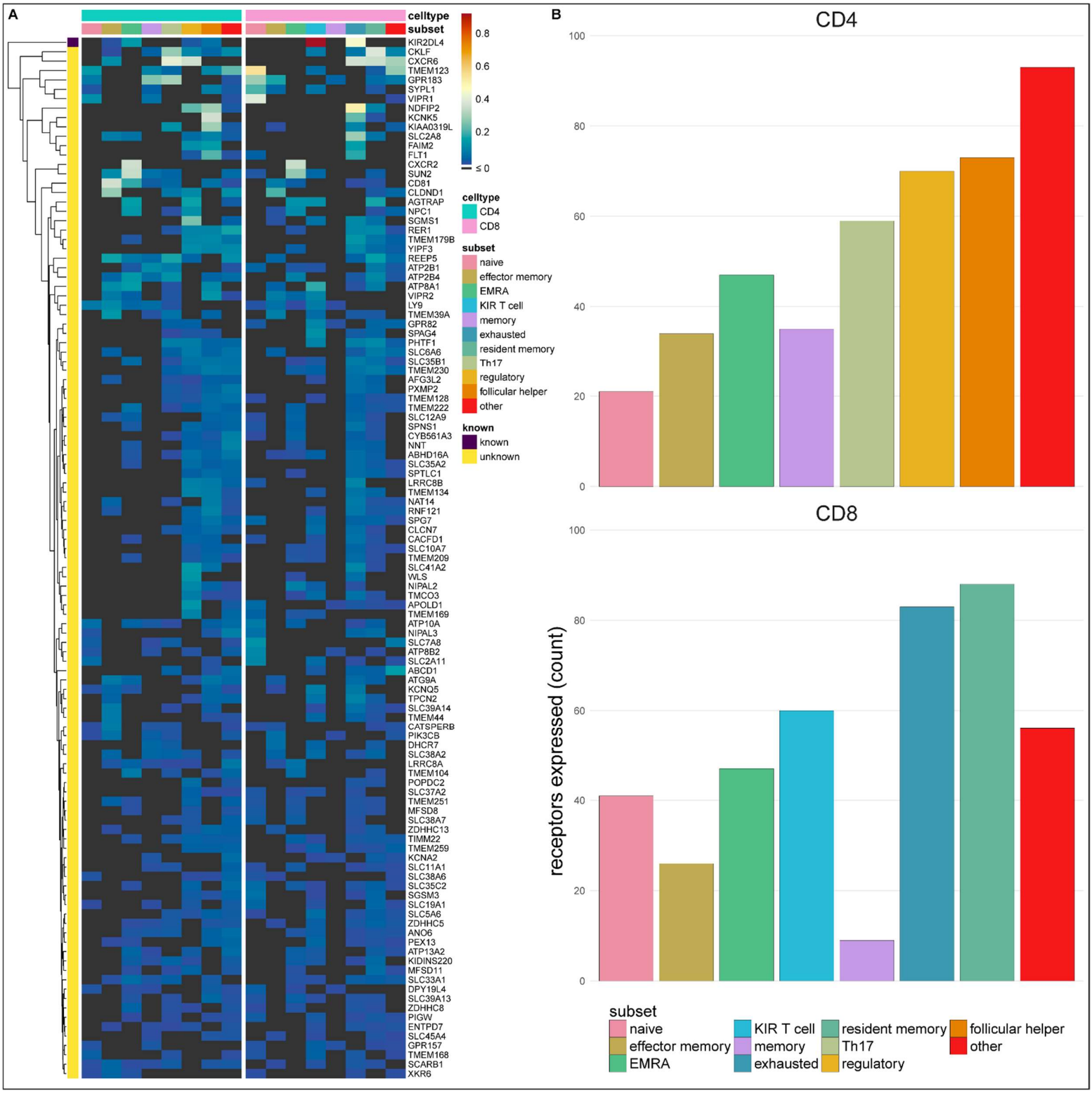
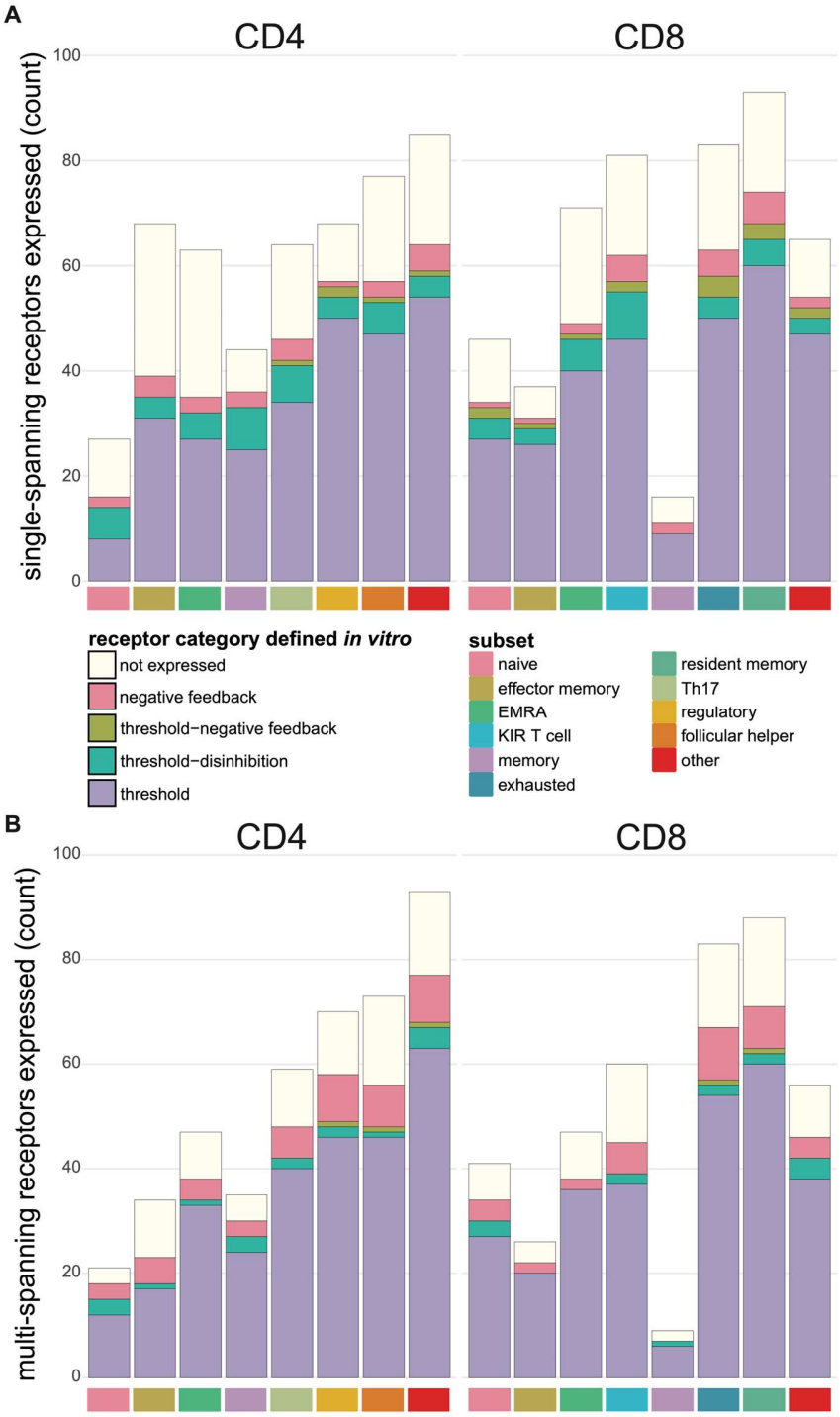


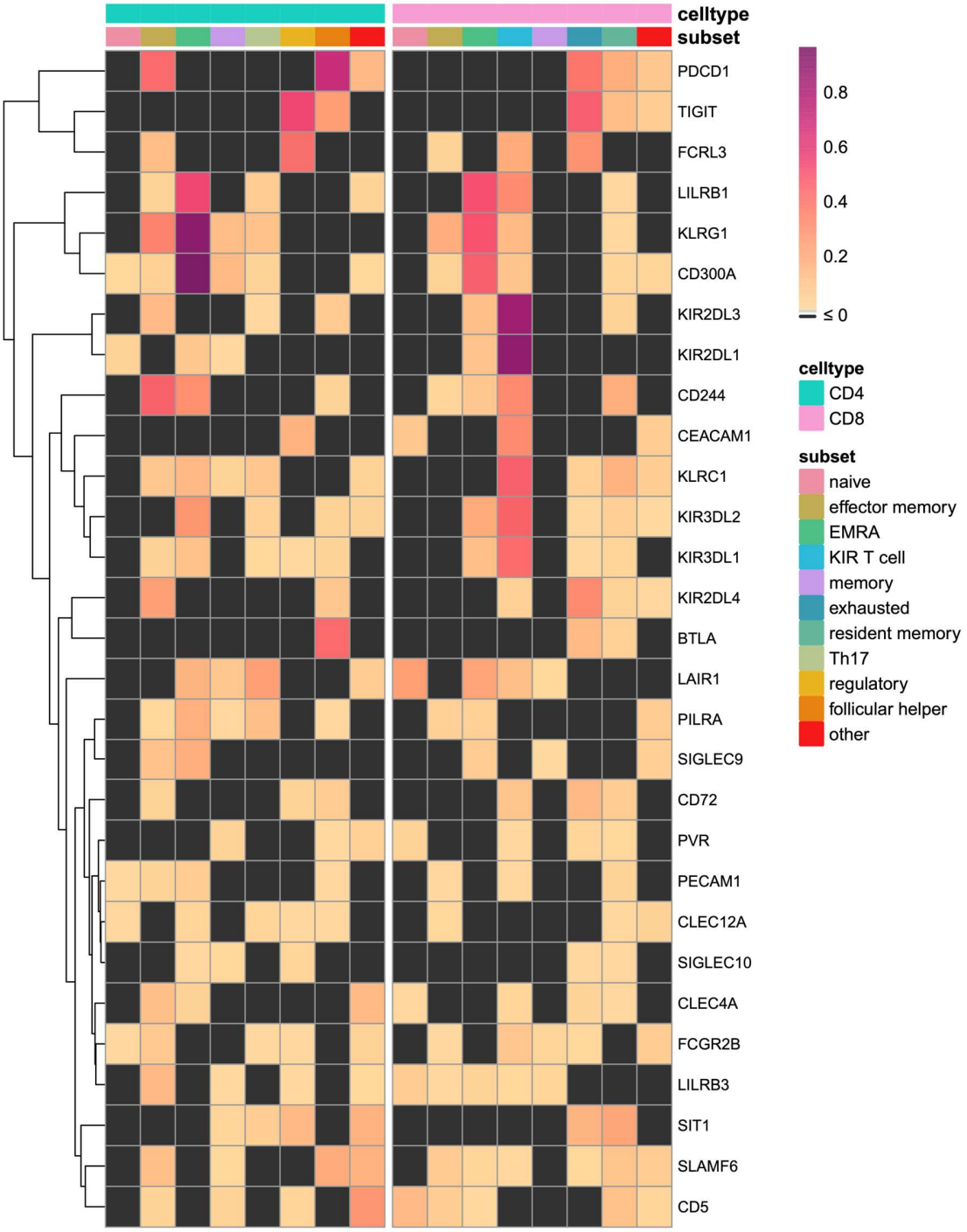
Figure 3. Single-spanning predicted inhibitory receptors are expressed across a wide variety of tumour infiltrating T cell subsets. A. Heatmap with row-normalized expression data for known and predicted single-spanning inhibitory receptor genes in different tumour infiltrating T cell subsets. Receptors were considered not expressed in a T cell subset (black) when expression was below median across all subsets. **B.** Number of inhibitory receptor genes expressed by different $CD4^+$ T cell subsets (upper graph) and $CD8^+$ T cell subsets (lower graph).



Supplementary Figure 3. Multi-spanning predicted inhibitory receptors are expressed across a wide variety of tumour infiltrating T cell subsets. A. Heatmap with row-normalized expression data for known and predicted multi-spanning inhibitory receptor genes in different tumour infiltrating T cell subsets. Receptors were considered not expressed in a T cell subset (black) when expression was below median across all subsets. **B.** Number of inhibitory receptor genes expressed by different $CD4^+$ T cell subsets (upper graph) and $CD8^+$ T cell subsets (lower graph).



Supplementary Figure 4. Expression of functional categories of inhibitory receptors across tumour infiltrating T cell subsets. A. Functional categories of known and novel single-spanning inhibitory receptors across tumour infiltrating T cell subsets. **B.** Functional categories of known and novel multi-spanning inhibitory receptors across tumour infiltrating T cell subsets. For **A** and **B**, the functional categorization was based on the *in vitro* RNA sequencing datasets, and was applied to the tumour infiltrating T cell subset expression.



Supplementary Figure 5. Known inhibitory receptors are expressed in tumour infiltrating T cell subsets of melanoma patients. Heatmap with row-normalized expression data for known inhibitory receptor genes in different tumour infiltrating T cell subsets of melanoma patients. Receptors were considered not expressed in a T cell subset (black) when expression was below median across all subsets.

Tables

Supplementary Table 1a. Overview of predicted immune inhibitory receptor genes bearing an intracellular ITIM -> find attached separately due to size limitations

Table 1. Genes encoding previously described ITIM-bearing inhibitory receptors

BTLA	CLEC4A	KIR3DL2	NCR2	SIGLEC11
CD22	FCGR2B	KIR3DL3	PDCD1	SIRPA
CD244	FCRL2	KLRC1	PECAM1	SIT1
CD300A	FCRL3	KLRG1	PILRA	SLAMF6
CD300LF	FCRL4	LAIR1	PVR	TIGIT
CD33	FCRL5	LILRB1	SIGLEC5	TREML1
CD5	KIR2DL1	LILRB2	SIGLEC6	VSIG4
CD72	KIR2DL2	LILRB3	SIGLEC7	VSTM1
CEACAM1	KIR2DL3	LILRB4	SIGLEC8	
CLEC12A	KIR2DL4	LILRB5	SIGLEC9	
CLEC12B	KIR3DL1	*MPIG6B*	SIGLEC10	

Table 2. Number of samples for different resting and activated immune cell subsets

cell type	subsets	resting	activated	stimulation	duration
Neutrophils		3	10	TSLP / *S. aureus*	4h / 24h
Monocytes		4	8	LPS	24h
NK cells		6	6	IL-2	24h
B cells	naïve, memory, plasmablast	13	10	anti-IgG/M + IL-4	24h
CD4 T cells	naïve, effector, memory, regulatory, Tfh	37	40	dynabeads + IL-2	24h
CD8 T cells	naïve, effector, memory	15	16	dynabeads + IL-2	24h

Supplementary Table 2. Number of multi-spanning receptors in different functional categories for each immune cell subset

cell type	not expressed	negative feedback	threshold negative-feedback	threshold-disinhibition	threshold
Neutrophils	79	4	1	3	110
Monocytes	97	10	2	16	72
NK cells	107	7	1	3	79
B cells	102	13	0	7	75
CD4 T cells	106	13	2	5	71
CD8 T cells	106	11	2	6	72

Table 3. Number of single-spanning receptors in different functional categories for each immune cell subset

cell type	not expressed	negative feedback	threshold-negative feedback	threshold-disinhibition	threshold
Neutrophils	78	6	3	4	124
Monocytes	95	26	1	21	72
NK cells	104	17	2	11	81
B cells	104	10	0	16	85
CD4 T cells	116	8	2	16	73
CD8 T cells	109	9	4	15	78