Gene Model Correction for PVRIG and TIGIT Single Cell Sequencing Data Enables Accurate Detection and Study of their Functional Relevance

Sergey Nemzer[1], Niv Sabath[1], Assaf Wool[1], Zoya Alteber[1], Hirofumi Ando[2], Amanda Nickles-Fader[2], Tian-Li Wang[2], Ie- Ming Shih[2], Drew Pardoll[2], Sudipto Ganguly[2], Yaron Turpaz[1], Zurit Levine[1], Roy Granit[1]*

[1]Compugen LTD, Holon, Israel [2]Bloomberg-Kimmel Institute for Cancer Immunotherapy, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine; Baltimore, MD, USA

Learn more in our bioRxiv publication

compugen
FROM CODE TO CURE®

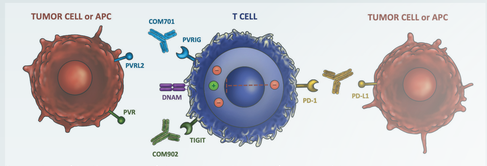
Abstract 4292

Introduction

Background: Single cell RNA sequencing (scRNA-seq) has gained increased popularity in recent years and has revolutionized the study of cell populations; however, this technology presents several caveats regarding specific gene expression measurement. PVRIG and TIGIT are two immune inhibitory checkpoints, previously discovered by our computational platform. While studying their expression in scRNAseq data we noted that PVRIG and murine TIGIT suffer from poor detection in 10x Chromium data due to miss annotations in the default GENCODE gene model, typically used in the analysis of such data.

Methods and finding: We generated a corrected reference genome, by removing the faulty read-through in the case of PVRIG and by extending the TIGIT 3' UTR and demonstrate that by employing these changes we can correctly capture genuine expression levels of these checkpoints, which align with our findings at the protein level using FACS and CITEseq. Furthermore, we show that specialized read multimap algorithms such as RSEM and STARsolo can also partially improve the detection of PVRIG.

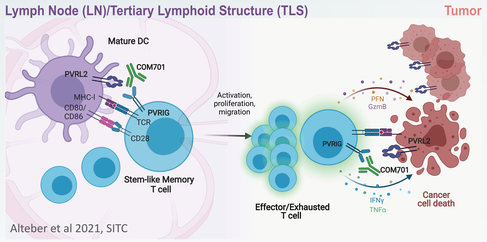
DNAM axis potential to be a game changer in the fight against cancer



Support combination approach to overcome immunotherapy resistance

- PVRIG and TIGIT discovered by Compugen's discovery platform
- DNAM axis – two parallel and complementary inhibitory pathways (PVRIG & TIGIT)
- Potential intersection between PVRIG/TIGIT and PD-1 pathway

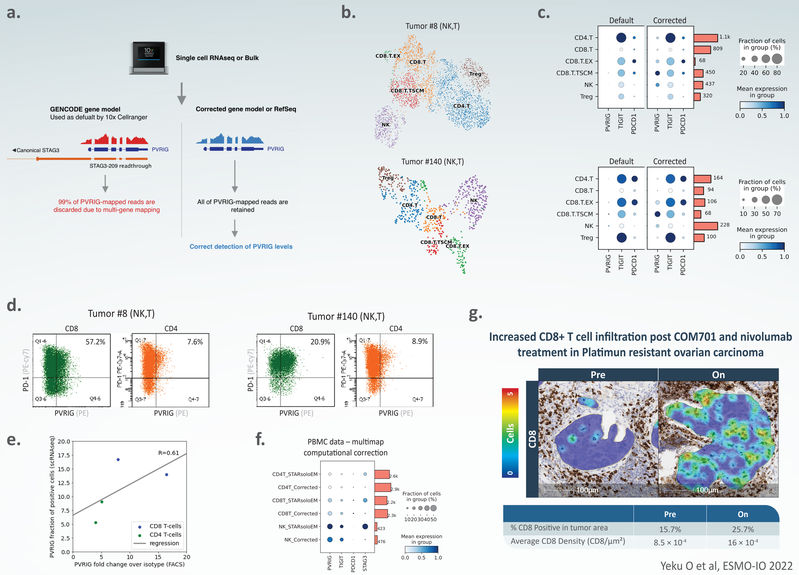
Proposed mechanism of action for COM701 (anti-PVRIG)



Alteber et al 2021, SITC

- PVRL2 broadly expressed in PD-L1 high and low tumors
- Based on our computational analysis and experimental data we believe that PVRIG+ Tscm interact with PVRL2+ DCs in the LN or TLS
- Blocking PVRIG is potentially unique in generating new waves of T cells to infiltrate the TME
- Hence, COM701 may sensitize cold tumors to PD-1/TIGIT blockade

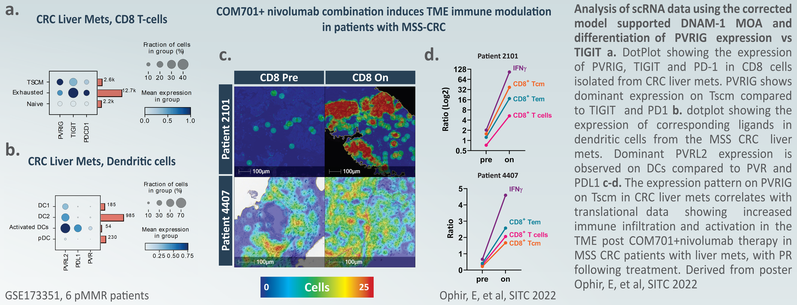
Corrected gene model enables correct detection of PVRIG and validates expression in Tscm in ovarian tumor samples



Yeku O et al, ESMO-IO 2022

Corrected gene model enables the detection of PVRIG. a. Summarizing illustration; using the GENCODE gene model reads are discarded due to multi-mapping, following the correction of the gene model PVRIG expression is correctly measured. **b.** UMAPs showing the population distribution in two ovarian tumor samples subset for T and NK cells. **c.** Dotplot showing the expression of PVRIG and additional T-cell checkpoints in ovarian tumor samples processed using the default or corrected gene model subset for T and NK cells. Bars indicate number of cells in each sample. **d.** Matched FACS analysis showing the expression of PVRIG and PD-1 in CD8 or CD4 T cells. **e.** Correlation between PVRIG-positive fraction as measured using FACS (X-axis) to its parallel in the scRNA-seq data following the gene model correction. Linear-regression line is shown in gray as well as the Pearson correlation coefficient. **f.** DotPlot showing the expression of PVRIG, TIGIT, PD-1 and STAG3 in NK/T populations in the PBMC10K dataset processed using the corrected gene model or using STARsolo-EM method. **g.** The expression pattern on PVRIG on Tscm in Ovarian cancer correlates with clinical data showing increased CD8. Image shows IHC staining of patient with HGSC with partial response (PR) following treated with COM701 and nivolumab, who received 7 prior lines of anti cancer treatment (including nivolumab**)**. CD8 intensity was calculated using Halo Density heatmap algorithm. Derived from poster Yeku O et al, ESMO-IO 2022.

Gene model correction allows the study of PVRIG in tumor samples, supporting the mechanism of action

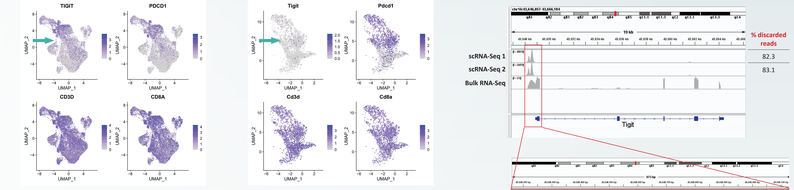


GSE173351, 6 pMMR patients

Ophir, E, et al, SITC 2022

Analysis of scRNA data using the corrected model supported DNAM-1 MOA and differentiation of PVRIG expression vs TIGIT a. DotPlot showing the expression of PVRIG, TIGIT and PD-1 in CD8 cells isolated from CRC liver mets. PVRIG shows dominant expression on Tscm compared to TIGIT and PD1 **b.** dotplot showing the expression of corresponding ligands in dendritic cells from the MSS CRC liver mets. Dominant PVRL2 expression is observed on DCs compared to PVR and PDL1 **c-d.** The expression pattern on PVRIG on Tscm in CRC liver mets correlates with translational data showing increased immune infiltration and activation in the TME post COM701+nivolumab therapy in MSS CRC patients with liver mets, with PR following treatment. Derived from poster Ophir, E, et al, SITC 2022

Mouse TIGIT is poorly detected due to miss-annotation of it's 3' UTR



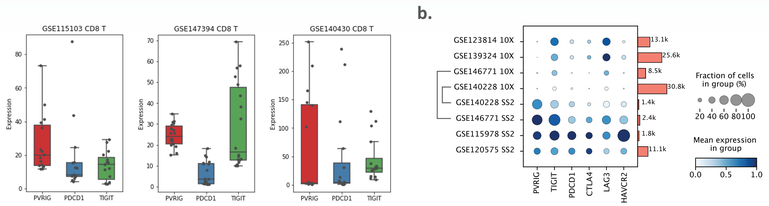
RNA-Seq analysis have shown that TIGIT is highly expressed on PD-1+ exhausted CD8 T cells (a). In contrast, scRNA-Seq mouse datasets show very low TIGIT expression in these cells(b)

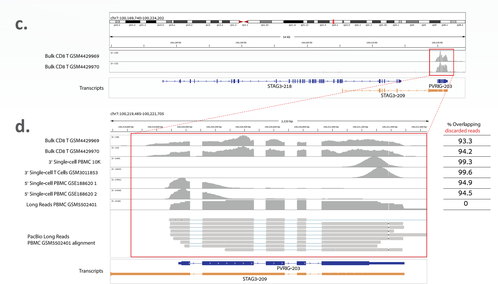
Mouse TIGIT expression pattern initially did not match that in human, possibly due to miss annotation of it's UTR a. UMAP of human CD8 T cells colored by the expression of TIGIT, PD-1 (PDCD1), CD3D, and CD8A **b.** UMAP of mouse CD8 T cells colored by the expression of TIGIT, PD-1 (Pdcd1), CD3D, and CD8A. **c.** Integrative Genomics Viewer (IGV) of the mapped reads around TIGIT coordinates in the mouse genome. In contrast to bulk RNA-Seq, most reads sequenced by scRNA-Seq are from the 300 bp at the 3' end of the transcript. Hence, the miss-annotation of TIGIT 3' UTR led to the observed low level of TIGIT expression in scRNA-Seq. In support of the extended UTR, this downstream peak consists of a predicted alternative ploy-A site (Wang et al., NAR, 2018).

Integrative Genomics Viewer (IGV) of the mapped reads around Tigit gene coordinates in the mouse genome.

PVRIG is poorly detected in 10x chromium data due to overlap with STAG 3

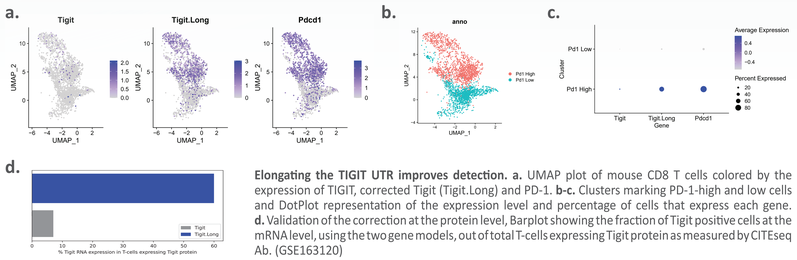
We noted that in 10x Chromium datasets, PVRIG is poorly detected due to multi-mapping of reads to STAG3-209, a read-through transcript with poor support levels that was described only once in the literature, and it not detected in 10x or PacBio long-read sequencing.





Unlike bulk and smart-seq PVRIG is poorly detected in 10x Single cell data due to overlap with STAG3-209. a. Boxplots showing that PVRIG, TIGIT and PD-1 are expressed across several CD8 T-cell datasets as measured using bulk RNA-seq. GSE140430 isolated CD8 T-cells from TME, rest are from blood. **b.** Dotplot showing the expression of PVRIG and additional T-cell checkpoints in several Smart-seq2 (SS2) datasets (bottom) but PVRIG is not detected in 10x datasets (top). Line connect datasets from the same studies conducted with the two technologies. Bars indicate number of cells in each sample **c.** The STAG3-PVRIG genomic locus. Histograms feature the distribution of reads in bulk RNA-seq data of isolated CD8 T-cells highlighting that PVRIG is expressed in such cells (red box). Gene structures appear below in blue. **d.** Closer view of PVRIG locus. Histograms highlight the distribution of reads in the indicated datasets. **e.** Fraction of reads that map to both PVRIG and STAG3-209 and are thus discarded.

Extending the mouse TIGIT UTR annotation leads to improved detection



Elongating the TIGIT UTR improves detection. a. UMAP plot of mouse CD8 T cells colored by the expression of TIGIT, corrected Tigit (Tigit.Long) and PD-1. **b-c.** Clusters marking PD-1-high and low cells and DotPlot representation of the expression level and percentage of cells that express each gene. **d.** Validation of the correction at the protein level, Barplot showing the fraction of Tigit positive cells at the mRNA level, using the two gene models, out of total T-cells expressing Tigit protein as measured by CITEseq Ab. (GSE163120)

Conclusion

Our study provides means to better interrogate the expression of PVRIG and murine TIGIT in scRNA-seq and emphasize the importance of optimizing gene models and alignment algorithms to enable accurate gene expression measurement in scRNA-seq and bulk sequencing. Additionally, our results will enable detailed study of the expression of immune checkpoints in clinical and pre-clinical studies towards the development of cancer immunotherapy treatments.