



An alternatively spliced PD-L1 isoform PD-L1 Δ 3, and PD-L2 expression in breast cancers: implications for eligibility scoring and immunotherapy response

Didem Naz Dioken¹ · Ibrahim Ozgul¹ · Irem Yilmazbilek¹ · Mustafa Cengiz Yakicier² · Ezgi Karaca^{3,4} · Ayse Elif Erson-Bensan¹

Received: 15 July 2023 / Accepted: 5 September 2023 / Published online: 28 September 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

Targeting PD-1/PD-L1 has shown substantial therapeutic response and unprecedented long-term durable responses in the clinic. However, several challenges persist, encompassing the prediction of treatment effectiveness and patient responses, the emergence of treatment resistance, and the necessity for additional biomarkers. Consequently, we comprehensively explored the often-overlooked isoforms of crucial immunotherapy players, leveraging transcriptomic analysis, structural modeling, and immunohistochemistry (IHC) data. Our investigation has led to the identification of an alternatively spliced isoform of *PD-L1* that lacks exon 3 (PD-L1 Δ 3) and the IgV domain required to interact with PD-1. *PD-L1* Δ 3 is expressed more than the canonical isoform in a subset of breast cancers and other TCGA tumors. Using the deep learning-based protein modeling tool AlphaFold2, we show the lack of a possible interaction between PD-L1 Δ 3 and PD-1. In addition, we present data on the expression of an additional ligand for PD-1, *PD-L2*. *PD-L2* expression is widespread and positively correlates with *PD-L1* levels in breast and other tumors. We report enriched epithelial-mesenchymal transition (EMT) signature in high *PD-L2* transcript expressing (PD-L2 > PD-L1) tumors in all breast cancer subtypes, highlighting potential crosstalk between EMT and immune evasion. Notably, the estrogen gene signature is downregulated in ER + breast tumors with high *PD-L2*. The data on PD-L2 IHC positivity but PD-L1 negativity in breast tumors, together with our results on PD-L1 Δ 3, highlight the need to utilize PD-L2 and PD-L1 isoform-specific antibodies for staining patient tissue sections to offer a more precise prediction of the outcomes of PD-1/PD-L1 immunotherapy.

Keywords PD-L1 · PD-L2 · AlphaFold · Immunotherapy · Breast cancer · Immune checkpoint blockade

Introduction

PD-L1 is an Ig-like transmembrane receptor ligand expressed on cell surfaces. Interaction of this ligand with its receptor (PD-1) inhibits T-cell activation and cytokine production. During infection or inflammation, this interaction is essential to prevent autoimmunity. However, PD-L1 is also frequently overexpressed on the surface of different tumors, including lymphoma, melanoma, lung, breast, kidney, ovary, bladder cancers, and glioblastoma [1–4]. These tumor cells expressing PD-L1 protein escape from the immune system through cytotoxic T-cell inactivation, which promotes tumor growth and metastasis [5]. On the other hand, PD-1 is highly expressed in activated T cells, B cells, thymocytes, natural killer (NK) cells, and other antigen-presenting cells (APCs). Hence, PD-1/PD-L1 signal transduction is critical for autoimmunity, antiviral responses, and antitumoral T-cell

Didem Naz Dioken and Ibrahim Ozgul are co-first authors.

✉ Ayse Elif Erson-Bensan
erson@metu.edu.tr

- ¹ Department of Biological Sciences, Middle East Technical University (METU), Dumlupinar Bly No:1 Universiteler Mah, Cankaya, 06800 Ankara, Türkiye
- ² AQUARIUS/NPG Genetic Diseases Evaluation Center, Kucukbakkalkoy Mah. Kayisdagi Cad. 137/6 Atasehir, Istanbul, Türkiye
- ³ Izmir Biomedicine and Genome Center, Dokuz Eylul University Health Campus, 35340 Balcova, Izmir, Türkiye
- ⁴ Izmir International Biomedicine and Genome Institute, Dokuz Eylul University, 35340 Balcova, Izmir, Türkiye

responses [6]. Given these vital roles, monoclonal antibodies directed against PD-1 or PD-L1 for immune checkpoint blockade immunotherapy have emerged as a promising and effective treatment strategy in a subset of advanced cancer patients [7–9]. Targeting PD-1/PD-L1 has shown substantial therapeutic response and unprecedented long-term durable responses in the clinic; however, critical challenges remain to be addressed. For example, the reasons for the lack of complete clinical response or resistance to immune checkpoint inhibitors in some patients are unclear.

Currently, most focus is on PD-L1 binding to PD-1 for immune checkpoint blockade immunotherapy. Diagnostic antibodies score PD-L1 levels to determine patient eligibility for immunotherapy. However, PD-L1 isoforms and a second ligand (PD-L2) [10] are often overlooked within this great potential to treat aggressive tumors. PD-L2 expression is already reported for several tumor types, including head and neck squamous cell carcinoma [11], lung squamous cell carcinoma [12], renal cell carcinoma [13], and pancreatic ductal adenocarcinoma [14]. Here we focus on PD-L1 isoforms and PD-L2 in breast cancers and propose that disregarding these variants may hinder accurate eligibility scoring for immunotherapy and effective treatment of patients.

Methods

Expression data

Tumor expression data were retrieved from public domain resources. TPM data (RSEM) in the Genotype-Tissue Expression Project (GTEx) (<https://gtexportal.org>) and The Cancer Genome Atlas (TCGA) Genomic Data Commons Data Portal (GDC Data Portal) (<https://portal.gdc.cancer.gov>) were downloaded from UCSC Xena, Xena Toil RNA-Seq Recompute Compendium (<https://toil.xenahubs.net>) (Jan.16, 2021). The batch effect caused by different computational analyses is eliminated because UCSC Xena contains data re-analyzed by the same RNA-Seq pipeline for TCGA and GTEx samples. The clinical data for TCGA-BRCA samples containing PAM50 status (Luminal A (estrogen receptor (ER)-positive and progesterone receptor (PR)-positive, HER2-negative), Luminal B (ER-positive and HER2-negative), HER2-enriched (ER-negative, PR-negative, and HER2-positive), and Triple-negative or basal-like breast cancer (ER/PR/HER2-negative) were downloaded from TCGA by the TCGAAbiolinks R package version 2.20.0 [15].

PD-L1 Δ 3 and PD-L2 expression analysis

To determine the number of patients exhibiting higher expression of the PD-L1 Δ 3 or PD-L2 transcripts, we extracted RNA-Seq data in the form of RSEM TPM values

using the UCSC Xena tool and compared the isoform-level RSEM TPM data for isoforms.

Differential gene expression

RNA expression data for cancer patients were obtained from the TCGA database through the XENA Toil web interface. Patients with available PD-L1 and PD-L2 expression data were included in the study. Based on the PD-L2/PD-L1 ratio, patients were classified into two groups: those with a ratio greater than or equal to 1 (PD-L2/PD-L1 \geq 1) and those with a ratio less than 1 (PD-L2/PD-L1 $<$ 1). The RNA-seq data of TCGA was used on the cBioPortal platform to determine differentially regulated genes in the PD-L2 $>$ PD-L1 group. Genes with a student *t* test *p* value $<$ 0.05 and fold-change $>$ 1.5 or $<$ 0.6 were considered differentially expressed between the two patient groups (PD-L2 $>$ PD-L1 vs. PD-L2/PD-L1 $<$ 1). The “GSEAPreranked” tool in Gene Set Enrichment Analysis (GSEA) software version 4.2.1 was used to compute enriched biological pathways for tumors grouped according to PD-L2/PD-L1 expression ratios [16]. The weighted GSEA analysis was performed with 1000 permutations using human gene symbols for Hallmark gene sets. mRNA expressions of PD-L1 and PD-L2 were also analyzed in the METABRIC (Molecular Taxonomy of Breast Cancer International Consortium) dataset [17]. We grouped patients using their PAM50 subtype and the expression level of PD-L2 and PD-L1. The normalized mRNA expression data for PD-L1 (ILMN_1701914) and PD-L2 (ILMN_2159272) from European Genome-Phenome Archive (EGAS00000000083) and the clinical data from cBioPortal were used. The DEGs between the two groups (PD-L2/PD-L1 \geq 1 vs. PD-L2/PD-L1 $<$ 1) were determined and analyzed by GSEA.

Protein structure and modelling

Human PD-1, PD-L1, and PD-L2 sequences were retrieved from the UniProt Knowledgebase [18]. The crystal structures of PD-L1 and PD-1 were retrieved from Protein Data Bank [19] through accession ids 3BIS [20] and 3RRQ (Supplementary Fig. 1). For modeling the protein isoforms with unknown structures and their interactions, the deep learning-based protein modeling tool AlphaFold2 (AF2) [21] was used through the ColabFold platform [22]. AF2 evaluates the per-residue confidence score (pLDDT) between 0 and 100, the predicted aligned error (PAE) rate for each residue, and the predicted structure accuracy (predicted TM (pTM)), ranging between 0 to 1. The interfaces of protein interactions were analyzed with the PDBePISA web server [23]. PyMOL was used for the visualization of the protein structures (The PyMOL Molecular Graphics System, Version 2.0,

Schrödinger, LLC, 1 December 2022). IBS 2.0 was used to illustrate the protein domain organizations [24].

IHC for PD-L1 and PD-L2

PD-L1 and PD-L2 immunohistochemistry results were taken from the Human Protein Atlas (proteintlas.org) [25]. CAB076385 antibody was used for PD-L1, and HPA013411 was used for PD-L2 staining.

Results

PD-L1 and PD-L1 isoforms

PD-L1 (CD274, ENSG00000120217.13) gene maps to Chromosome 9: 5,450,503–5,470,566 and generates several mRNA isoforms. The canonical mRNA isoform (ENST00000381577.3) is 3622 nucleotides long, has seven exons, and codes for a 290 amino-acid long protein with N-terminus Immunoglobulin Variable domain (IgV), Immunoglobulin Constant-2 set (IgC2) domain, transmembrane and cytoplasmic domains at the C-terminus (Fig. 1a). The N-terminal IgV domain of PD-L1 is responsible for binding to the IgV domain of PD-1 [26, 27].

Along with this well-known canonical *PD-L1* transcript, a second *PD-L1* mRNA isoform (ENST00000381573.8, 3.3 kb) is also expressed. Because of alternative splicing, this isoform lacks the third exon and encodes a 176 amino acid protein isoform (NP_001254635) which lacks the IgV domain (hereon referred to as PD-L1 Δ 3) (Fig. 1a).

First, to understand the functional relevance of not having the IgV domain for the PD-1 interaction, we investigated the structure and function of the PD-L1 Δ 3 protein isoform. The crystal structure of the PD-1: PD-L1 complex was available through the PDB (ID:3BIK) [19]. The individual structures of PD-L1 and PD-1 were previously resolved (PDB IDs: 3BIS [20], and 3RRQ, Supplementary Fig. 1). PD-1 and PD-L1 interact through the front and the side of (IgV) domains, so the interface residues are located on their IgV domains (Supplementary Table 1). To understand the implication for the loss of the IgV domain in PD-L1 Δ 3 for the PD-1 interaction, we used the deep learning-based protein modeling tool AlphaFold2 (AF2) [21] (Supplementary Fig. 1). As a result, the lack of IgV domain in the PD-L1 Δ 3 resulted in a significantly low confidence score at the interaction surface with PD-1, compared with the high score of PD-1-IgV: PD-L1-IgV, calculated as a reference (Fig. 1b). In addition, the important salt bridges (SB1: D85-F19 and SB2: E136-R125) between the PD-1-IgV and PD-L1-IgV interface are completely lost in the PD-1-IgV: PD-L1 Δ 3-IgC2 complex model (Fig. 1c). Based on these results, we suggest that PD-L1 Δ 3 protein does not bind to PD-1.

Next, we were curious whether the *PD-L1 Δ 3* transcript is expressed in tumors. Within the TCGA (The Cancer Genome Atlas) tumor types, we determined patients with higher levels of *PD-L1 Δ 3* compared to the canonical *PD-L1* transcript (PD-L1 Δ 3 > PD-L1) (Fig. 2a). Out of 33 different malignancies, 20 cancer types had 10% or more patients with higher PD-L1 Δ 3 > PD-L1 transcript expression. A breast cancer-specific graph (n = 86) shows *PD-L1 Δ 3* and *PD-L1* expression levels (Fig. 2b).

These results indicated that *PD-L1 Δ 3* transcript levels exceed the canonical *PD-L1* transcript expression in a group of breast cancers and other malignancies. Hence, it is reasonable to expect that PD-L1 Δ 3 protein is present in tumors and that anti-PD-L1 antibodies recognizing common epitopes would bind to both protein isoforms. Consequently, unspecific detection of the PD-L1 Δ 3 protein isoform could be problematic in immunotherapy eligibility tests based on PD-L1 staining. Of note, the functional role of this PD-L1 protein isoform without the IgV domain remains to be investigated.

A second ligand for PD-1; PD-L2

We continued to look into additional isoform-level complexities that may be involved in immunotherapy checkpoint responses. We focused on a second ligand of PD-1; PD-L2 (Programmed Cell Death-1, Ligand 2, *PDCD1LG2*, a.k.a. *CD273*). *PD-L2* gene maps distal to *PD-L1* on Chromosome 9: 5,510,531–5,571,282 (hg38) and is known to generate only one mRNA transcript. In addition to the physical proximity of *PD-L2* and *PD-L1* genes, previous work showed that both genes are co-regulated [27].

PD-L2 mRNA isoform (ENST00000397747.5) is 2432 nucleotides long, has seven exons, and codes for a 273 amino-acid long protein with an N-terminus Ig-like V-type domain and a membrane-proximal IgC domain. The N-terminal Ig-like V-domain of PD-L2 is responsible for binding to PD-1, similar to PD-L1 [10]. Numerous structural and biochemical methods confirmed that PD-L2 binding to PD-1 has a stronger affinity than PD-L1 [28]. Notably, current diagnostic or therapeutic antibodies targeting PD-L1 fail to bind to PD-L2 due to a lack of high sequence homology between the two proteins [29].

Given its ability to bind to PD-1, we sought to determine the expression pattern of PD-L2 in breast tumors. Based on RNA-seq data of the TCGA dataset, *PD-L1* and *PD-L2* transcript levels positively correlate in breast cancers as well as other cancer types (Fig. 3a and Supplementary Fig. 2). Next, we wanted to see whether there are tumors that express more *PD-L2* than *PD-L1*. For breast cancer subtypes, basal tumors had (150 out of 173, 87%) the highest ratio of PD-L2/PD-L1 transcript expression, but within all subtypes, most tumors had higher levels of *PD-L2* transcript (Fig. 3b).

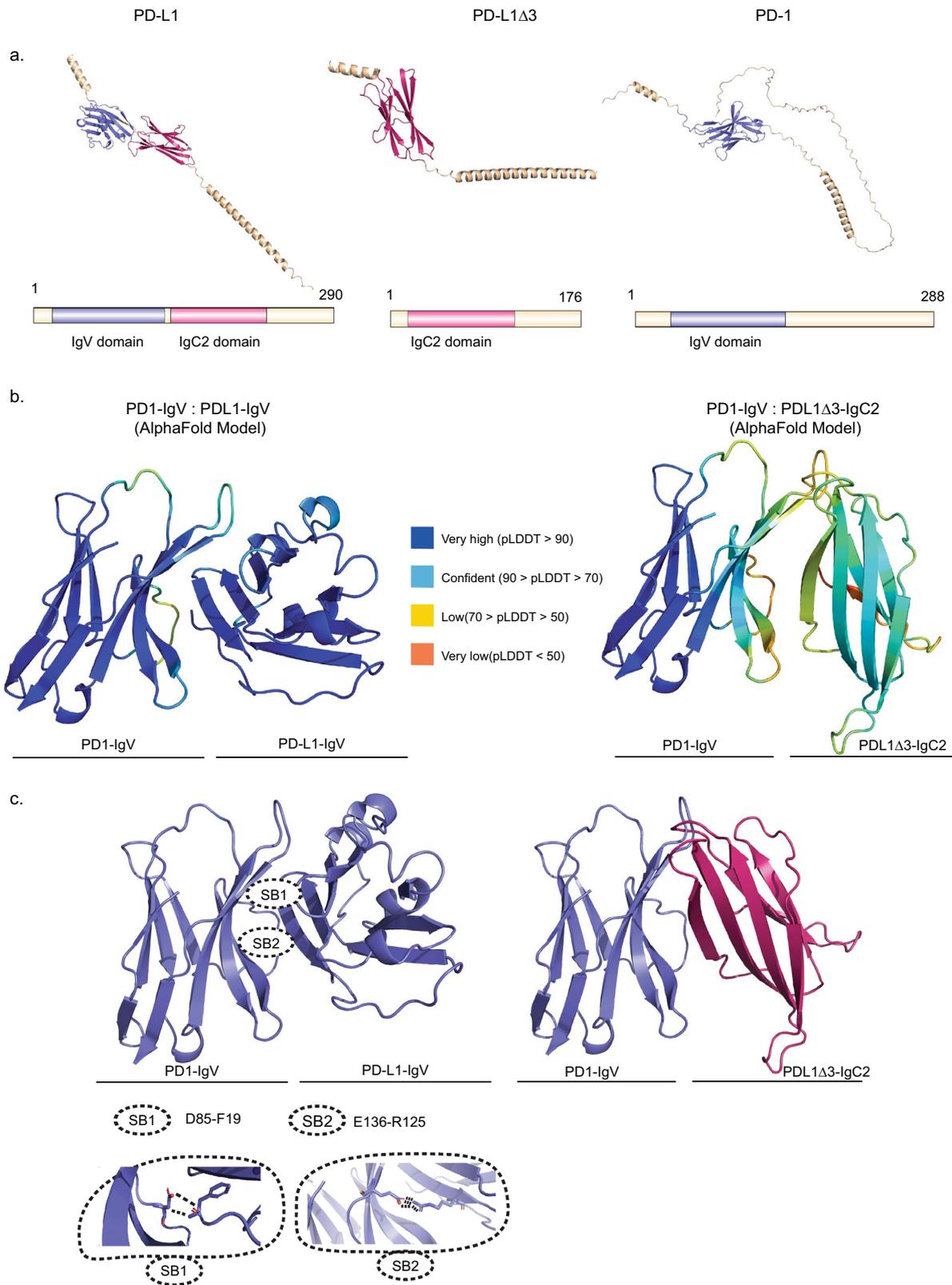


Fig. 1 PD-1/PD-L1 interaction models. **a** Domain organizations of PD-L1, PD-L1 Δ 3 isoform, and PD-1 (IgV in slate and IgC2 in pink). The same color coding for domains is followed for panels. The corresponding full-length protein models were generated with AF2. **b** The best structural models produced by AF2 for PD1-IgV:PDL1-IgV and PD1-IgV:PDL1 Δ 3-IgC2 interactions. The confidence score of models is presented in color-coded pLDDT scores, where yellow and orange indicate low and very low confidence regions. **c** The salt bridge (SB) distribution across PD1-IgV:PDL1-IgV interface is shown. SB1 and SB2 are lost for the PD1-IgV: PDL1 Δ 3-IgC2 complex model

To start understanding the biological relevance of *PD-L2* expression in breast tumors, we grouped tumors based on the infiltration of cytotoxic T lymphocytes (CTL) according to *CD8A*, *CD8B*, *GZMA*, *GZMB*, and *PRF1* expression levels, using the TIDE (Tumor Immune Dysfunction and Exclusion) algorithm [30, 31]. Then we determined *PD-L1* and *PD-L2* expression levels in high and low CTL infiltration groups. *PD-L2* expression levels were comparable to or higher than *PD-L1* levels in all CTL high groups. Of note, within the luminal A (ER/PR+, HER2-) subtype, *PD-L2* expression was more elevated than *PD-L1* in the CTL high group. In contrast, both *PD-L1* and *PD-L2* levels were low in the CTL low tumors (Fig. 3c). These results suggested that high expression of *PD-L2* in the CTL high groups could have functional relevance.

Next, we grouped TCGA BRCA patients based on their PAM50 subtype and the expression level of *PD-L2* and *PD-L1* and then compared two groups (*PD-L2*/*PD-L1* > 1 vs. *PD-L2*/*PD-L1* < 1) of patients to identify differentially expressed genes. Differential expression analysis using cBioPortal resulted in DEGs (differentially expressed genes) according to the criteria of a fold change ($fc > 1.5$

or $fc < 0.6$, and a p value of < 0.05) (Fig. 4a). DEGs were further analyzed by gene set enrichment and ontology tools. GSEA showed that the gene signature for EMT (epithelial-mesenchymal transition) was enriched for the transcript ratio of *PD-L2*/*PD-L1* > 1 breast tumors in all subtypes (Fig. 4b).

An independent METABRIC breast cancer dataset confirmed these results, showing enrichment of an EMT signature for high *PD-L2* expressing tumors (Supplementary Table 2). EMT, a characteristic of tumor cells, is essential for migration, colonization, and metastasis [32]. Evidence of a bidirectional regulation between EMT and immune checkpoint proteins is increasing [33]; hence the co-existence of EMT and high *PD-L2* expression may also represent an opportunity for effective immunotherapy, biomarker development, and drug targeting.

While the EMT signature was enriched in *PD-L2*-high tumors of all breast cancer subtypes, there were also subtype-specific gene signatures. Of note, the KRAS-Signaling-Up signature was enriched in luminal A, and luminal B, breast cancers with higher *PD-L2* expression (ES = 0.53, 0.49, respectively). GSEA Hallmark gene signatures for allograft rejection, inflammatory response, and interferon-gamma response were also enriched along with the KRAS-Signaling-Up gene set in *PD-L2*-high tumors, suggesting inflammation and a favorable tumor immune microenvironment [34]. Luminal A and luminal B tumors with higher *PD-L2* expression were also enriched for the Gene Ontology Biological Process tool's positive T cell selection gene signature (ES = 0.78 and 0.79, respectively) (Supplementary Table 2). Interestingly estrogen response genes were down-regulated in the *PD-L2* > *PD-L1* luminal A and luminal B breast cancer patients (Supplementary Table 2).

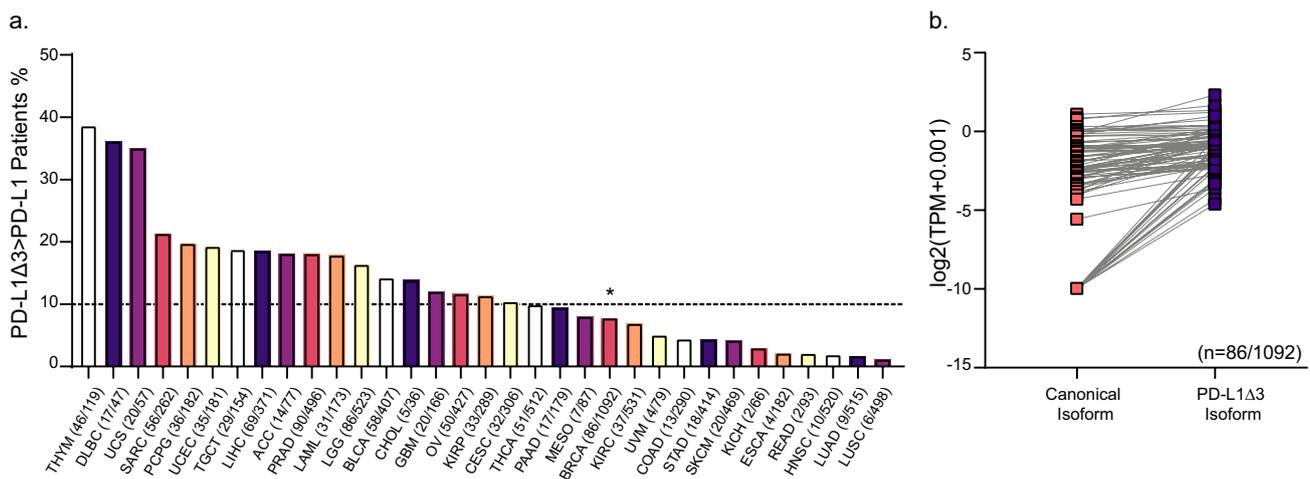


Fig. 2 *PD-L1* Δ 3 isoform and canonical *PD-L1* isoform expression levels across multiple human cancers. **a** Pan-cancer analysis shows the percentage and the number of patients exhibiting higher expression of the *PD-L1* Δ 3 (ENST00000381573) transcript compared to

the canonical *PD-L1* transcript (ENST00000381577). The RSEM TPM values of isoforms for the TCGA dataset were extracted using the UCSC Xena tool. **b** Higher expression of the *PD-L1* Δ 3 transcript compared to the canonical *PD-L1* transcript in breast tumors (n = 86)

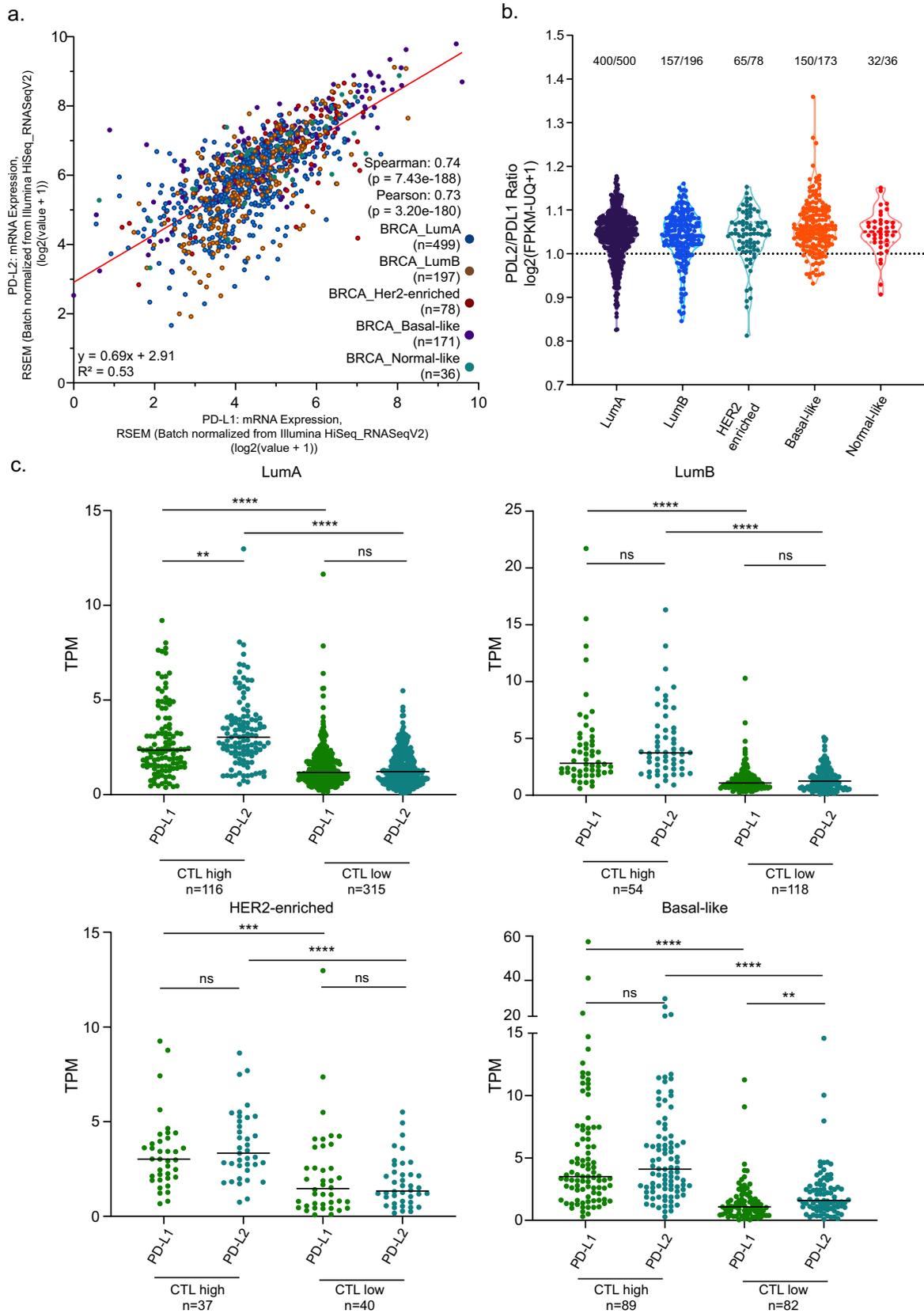


Fig. 3 *PD-L1* and *PD-L2* expression levels correlate in TCGA breast cancers. **a** The x-axis represents the log₂-transformed RSEM gene expression values of *PD-L1*, and the y-axis represents the log₂-transformed RSEM gene expression values of *PD-L2*. Pearson ($r=0.74$), Spearman correlation coefficients ($r=0.73$), and corresponding p-values are shown. The line represents the regression line of the positive correlation. **b** Breast cancer subtypes and high *PD-L2*/*PD-L1* expressing tumors (log₂(FPKM-UQ+1)) are shown., **c**. *PD-L1* and *PD-L2* levels in CTL high and low breast cancers grouped according to PAM50 status (Mann–Whitney test, ns: not significant, ** < 0.01, *** < 0.001, **** < 0.0001). CTL infiltration groups were determined using the TIDE algorithm

For Her2-enriched tumors, high *PD-L2* expressing tumors had downregulated E2F target genes, G2M checkpoint genes, MYC targets, and oxidative phosphorylation gene signatures (Supplementary Table 1). For basal-like tumors, in addition to upregulated EMT genes, Myogenesis, UV-Response, and Angiogenesis Hallmark gene signatures were upregulated (Supplementary Table 2). In contrast, interferon-gamma/alpha response genes, E2F, and MYC target genes were downregulated in high *PD-L2*-expressing tumors (Supplementary Table S2).

These results indicate significant transcriptomic differences in high *PD-L2*/*PD-L1* expressing tumors compared to low *PD-L2*/*PD-L1* tumors. These differences suggest functional implications for high *PD-L2* expression in breast cancers through EMT and subtype-specific pathways. To support the significance of our findings, we provide evidence for *PD-L2* staining in breast cancer patients with undetectable levels of *PD-L1* in the Human Protein Atlas [25] (Fig. 5a, b). 73% of breast tumors (8 out of 11) had medium, 18% (2 out of 11) had low staining intensity for *PD-L2*, whereas none of the 12 samples had *PD-L1* staining (Fig. 5a). Figure 5b shows the same patient samples stained for medium intensity for *PD-L2* whereas *PD-L1* staining was not detected.

Overall, these results show that the expression of *PD-L2* and isoforms of *PD-L1* adds another layer of complexity that may play decisive roles in the ultimate outcomes of immunotherapy blockade.

Discussion

Immunotherapy primarily aims to block PD-1/*PD-L1* interaction to reactivate the immune system to recognize and attack cancer cells. However, multiple factors (e.g., CTL infiltration, DNA repair defects, mutation, neo-antigen load) can affect the success of cancer immunotherapy approaches. None of these factors, including *PD-L1* levels, is sufficient to predict the therapy response. This study looked into transcript-level complexities that may improve the current understanding of PD-1/*PD-L1*-focused treatment strategies. Here, we mainly provide transcript level evidence from patient samples but protein levels and post-translational

modifications such as ubiquitination, glycosylation, phosphorylation acetylation, and palmitoylation are to be considered for *PD-L1* and *PD-L2* positivity in future studies.

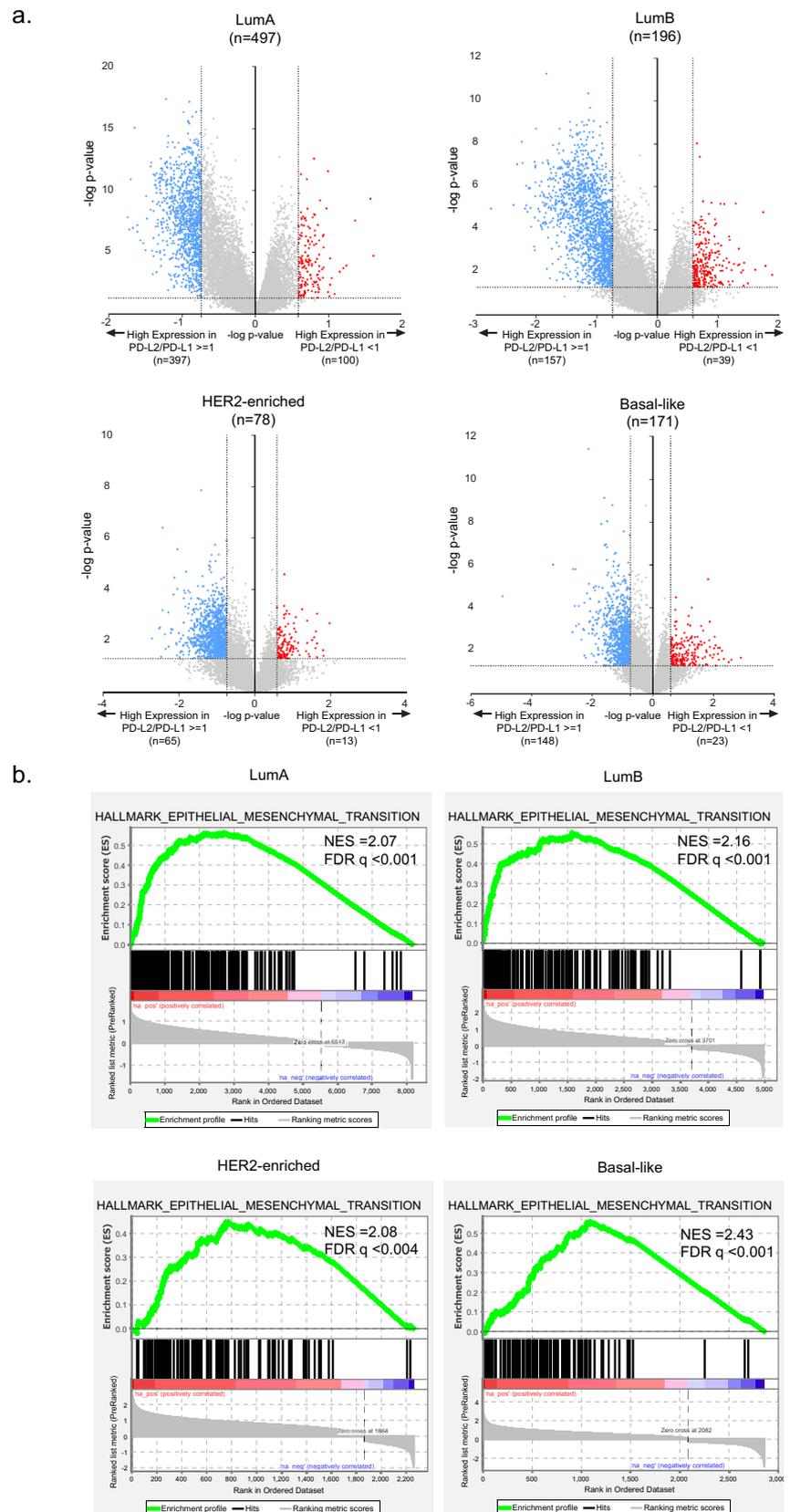
PD-L1 isoforms

The FDA approved different *PD-L1* immunohistochemical assays/antibodies. These assays are generally based on four *PD-L1* antibodies (22C3, 28-8, SP263, SP142). SP142 and SP263 recognize the cytoplasmic domain of *PD-L1*, 22C3 recognizes the IgC2 domain, and 28-8 recognizes both the IgV and the IgC2 domains [35]. Hence, a positive *PD-L1* IHC score could be due to the recognition of the protein isoform translated from *PD-L1*Δ3 alone or with the canonical *PD-L1*, causing false positivity. Only 28-8 antibody is likely to recognize the canonical full-length *PD-L1* protein among these four antibodies. Consequently, unspecific detection of the *PD-L1*Δ3 protein could explain commonly reported staining inconsistencies [35, 36]. Hence, as diagnostic accuracy is a critical parameter for *PD-L1* positivity in patients, the presence and unintentional detection of a non-*PD-1* interacting *PD-L1* protein variant may skew the test results. Earlier, a *PD-L1* splice variant lacking the IgV domain was identified in peripheral blood mononuclear cells [37]. The alternatively spliced isoform of *PD-L1*, missing exon 3, was also recently described in oral squamous cell carcinoma cell lines [38]. This study further provides evidence on enhanced exon 3 inclusion upon IFN-γ treatment in cell lines. Altogether, these findings highlight the importance of understanding the mechanisms controlling gene expression and splicing patterns.

In addition to *PD-L1*Δ3, a soluble form of *PD-L1* (s*PD-L1*) was detected in NSCLC (Non-Small Cell Lung Cancer) patients who did not respond to anti-*PD-L1* treatment [39]. This isoform is a product of alternative splicing and polyadenylation and, when translated, lacks the transmembrane domain. As a result, it is secreted, can bind to *PD-1*, and works as a decoy of anti-*PD-L1* antibodies [39]. Interestingly, this soluble *PD-L1* encoded form has been detected in healthy human serum but is elevated in autoimmune disease and cancer [40–44]. Expression or up-regulation of this isoform, translated into a C-terminus truncated protein (s*PD-L1*) alone or with the full-length protein, may also explain why some *PD-L1*-positive patients do not respond well to anti-*PD-L1* drugs.

In addition to earlier work, large datasets for transcriptome and proteome level analysis in normal tissues and patient samples are now available. Together with previous work, we highlight to need to look into *PD-L1* isoforms that may have functional implications.

Fig. 4 Differential expression analysis results for patients with different PD-L2/PD-L1. **a** Volcano graphs of DEGs in PD-L2/PD-L1 > 1 v.s. PD-L2/PD-L1 < 1 tumors of the luminal A, luminal B, HER2-enriched, and basal-like breast cancer subtypes of the TCGA dataset. DEGs were selected based on fold-change (> 1.5 or < 0.6) and *p* value (student *t* test *p* value < 0.05) criteria. **b**. Gene set enrichment analysis (GSEA) was performed to identify enriched biological pathways and gene ontology terms in high PD-L2 expressing tumors. Enrichment plots for “Hallmark-Epithelial Mesenchymal Transition” (EMT) gene sets are shown for each breast cancer subtype. The y-axis represents the enrichment score (ES), and the curves represent the running sum of ESs. The x-axis shows the rank positions of gene set members representing EMT. Vertical lines indicate the position of individual genes in the ranked list. Normalized enrichment scores (NES) and false discovery rates (FDR) are indicated on the graphs



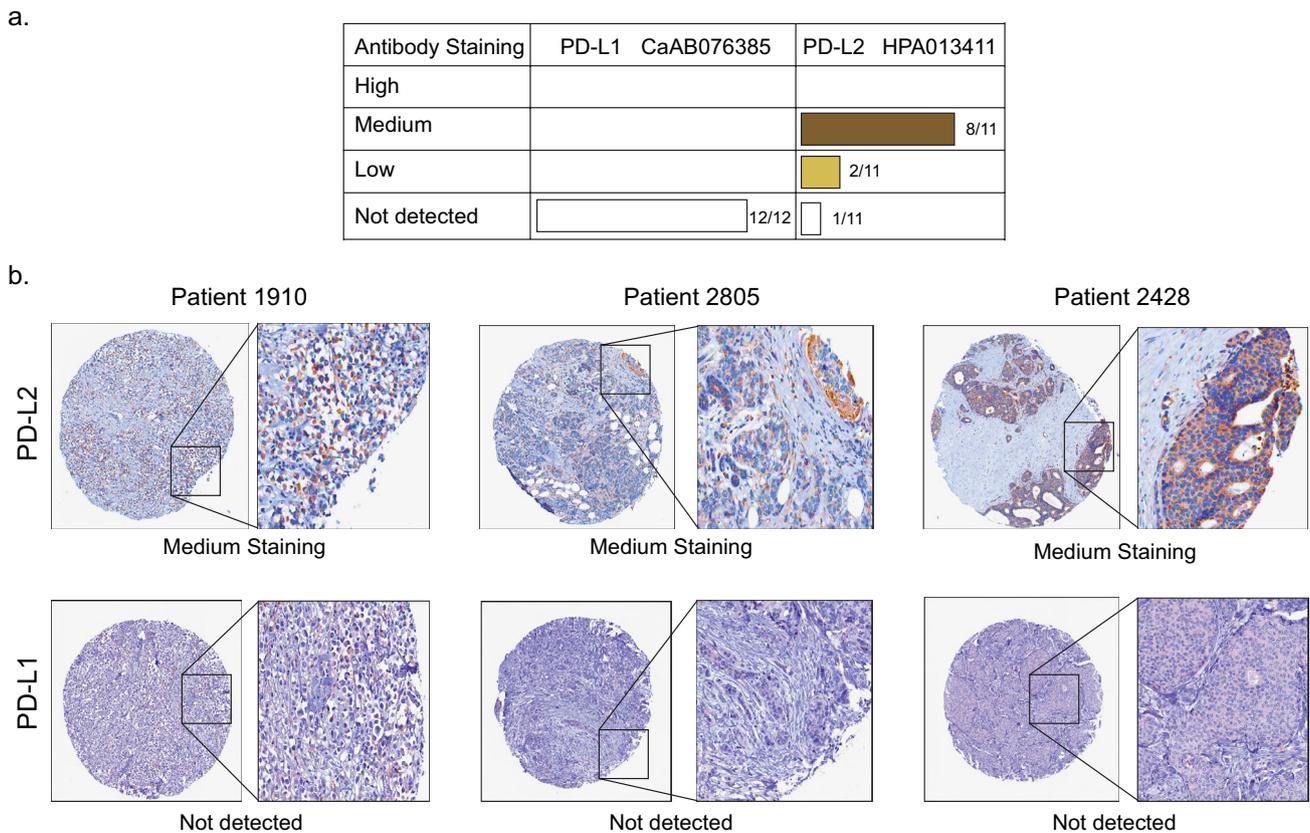


Fig. 5 The PD-L1 and PD-L2 protein staining data for breast tumors with representative regions amplified for details. **a** According to the Protein Atlas annotation, PD-L2 staining is medium, moderate intensity, the quantity is >75%, and the location is cytoplasmic/membra-

nous. In contrast, PD-L1 staining is not detected for the same patient samples. **b** The cell nuclei are labeled blue, and the proteins are shown with brown color. Image source credit: Human Protein Atlas

PD-L2

PD-L2 has a sequence identity of only 38% to PD-L1 [20]. Structural, functional, and evolutionary differences between the two ligands showed PD-L2 to have a stronger affinity to PD-1 than the PD-1/PD-L1 interaction [28]. Hence, we investigated high PD-L2 expression correlated genes and gene signatures to determine responders and potential biomarker candidates for immune checkpoint blockade.

We demonstrate the EMT gene signature enrichment in the high *PD-L2*-expressing tumors among all breast cancer subtypes. The transition from polarized epithelial cells to mobile mesenchymal cells, mediated by EMT, allows enhanced stem cell properties, therapy resistance, invasiveness, and metastasis. A correlation between PD-L1 and EMT-related gene signatures in breast and other cancers has been suggested [reviewed in 45]. In addition to tumor cells, tumor-infiltrating immune cells can also promote EMT through the secretion of soluble factors (e.g., TGF β , IL-6, TNF- α , IFN- γ , VEGF, etc.). Interestingly, pathways

implicated in EMT have been linked to PD-L1 upregulation. It remains to be tested whether EMT also modulates *PD-L2* transcription.

Of note, the high PD-L2/PD-L1 ratio group in ER+ breast tumors had a downregulated estrogen response-related gene signature, suggesting the presence of a more aggressive subgroup. These results support a recent study by Chervoneva et al. Approximately one-third of treatment-naive ER+ breast tumors (n = 684, and a validation cohort of 273 patients) were reported to have high PD-L2 IHC staining, independently predicting poor clinical outcomes and elevated progression risk in patients despite receiving adjuvant chemotherapy [46]. So far, PD-1 inhibitors have been approved only for TNBCs, and ER+ subtype patients are considered to have lower response rates to immunotherapy [47]. Hence, finding markers to identify likely responders expressing PD-L2 could significantly impact these patients.

On the other hand, PD-L1 signaling in some tumor cells has been linked to cancer initiation, EMT, invasion, metastasis, glucose metabolism, and drug resistance [48]. However,

PD-L2 signaling in cancer cells is less understood and warrants more interest in light of its expression in cancers. High PD-L2 in some patients may be competitive over antibody drugs targeting PD-L1 or the PD-1 receptor. Hence, tumors exclusively expressing PD-L1 or with PD-L2 may have different responses to immunotherapy. Of note, the expression of PD-L2 may explain why some PD-L1-negative patients still benefit from anti-PD-1 therapy. Hence, high PD-L2 expression in breast and other tumor types may provide a rationale for immune checkpoint blockade with anti-PD-1 and yet-to-be-developed PD-L2 inhibitors. Of note, no alternatively spliced PD-L2 isoform is reported in normal tissues tested in the GTEx database, but in an early study, PD-L2 isoforms were reported in activated leukocytes [49]. Moreover, pathogenic alterations may also affect the function of PD-L2 or cause deregulated alternative splicing. Indeed, a germline mutation in the PD-L2 gene causes a novel transcript variant that likely underlies the genetic etiology of the lymphomas in a specific family [50]. Hence identification of PD-L2 mutations and/or isoforms may also be critical during personalized immunotherapy decisions.

In closing, we provide transcriptome-level evidence supported by structural modeling, bioinformatic analyses, and IHC data on the significance of PD-L1 isoforms and PD-L2 in breast cancer subtypes. Because effective predictive biomarkers for PD-1 blockade are needed to improve immunotherapy response, the expression of PD-L1 isoforms and PD-L2 in tumors warrants further research. Based on accumulating evidence and our findings, we propose the urgent need to develop and utilize isoform-specific antibodies for PD-L1 and PD-L2 for staining patient tissue sections to predict immunotherapy outcomes better. Immune-dependent overlapping and independent, unique molecular functions of PD-L1 and PD-L2 also remain to be fully understood.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00262-023-03543-y>.

Acknowledgements EK and IY are supported by the EMBO Installation Grant (No: 4421). We thank METABRIC Consortium for sharing the breast cancer data.

Author contributions All authors contributed to the study's conception and design. Data collection and analysis were performed by DND, IO, and IY. MCY and EK contributed to the data analysis. The manuscript was written by AEED, and all authors contributed, read, and approved the manuscript. The results of the dataset analysis of the current study are given in the supplementary materials.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

- Freeman GJ et al (2000) Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation. *J Exp Med* 192(7):1027–1034. <https://doi.org/10.1084/jem.192.7.1027>
- Matsuzaki J et al (2010) Tumor-infiltrating NY-ESO-1-specific CD8+ T cells are negatively regulated by LAG-3 and PD-1 in human ovarian cancer. *Proc Natl Acad Sci USA* 107(17):7875–7880. <https://doi.org/10.1073/pnas.1003345107>
- Muenst S, Soysal SD, Gao F, Obermann EC, Oertli D, Gillanders WE (2013) The presence of programmed death 1 (PD-1)-positive tumor-infiltrating lymphocytes is associated with poor prognosis in human breast cancer. *Breast Cancer Res Treat* 139(3):667–676. <https://doi.org/10.1007/s10549-013-2581-3>
- Sun S et al (2014) PD-1(+) immune cell infiltration inversely correlates with survival of operable breast cancer patients. *Cancer Immunol Immunother* CII 63(4):395–406. <https://doi.org/10.1007/s00262-014-1519-x>
- Pardoll DM (2012) The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer* 12(4):252–264. <https://doi.org/10.1038/nrc3239>
- Okazaki T, Honjo T (2007) PD-1 and PD-1 ligands: from discovery to clinical application. *Int Immunol* 19(7):813–824. <https://doi.org/10.1093/intimm/dxm057>
- Herbst RS et al (2016) Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. *Lancet Lond Engl* 387(10027):1540–1550. [https://doi.org/10.1016/S0140-6736\(15\)01281-7](https://doi.org/10.1016/S0140-6736(15)01281-7)
- Larkin J et al (2015) Combined nivolumab and ipilimumab or monotherapy in untreated melanoma. *N Engl J Med* 373(1):23–34. <https://doi.org/10.1056/NEJMoa1504030>
- Schachter J et al (2017) Pembrolizumab versus ipilimumab for advanced melanoma: final overall survival results of a multicentre, randomised, open-label phase 3 study (KEYNOTE-006). *Lancet Lond Engl* 390(10105):1853–1862. [https://doi.org/10.1016/S0140-6736\(17\)31601-X](https://doi.org/10.1016/S0140-6736(17)31601-X)
- Latchman Y et al (2001) PD-L2 is a second ligand for PD-1 and inhibits T cell activation. *Nat Immunol* 2(3):261–268. <https://doi.org/10.1038/85330>
- Yearley JH et al (2017) PD-L2 expression in human tumors: relevance to anti-PD-1 therapy in cancer. *Clin Cancer Res Off J Am Assoc Cancer Res* 23(12):3158–3167. <https://doi.org/10.1158/1078-0432.CCR-16-1761>
- Matsubara T et al (2019) A clinicopathological and prognostic analysis of PD-L2 expression in surgically resected primary lung squamous cell carcinoma. *Ann Surg Oncol* 26(6):1925–1933. <https://doi.org/10.1245/s10434-019-07257-3>
- Shin S-J et al (2016) Clinicopathologic analysis of PD-L1 and PD-L2 expression in renal cell carcinoma: association with oncogenic proteins status. *Ann Surg Oncol* 23(2):694–702. <https://doi.org/10.1245/s10434-015-4903-7>
- Zhang Y et al (2019) A PD-L2-based immune marker signature helps to predict survival in resected pancreatic ductal adenocarcinoma. *J Immunother Cancer* 7(1):233. <https://doi.org/10.1186/s40425-019-0703-0>
- Colaprico A et al (2016) TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res* 44(8):e71. <https://doi.org/10.1093/nar/gkv1507>
- Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P (2015) The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst* 1(6):417–425. <https://doi.org/10.1016/j.cels.2015.12.004>

17. Curtis C et al (2012) The genomic and transcriptomic architecture of 2000 breast tumours reveals novel subgroups. *Nature* 486(7403):346–352. <https://doi.org/10.1038/nature10983>
18. UniProt Consortium (2023) UniProt: the universal protein knowledgebase in 2023. *Nucleic Acids Res* 51(D1):D523–D531. <https://doi.org/10.1093/nar/gkac1052>
19. Berman HM et al (2000) The protein data bank. *Nucleic Acids Res* 28(1):235–242. <https://doi.org/10.1093/nar/28.1.235>
20. Lin DY-W et al (2008) The PD-1/PD-L1 complex resembles the antigen-binding Fv domains of antibodies and T cell receptors. *Proc Natl Acad Sci USA* 105(8):3011–3016. <https://doi.org/10.1073/pnas.0712278105>
21. Jumper J et al (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596(7873):583–589. <https://doi.org/10.1038/s41586-021-03819-2>
22. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M (2022) ColabFold: making protein folding accessible to all. *Nat Methods* 19(6):679–682. <https://doi.org/10.1038/s41592-022-01488-1>
23. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372(3):774–797. <https://doi.org/10.1016/j.jmb.2007.05.022>
24. Xie Y et al (2022) IBS 2.0: an upgraded illustrator for the visualization of biological sequences. *Nucleic Acids Res* 50(W1):W420–W426. <https://doi.org/10.1093/nar/gkac373>
25. Uhlen M et al (2017) A pathology atlas of the human cancer transcriptome. *Science* 357(6352):eaan2507. <https://doi.org/10.1126/science.aan2507>
26. Freeman GJ (2008) Structures of PD-1 with its ligands: sideways and dancing cheek to cheek. *Proc Natl Acad Sci* 105(30):10275–10276. <https://doi.org/10.1073/pnas.0805459105>
27. Garcia-Diaz A et al (2017) Interferon receptor signaling pathways regulating PD-L1 and PD-L2 expression. *Cell Rep* 19(6):1189–1201. <https://doi.org/10.1016/j.celrep.2017.04.031>
28. Phillips EA et al (2020) The structural features that distinguish PD-L2 from PD-L1 emerged in placental mammals. *J Biol Chem* 295(14):4372–4380. <https://doi.org/10.1074/jbc.AC119.011747>
29. Lim H, Chun J, Jin X, Kim J, Yoon J, No KT (2019) Investigation of protein-protein interactions and hot spot region between PD-1 and PD-L1 by fragment molecular orbital method. *Sci Rep* 9(1):16727. <https://doi.org/10.1038/s41598-019-53216-z>
30. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N (2015) Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* 160(1–2):48–61. <https://doi.org/10.1016/j.cell.2014.12.033>
31. Jiang P et al (2018) Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med* 24(10):1550–1558. <https://doi.org/10.1038/s41591-018-0136-1>
32. Redfern AD, Spalding LJ, Thompson EW (2018) The Kraken Wakes: induced EMT as a driver of tumour aggression and poor outcome. *Clin Exp Metastasis* 35(4):285–308. <https://doi.org/10.1007/s10585-018-9906-x>
33. Wang G et al (2021) The pan-cancer landscape of crosstalk between epithelial-mesenchymal transition and immune evasion relevant to prognosis and immunotherapy response. *NPJ Precis Oncol* 5(1):56. <https://doi.org/10.1038/s41698-021-00200-4>
34. Tokumaru Y et al (2020) KRAS signaling enriched triple negative breast cancer is associated with favorable tumor immune microenvironment and better survival. *Am J Cancer Res* 10(3):897–907
35. Lawson NL, Dix CI, Scorer PW, Stubbs CJ, Wong E, Hutchinson L, Barker C (2020) Mapping the binding sites of antibodies utilized in programmed cell death ligand-1 predictive immunohistochemical assays for use with immuno-oncology therapies. *Modern Pathol Off J US Can Acad Pathol Inc* 33(4):518–530. <https://doi.org/10.1038/s41379-019-0372-z>
36. Pang C et al (2018) Assessment of programmed cell death ligand-1 expression with multiple immunohistochemistry antibody clones in non-small cell lung cancer. *J Thorac Dis* 10(2):816–824. <https://doi.org/10.21037/jtd.2018.01.124>
37. He X, Xu L, Liu Y (2005) Identification of a novel splice variant of human PD-L1 mRNA encoding an isoform-lacking IgV-like domain. *Acta Pharmacol Sin* 26(4):462–468. <https://doi.org/10.1111/j.1745-7254.2005.00086.x>
38. Yan L, Sun Y, Guo J, Jia R (2023) PD-L1 exon 3 is a hidden switch of its expression and function in oral cancer cells. *Int J Mol Sci* 24(9):8193. <https://doi.org/10.3390/ijms24098193>
39. Sagawa R et al (2022) Soluble PD-L1 works as a decoy in lung cancer immunotherapy via alternative polyadenylation. *JCI Insight* 7(1):e153323. <https://doi.org/10.1172/jci.insight.153323>
40. Chen J et al (2011) Intrahepatic levels of PD-1/PD-L correlate with liver inflammation in chronic hepatitis B. *Inflamm Res Off Eur Histamine Res Soc* 60(1):47–53. <https://doi.org/10.1007/s00011-010-0233-1>
41. Koukourakis MI et al (2018) Increased soluble PD-L1 levels in the plasma of patients with epithelial ovarian cancer correlate with plasma levels of miR34a and miR200. *Anticancer Res* 38(10):5739–5745. <https://doi.org/10.21873/anticancer.12912>
42. Okuma Y, Hosomi Y, Nakahara Y, Watanabe K, Sagawa Y, Homma S (2017) High plasma levels of soluble programmed cell death ligand 1 are prognostic for reduced survival in advanced lung cancer. *Lung Cancer Amst Neth* 104:1–6. <https://doi.org/10.1016/j.lungcan.2016.11.023>
43. Wang X et al (2015) Tumor suppressor miR-34a targets PD-L1 and functions as a potential immunotherapeutic target in acute myeloid leukemia. *Cell Signal* 27(3):443–452. <https://doi.org/10.1016/j.cellsig.2014.12.003>
44. Zhu X, Lang J (2017) Soluble PD-1 and PD-L1: predictive and prognostic significance in cancer. *Oncotarget* 8(57):97671–97682. <https://doi.org/10.18632/oncotarget.18311>
45. Segovia-Mendoza M, Romero-Garcia S, Lemini C, Prado-Garcia H (2021) Determining factors in the therapeutic success of checkpoint immunotherapies against PD-L1 in breast cancer: a focus on epithelial-mesenchymal transition activation. *J Immunol Res* 2021:6668573. <https://doi.org/10.1155/2021/6668573>
46. Chervoneva I et al (2023) High PD-L2 predicts early recurrence of ER-positive breast cancer. *JCO Precis Oncol* 7:e2100498. <https://doi.org/10.1200/PO.21.00498>
47. Yang T, Li W, Huang T, Zhou J (2023) Immunotherapy targeting PD-1/PD-L1 in early-stage triple-negative breast cancer. *J Pers Med* 13(3):526. <https://doi.org/10.3390/jpm13030526>
48. Hudson K, Cross N, Jordan-Mahy N, Leyland R (2020) The extrinsic and intrinsic roles of PD-L1 and its receptor PD-1: implications for immunotherapy treatment. *Front Immunol* 11:568931. <https://doi.org/10.3389/fimmu.2020.568931>
49. He X-H, Liu Y, Xu L-H, Zeng Y-Y (2004) Cloning and identification of two novel splice variants of human PD-L2. *Acta Biochim Biophys Sin* 36(4):284–289. <https://doi.org/10.1093/abbs/36.4.284>
50. Shao J et al (2022) Exome sequencing identifies PD-L2 as a potential predisposition gene for lymphoma. *Hematol Oncol* 40(3):475–478. <https://doi.org/10.1002/hon.3033>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.