PD-L1 immunohistochemistry assay concordance in urothelial carcinoma of the bladder and hypopharyngeal squamous cell carcinoma

Anjelica Hodgson, MD[1,2], **Elzbieta Slodkowska, MD**[1,2], **Achim Jungbluth, MD PhD**[3], **Stanley K. Liu, MD PhD**[4], **Danny Vesprini, MD**[4], **Danny Enepekides, MD**[5], **Kevin Higgins, MD**[5], **Nora Katabi, MD**[3], **Bin Xu, MD PhD**[1,2,*], **Michelle R. Downes, MD**[1,2,*]

[1]Department of Anatomic Pathology, Sunnybrook Health Sciences Centre, Toronto, ON, Canada

[2]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada

[3]Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[4]Department of Radiation Oncology, University of Toronto, Toronto, ON, Canada

[5]Department of Head and Neck Surgery, Sunnybrook Health Sciences Centre, Toronto, ON, Canada

Abstract

PD-L1 immunohistochemistry is used to guide treatment decisions regarding the use of checkpoint immunotherapy in the management of urothelial carcinoma of the bladder and hypopharyngeal squamous cell carcinoma. With increasing PD-L1 testing options, a need has arisen to assess the analytical comparability of diagnostic assays in order to develop a more sustainable testing strategy. Using tissue microarrays, PD-L1 expression in tumor cells (TC) and immune cells (IC) was manually scored in 197 cases and 27 cases of bladder and hypopharyngeal cancer, respectively. Three commercial kits (Ventana SP263, Ventana SP142, Dako 22C3) and one platform-independent test (Cell Signalling Technologies E1L3N) were utilized. Across the three commercially available clones, 14% and 74% of urothelial carcinomas were positive and negative respectively, while 7% and 78% of hypopharyngeal carcinomas were positive and negative respectively. 12% of bladder and 15% hypopharyngeal cases showed discrepant PD-L1 classification results. Regardless of the scoring algorithm used, E1L3N provided comparable PD-L1 staining results. Fleiss' kappa and intraclass correlation coefficient (ICC) analyses demonstrated substantial agreement among all antibody clones (k=0.639–0.791) and excellent reliability among SP263, 22C3, and E1L3N antibodies (ICC 0.929–0.949) in TC staining. Compared to the other three clones, SP142 TC staining was lower with only moderate correlation

*Joint corresponding authors: Michelle R. Downes, Room E417, Anatomic Pathology, Sunnybrook Health Sciences Centre, 2075 Bayview Avenue, Toronto, ON, Canada, M4N 3M5, michelle.downes@sunnybrook.ca, Phone: 416-480-6100 ext. 87597, Fax: 416-480-2179, Bin Xu, Room E428, Anatomic Pathology, Sunnybrook Health Sciences Centre, 2075 Bayview Avenue, Toronto, ON, Canada, M4N 3M5, bin1.xu@sunnybrook.ca, Phone: 416-480-4011, Fax: 416-480-2179.

(ICC 0.500–0.619). Generally, the reliability of IC staining was lower compared to TC staining (ICC 0.519–0.866). Our results demonstrate good analytic comparability of all four antibodies. The results are encouraging and support growing optimism in the pathology and oncology communities concerning strategies in PD-L1 assay use.

Keywords

INTRODUCTION

The use of checkpoint inhibitors in the management of various malignancies has become increasingly prevalent following successful early clinical trials[1]. The programmed death-1 (PD-1) and programmed death ligand-1 (PD-L1) checkpoint has emerged as a key checkpoint that can be manipulated using inhibitory monoclonal antibodies. In this context, the expression profile of PD-L1, is typically evaluated to identify patients most likely to derive clinically meaningful benefit from such inhibitors[2,3].

PD-1 is usually expressed on adaptive immune cells as well as natural killer cells while PD-L1, in addition to being expressed by lymphocytes, is expressed by different antigen-presenting cells and endothelial cells[1]. Normal physiologic interactions between this receptor-ligand pair inhibit the immune response, and thus these molecules play an important role in the prevention of autoimmunity through the promotion of self-tolerance[4,5]. Remarkably, it has been shown that the PD-1/PD-L1 pathway is used by tumor cells to evade recognition by the host immune system, as PD-L1 can also be expressed by a multitude of tumors[6]. Because of this, the PD-1/PD-L1 pathway has become the target of intense oncologic and pathologic research, as it has been shown that blockade of PD-1 and/or PD-L1 has positive therapeutic benefit[7,8]. Certain checkpoint inhibitors are already available for use in the treatment of both urothelial carcinoma (UC)[9–14] and squamous cell carcinoma (SCC) of the head and neck[15–18] and others are being evaluated in clinical trials. In order to determine patient candidacy for treatment with these drugs, immunohistochemical testing is used to evaluate expression of PD-L1 in tumor samples. Multiple assays are available for diagnostic use, each with its own primary antibody, specific assay performance conditions and scoring methods. Within each assay, the cut-off levels required for positivity will vary depending on tumor type.

Evaluation of PD-L1 expression in tumor tissue is complicated by many factors[6]. With multiple assays available, each linked with a specific therapeutic agent, identification of patients who may benefit from immune checkpoint inhibition can be challenging. Given limited tissue and resource availability coupled with platform specific assay requirements, performing multiple assays for each patient is ultimately impractical. As such, it is imperative to understand the potential interchangeability of various assays, given the critical clinical implications.

This issue of assay comparison and interchangeability has been investigated in the context of several organs, mainly non-small cell carcinoma of the lung[19–21] However data for both UC and hypopharyngeal SCC (HP-SCC) are lacking.

With this in mind, we sought to assess and compare the analytic performance of three commercially available *in vitro* diagnostic (IVD) PD-L1 expression assays (Ventana SP263, Ventana SP142, and Dako 22C3) in UC of the bladder and HP-SCC. In addition, we assessed the performance of a platform independent laboratory developed test (LDT), Cell Signalling Technologies E1L3N antibody.

MATERIALS AND METHODS

Case selection

This work was approved by the research ethics board of Sunnybrook Health Sciences Centre (SHSC, 187–2016 and 359–2017). A retrospective search of the laboratory information system (Sunquest CoPath) was performed to identify UC treated by cystectomy between 1999 and 2015, as well as HP-SCC resected between 2000 and 2017. For UC cases, the inclusion criteria were: high grade disease, stage of pT1 or higher, and urothelial histology with less than 50% divergent differentiation. For patients with HP-SCC, only patients with tumor epicenter located within the hypopharynx who underwent surgical resection at our institution were included. No patient had received prior immune checkpoint therapy. Clinicopathological features including age, sex, histotype of UC, American Joint Committee on Cancer (AJCC) pT stage and nodal metastasis (if present), were recorded.

Tissue microarray (TMA) construction

Triplicate 1-mm core TMAs were constructed using a TMA instrument (Beecher instruments, Silver Springs, MD, USA) as previously described[22]. The cores were taken as follows: one punch from tumor front, one from superficial tumor and one from mid portion of tumor to control for temporal heterogeneity. Sections from the TMA blocks were stained with haematoxylin and eosin to confirm the adequacy of tumor sampling and to facilitate orientation during the interpretation of the anti PD-L1 staining.

Immunohistochemistry

Serial unstained slides (4 microns) were prepared from each block for subsequent immunohistochemistry (IHC) with the following PD-L1 antibody clones: SP142 and SP263 (Ventana Medical Systems, AZ, USA), 22C3 (Agilent Technologies, CA, USA) and E1L3N (Cell Signaling Technologies, MA, USA). Staining for anti PD-L1 rabbit mAb clones SP142, SP263 and E1L3N was performed on the Ventana Benchmark Ultra automated staining platform according to the manufacturer's protocol (with the Optiview DAB IHC detection kit for all the clones, followed by the Optiview amplification kit for SP142 only; for mouse mAb clone 22C3, the DAKO EnVision FLEX system on a DAKO Autostainer Link 48 system was used according to the manufacturer's instructions (Figure 1).

Scoring tumor cell and immune cell staining

The UC (n=197) and HP-SCC cases (n=27) were read and interpreted by fellowship-trained Genitourinary (MD) and Head and Neck (BX) pathologists, respectively. Tumor cell (TC) staining was defined as either partial or complete membranous staining of tumor cells of any intensity. Immune cell (lymphocytes and macrophages, IC) staining was either cytoplasmic or membranous of any intensity. The identification of TC and IC was based on morphologic features alone with assistance of H&E slides. Both TC and IC staining was assessed as either presence or absence with the estimated percentage of positive TC and positive IC recorded for each TMA core. For both the UC and HP-SCC cases, there was variability in PD-L1 expression from core to core. To counter this, the results were averaged across the triplicate cores to give one estimate of percentage TC and percentage IC for each case. Each of the four assays was assessed with a wash out period of several days between each round of scoring.

Application of clinical algorithms

For each of the three commercially available in vitro diagnostic assays (SP142, SP263, and 22C3), the site-specific manufacturer's algorithm was applied to each case (Supplementary Table 1). This was used with the recommended cut-off to designate cases as "positive" or "negative" as would be required for clinical use of the associated immune checkpoint inhibitor. The scoring was performed by the same pathologists as the TC/IC scoring.

For UC of the bladder, the SP263 staining was considered a positive result if 25% or more tumor or immune cells were staining. SP142 stained sections were deemed positive when the positive immune cells represented 5% or more of the tumor cell area. The 22C3 staining was recorded as positive when the combined proportion score (CPS) was 10% or greater. The CPS was calculated by combining the positive TC and IC staining and expressing it relative to the total tumor cells.

For HP-SCC, the SP263 staining was considered a positive result if 25% or more TCs were stained. SP142 stained sections were deemed positive when 50% or more TCs or 10% or more ICs showed immunopositivity. The 22C3 staining was recorded as positive when 50% or more TC showed partial or complete membranous staining.

The average TC and IC percentage scores of E1L3N anti PD-L1 stained sections were then assessed applying one of the algorithms for SP142, SP263 and 22C3 to allow comparison of E1L3N antibody clone with the other three clones.

Statistics

All statistical analyses were performed using the SPSS software 24.0 (IBM Corporation, New York, NY, USA). The agreement of the overall positivity and negativity determined using individual algorithms for PD-L1 clones 22C3, SP142, and SP263 was assessed using Fleiss' kappa statistics. Kappa analysis was also applied to compare the performance of E1L3N clone with the other three clones by means of applying appropriate algorithms as described in Supplementary Table 1. To compare the actual TC or IC percentage score, pairwise comparison of TC or IC between the antibody clones were conducted by intraclass

correlation coefficient (ICC) analysis. P values less than 0.05 were considered to be statistically significant.

RESULTS

Characteristics of the study cohort

The clinical and pathologic features of the study cohort are presented in Table 1. In brief, there were 197 UC patients including 143 males and 54 females. The median age at presentation was 71 (range 33–93). Seventy-four percent of UCs were the usual type, while the remaining showed divergent differentiation. The majority of the patients had advanced primary tumors at the time of resection, including 55% with pT3 and 28% with pT4 disease. Lymph node metastases were found in 79 (40%) of cases.

Among the 27 patients with HP-SCC, the median age as 64 (range 49–79). There was a male predominance with a male: female ratio of 3:1. Twenty (74%) primary tumors were moderately differentiated, while the rest was poorly differentiated. Nineteen (70%) patients had nodal metastases at the time of primary resection.

PD-L1 immunopositivity in UC and HP-SCC: a comparison of performance among antibody clones with established clinical algorithms

The performance of 22C3, SP263, and SP142 clones is illustrated in Figure 2. Among the 197 UCs studied, 27 (14%) tumors were considered as positive while 137 (74%) were negative across all three antibody clones. The remaining 23 tumors (12%) showed inconsistent results. In the 27 HP-SCCs tested, two (7%) were universally positive; 21 (78%) were consistently negative, while the remaining four (15%) exhibited discrepant results. The number and percentage of UC deemed PD-L1 positive were 41 (21%) using SP263 clone, 36 (18%) using SP142 clone, and 40 (20%) using 22C3 clone. The number and percentage of HP-SCC that were positive for PD-L1 were 5 (19%) using SP263 clone, 3 (11%) using SP142 clone, and 3 (11%) using 22C3 clone.

Figure 3 provides comparison between the E1L3N clone and each of the other clones using the same algorithm to determine PD-L1 positivity and negativity. Regardless of the algorithm used, E1L3N provided comparable results of PD-L1 immunopositivity: in UC, 20%, 17%, and 28% of cases were positive using SP263, SP142, and 22C3 algorithms, respectively. For HP-SCC, 15%, 15%, and 7% of cases were positive using SP263, SP142, 22C3 algorithms, respectively. Only a minority of cases showed discrepant results (positive with one clone but negative with the other) among E1L3N and one of the other antibody clones: in UC, this was 8% (SP263), 11% (SP142), and 10% (22C3); in HP-SCC, 4% of cases were discrepant when any of the three clones were compared.

Pairwise comparisons of PD-L1 antibody clones

The results of Fleiss' kappa analyses are shown in Table 2. There was substantial agreement between any of the two antibody clones compared (p < 0.001). The kappa value ranged from 0.639 to 0.791.

The actual percentage of TCs and ICs were compared using intraclass correlation coefficient analysis, and the results are illustrated in Table 2 and Figure 4. An excellent reliability was achieved for TC between 1) SP263 and E1L3N clones (ICC=0.949, 95% CI=0.933–0.961), 2) 22C3 and E1L3N (ICC=0.939, 95% CI=0.918–0.955), and 3) SP263 and 22C3 (ICC=0.929, 95% CI=0.905–0.947). In contrast, there was only moderate reliability when comparing the TC score obtained using SP142 clone to any of the other three clones (22C3: ICC=0.619, 95% CI=0.481–0.717; SP263: ICC=0.525, 95% CI=0.340–0.652; and E1L3N: ICC=0.500, 95% CI=0.320–0.629). The scatterplots of TC scores are shown in Figure 4 (left panel). Overall, a near-perfect correlation of the actual TC scores was achieved 1) between SP263 and 22C3; 2) between SP263 and E1L3N; and 3) between 22C3 and E1L3N. In contrast, the actual TC percentage rated using SP142 antibody appeared to be lower than the other 3 antibodies.

Generally, the reliability of IC scoring appeared to be lower compared with the reliability of TC score. There was good reliability of 22C3 and E1L3N IC scoring (ICC=0.866, 95% CI=0.826–0.897), and moderate reliability in any other pairwise comparisons including 1) SP263 vs. SP142: ICC=0.722, 95% CI=0.638–0.786; 2) SP263 vs. E1L3N: ICC=0.695, 95% CI=0.595–0.769; 3) SP142 vs. E1L3N: ICC=0.667, 95% CI=0.567–0.744; 4) SP263 vs. 22C3: ICC=0.567, 95% CI=0.435–0.668; and 5) SP142 vs. 22C3: ICC=0.519, 95% CI=0.376–0.630. The scatter plots with regression lines are shown in Figure 4 (right panel). The IC score appeared to be highest in SP142, followed by E1L3N and 22C3, and lowest in SP263.

DISCUSSION

PD-L1 inhibitors are being used to treat many different malignancies including SCC of the head and neck and UC of the bladder. As these therapeutic options become increasingly used, the demands for pathologic testing will also increase. Currently, testing for PD-L1 immunohistochemical expression is governed by little standardization, leading to multiple assays which utilize different clones, staining platforms, and scoring algorithms[23]. From a practical point of view, a need has arisen to evaluate the concordance between assays, in order to develop a more sustainable and logistically sound testing strategy. Because of this, we endeavored to compare the diagnostic performance of three commonly used commercially available in-vitro assays (Ventana SP 263 and SP142, and Dako 22C3). In addition, we sought to assess how a lab-developed test (Cell Signalling Technologies E1L3N) would compare when the same parameters for each *in vitro* assay were applied.

Our results have demonstrated several important points including (1) a high level of concordant PD-L1 scoring results among the three commercially available assays, (2) comparable PD-L1 positivity regardless of algorithm when E1L3N antibody was used, (3) substantial agreement of PD-L1 positivity interpretation between any of the two utilized antibody clones by Fleiss' kappa analysis, (4) excellent reliability as determined by intraclass correlation efficient analysis for TC staining for three clones, and (5) an overall lower reliability for immune cell staining. It is worth noting that TC staining is not a component of the UC SP142 algorithm even though TC staining is seen in some cases.

Therefore, it is not surprising that TC staining in UC is less than that with the other three antibodies.

We have shown here that there are diagnostic similarities between the four tested PD-L1 systems. The low level of discrepant results between SP263, SP142, and 22C3 antibodies that was observed in our study is reassuring, and is not dissimilar to what has been recently reported in studies of bladder cancer[24] and lung cancer[25,26] (different combinations of antibodies used). Interestingly, the number of discrepant results for both bladder and hypopharyngeal cases in our study is significantly lower than what was reported in the Blueprint study on non-small cell carcinoma of the lung[19]: 37% vs. 12% and 15% for bladder and hypopharyngeal cases, respectively. Multiple factors, including specimen type and antibody clones compared may attribute to the different discrepancy rate between the two studies. While our study utilized TMAs, the Blueprint study included whole sections from surgical resections and biopsies which may contribute to greater intratumor heterogeneity and additional interpretation subjectivity. The present study included three commercially available antibodies while the Blueprint study also involved PD-L1 28–8 clone which might add additional interpretation difference between antibody clones.

The E1L3N antibody clone which is staining platform independent, showed remarkably concordant performance compared to the IVD clones. These results are in keeping with another study in UC[24]. Such comparative studies in hypopharyngeal carcinoma have not been published to date. Not unexpectedly, the SP142 antibody was found to stain less tumor cells compared to the other antibodies in both UC and SCC groups, a finding recently reported in other organ systems[19,21,27]. The scoring algorithm for SP142 in the evaluation of both UC of the bladder and HP-SCC incorporates IC staining. While SP142 showed a similar overall diagnostic performance, pairwise comparisons only demonstrated moderate reliability in TC and IC staining when compared with the other 3 clones. This is explained by the fact that SP142 tends to stain more IC, and less TC. It could be argued that the same challenges that are experienced with immune cell staining should apply to SP263, as IC staining is incorporated into the scoring algorithm of SP263 specifically in UC of the bladder. This is not the case for two reasons: firstly, SP263 scoring in UC of the bladder also incorporates tumor cell staining (and tumor cell staining for this clone showed excellent reliability compared to 22C3 and E1L3N), and secondly, it is has been noted that SP142 immune cell staining may exhibit unusual staining patterns or may be low level/frequency, again making interpretation more difficult[28].

It has been noted that PD-L1 staining suffers from intratumoral heterogeneity in different tumor types[29–31]. This finding suggests that that tumor sampling based only on small biopsy material may wrongly classify patients based on PD-L1 status. A study looking at non-small cell lung carcinoma biopsy and subsequent surgical resection specimens demonstrated an eventual difference in PD-L1 expression[32]. Additional studies are needed to evaluate whether such difference may be also applied to bladder and head and neck cancers. Prospective studies on biopsy and subsequent surgical materials from bladder and head and neck cancers are needed. However, in a study evaluating PD-L1 staining heterogeneity, no significant intra-block or intra-case heterogeneity was noted in cases of UC and SCC of the head and neck[31].

Antibody evaluation and comparison by TMA has been previously done in the evaluation of primary and metastatic bladder cancer samples with different PD-L1 antibody clones[24]. In this study, we have minimized the impact of intratumoral heterogeneity by assessing tumors using triplicate cores from different tumor areas, and the staining results have been averaged. In addition, heterogeneity was also mitigated by evaluating staining by the different clones in consecutive serial sections.

A limitation of our study was the relatively small number of hypopharyngeal SCC cases which was due to the tumor's relative rarity[33] and the fact that, like other malignancies of the aerodigestive tract, these lesions are typically treated with a multimodal approach[34] and may not be surgically resected depending on the extent of disease. Another potential limitation of the present study is that the evaluation of PD-L1 was performed on TMAs, as it is known that PD-L1 immunoexpression may be heterogeneous[35,36]. Hence, the PD-L1 score obtained using TMAs and small biopsies may not always be consistent with the PD-L1 immunostain based on whole slides from surgical resection. Indeed, variability of PD-L1 scores was observed among cores from the same tumor across different antibody clones in our study (data not shown), which reflected the intrinsic staining heterogenicity of PD-L1 within bladder and head and neck carcinomas. In an attempt to address this issue, multiple cores were sampled for each tumor and the PD-L1 score was averaged for each case.

Similar to Phase 1 of the Blueprint comparison project[19], our study utilized a small number of expert pathologists to evaluate the cases. As was acknowledged in the similar non-small cell lung cancer study, the generalizability of the results to non-experts is unknown. Studies evaluating the generalizability as well as reproducibility of assay interchangeability are needed in the future. Importantly, our study does not address nor compare the outcome predictive value of this interchangeable assay strategy. This type of analysis will be critical in the future, in order to validate these promising alternative testing approaches.

In conclusion, the performance of the four compared PD-L1 antibody assays demonstrated in our study is encouraging. The results presented here support previously reported data from other malignancies, and complement ongoing comparative studies. From a practical point of view, several issues related to PD-L1 testing could be eliminated through the use of a single harmonized assay delivery and interpretation system. In addition, the potential for cost savings[37] and a more stream-lined and efficient testing process cannot be ignored in the era of personalized medicine.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Financial support statements:

REFERENCES

1. Tumeh PC, Harview CL, Yearley JH, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. Nature 2014;515:568–571. [PubMed: 25428505]

2. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer 2012;12:252–264. [PubMed: 22437870]

3. Postow MA, Callahan MK, Wolchok JD. Immune Checkpoint Blockade in Cancer Therapy. J Clin Oncol. 2015;33:1974–1982. [PubMed: 25605845]

4. Speiser DE, Utzschneider DT, Oberle SG, et al. T cell differentiation in chronic infection and cancer: functional adaptation or exhaustion? Nat Rev Immunol. 2014;14:768–774. [PubMed: 25257362]

5. Pauken KE, Wherry EJ. Overcoming T cell exhaustion in infection and cancer. Trends Immunol. 2015;36:265–276. [PubMed: 25797516]

6. Patel SP, Kurzrock R. PD-L1 Expression as a Predictive Biomarker in Cancer Immunotherapy. Mol Cancer Ther. 2015;14:847–856. [PubMed: 25695955]

7. Topalian SL, Hodi FS, Brahmer JR, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. N Engl J Med. 2012;366:2443–2454. [PubMed: 22658127]

8. Brahmer JR, Tykodi SS, Chow LQM, et al. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. N Engl J Med. 2012;366:2455–2465. [PubMed: 22658128]

9. Bellmunt J, Powles T, Vogelzang NJ. A review on the evolution of PD-1/PD-L1 immunotherapy for bladder cancer: The future is now. Cancer Treat Rev. 2017;54:58–67. [PubMed: 28214651]

10. Crist M, Balar A. Atezolizumab in invasive and metastatic urothelial carcinoma. Expert Rev Clin Pharmacol. 2017:1–7.

11. Balar AV, Galsky MD, Rosenberg JE, et al. Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: a single-arm, multicentre, phase 2 trial. Lancet 2017;389:67–76. [PubMed: 27939400]

12. Rosenberg JE, Hoffman-Censits J, Powles T, et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. Lancet 2016;387:1909–1920. [PubMed: 26952546]

13. Ning Y-M, Suzman D, Maher VE, et al. FDA Approval Summary: Atezolizumab for the Treatment of Patients with Progressive Advanced Urothelial Carcinoma after Platinum-Containing Chemotherapy. The Oncologist 2017;22:743–749. [PubMed: 28424325]

14. Powles T, Eder JP, Fine GD, et al. MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer. Nature 2014;515:558–562. [PubMed: 25428503]

15. Economopoulou P, Kotsantis I, Psyrri A. The promise of immunotherapy in head and neck squamous cell carcinoma: combinatorial immunotherapy approaches. ESMO Open 2016;1:e000122. [PubMed: 28848660]

16. Müller T, Braun M, Dietrich D, et al. PD-L1: a novel prognostic biomarker in head and neck squamous cell carcinoma. Oncotarget 2017;8:52889–52900. [PubMed: 28881780]

17. De Meulenaere A, Vermassen T, Aspeslagh S, et al. Turning the tide: Clinical utility of PD-L1 expression in squamous cell carcinoma of the head and neck. Oral Oncol. 2017;70:34–42. [PubMed: 28622889]

18. Ferris RL, Blumenschein G, Fayette J, et al. Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. N Engl J Med. 2016;375:1856–1867. [PubMed: 27718784]

19. Hirsch FR, McElhinny A, Stanforth D, et al. PD-L1 Immunohistochemistry Assays for Lung Cancer: Results from Phase 1 of the Blueprint PD-L1 IHC Assay Comparison Project. J Thorac Oncol. 2017;12:208–222. [PubMed: 27913228]

20. Ratcliffe MJ, Sharpe A, Midha A, et al. Agreement between Programmed Cell Death Ligand-1 Diagnostic Assays across Multiple Protein Expression Cutoffs in Non-Small Cell Lung Cancer. Clin Cancer Res. 2017;23:3585–3591. [PubMed: 28073845]
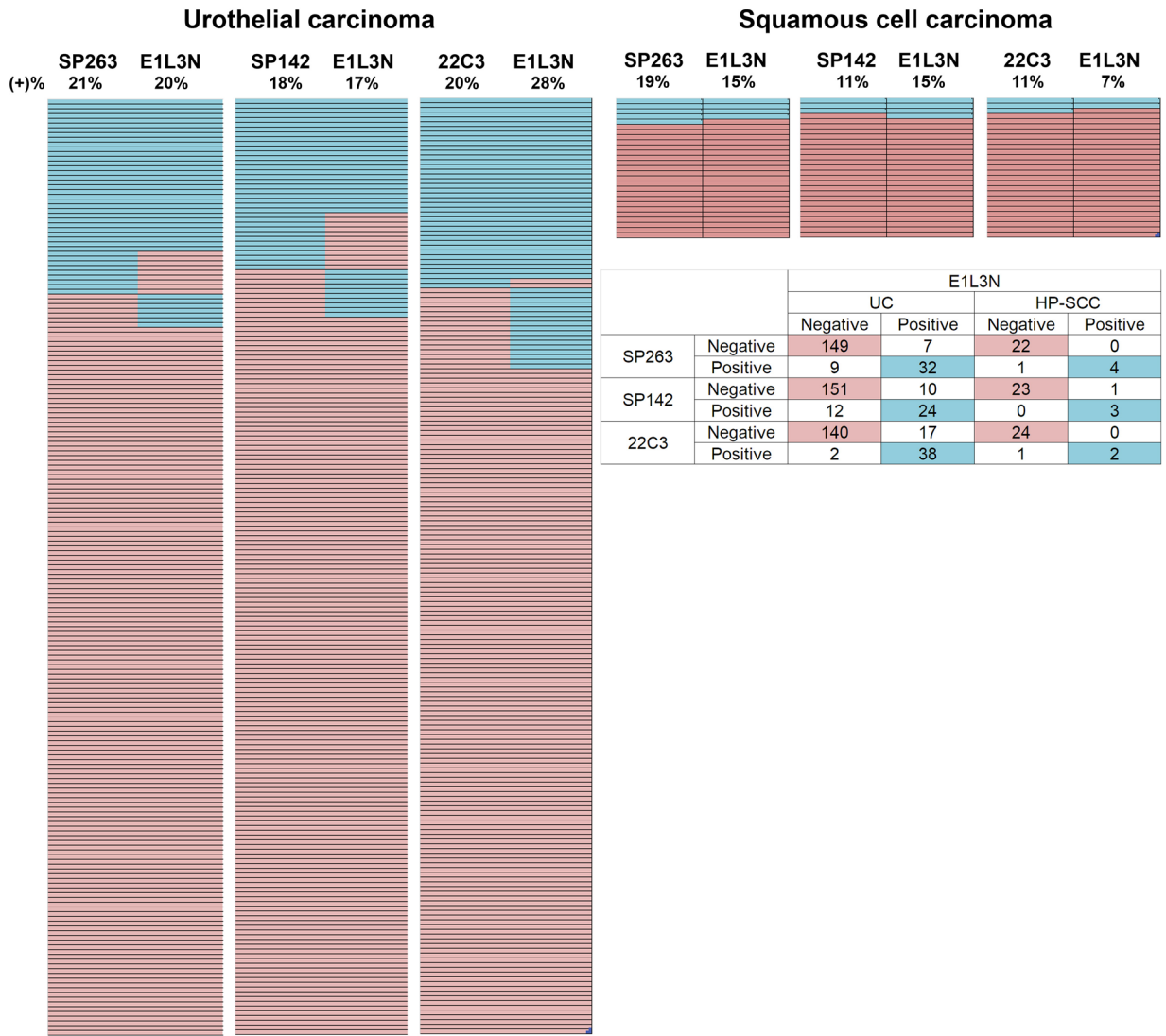
21. Scheel AH, Baenfer G, Baretton G, et al. Interlaboratory concordance of PD-L1 immunohistochemistry for non-small-cell lung cancer. Histopathology 2018;72:449–459. [PubMed: 28851100]

22. Hodgson A, Xu B, Downes MR. p53 immunohistochemistry in high-grade urothelial carcinoma of the bladder is prognostically significant. Histopathology 2017;71:296–304. [PubMed: 28342221]

23. Kerr KM, Hirsch FR. Programmed Death Ligand-1 Immunohistochemistry: Friend or Foe? Arch Pathol Lab Med. 2016;140:326–331. [PubMed: 26756647]

24. Tretiakova M, Fulton R, Kocherginsky M, et al. Concordance study of PD-L1 expression in primary and metastatic bladder carcinomas: comparison of four commonly used antibodies and RNA expression. Mod. Pathol. 2017 12 22. doi: 10.1038/modpathol.2017.188. Epub ahead of print.

25. Sheffield BS, Fulton R, Kalloger SE, et al. Investigation of PD-L1 Biomarker Testing Methods for PD-1 Axis Inhibition in Non-squamous Non-small Cell Lung Cancer. J Histochem Cytochem. 2016;64:587–600. [PubMed: 27591097]

26. Gaule P, Smithy JW, Toki M, et al. A Quantitative Comparison of Antibodies to Programmed Cell Death 1 Ligand 1. JAMA Oncol. 2017;3:256–259. [PubMed: 27541827]

27. Rimm DL, Han G, Taube JM, et al. A Prospective, Multi-institutional, Pathologist-Based Assessment of 4 Immunohistochemistry Assays for PD-L1 Expression in Non-Small Cell Lung Cancer. JAMA Oncol. 2017;3:1051–1058. [PubMed: 28278348]

28. Rehman JA, Han G, Carvajal-Hausdorf DE, et al. Quantitative and Pathologist-Read comparison of the Heterogeneity of Programmed Death-Ligand 1(PD-L1) expression in Non-Small Cell Lung Cancer. Mod Pathol. 2017;30:340–349. [PubMed: 27834350]

29. Madore J, Vilain RE, Menzies AM, et al. PD-L1 expression in melanoma shows marked heterogeneity within and between patients: implications for anti-PD-1/PD-L1 clinical trials. Pigment Cell Melanoma Res. 2015;28:245–253. [PubMed: 25477049]

30. Casadevall D, Clavé S, Taus Á, et al. Heterogeneity of Tumor and Immune Cell PD-L1 Expression and Lymphocyte Counts in Surgical NSCLC Samples. Clin Lung Cancer 2017;18:682–691.e5. [PubMed: 28549836]

31. Scott ML, Scorer P, Lawson N, et al. Assessment of heterogeneity of PD-L1 expression in NSCLC, HNSCC, and UC with Ventana SP263 assay. J Clin Oncol. 2017 Available at: http://ascopubs.org/doi/abs/10.1200/JCO.2017.35.15_suppl.e14502. Accessed December 29, 2017.

32. Ilie M, Long-Mira E, Bence C, et al. Comparative study of the PD-L1 status between surgically resected specimens and matched biopsies of NSCLC patients reveal major discordances: a potential issue for anti-PD-L1 therapeutic strategies. Ann Oncol. 2016;27:147–153. [PubMed: 26483045]

33. Cooper JS, Porter K, Mallin K, et al. National Cancer Database report on cancer of the head and neck: 10-year update. Head Neck 2009;31:748–758. [PubMed: 19189340]

34. Day D, Hansen AR, Siu LL. Hypopharyngeal cancer: looking back, moving forward. Curr Oncol. 2016;23:221–222. [PubMed: 27536170]

35. Liu Y, Dong Z, Jiang T, et al. Heterogeneity of PD-L1 Expression Among the Different Histological Components and Metastatic Lymph Nodes in Patients With Resected Lung Adenosquamous Carcinoma. Clin Lung Cancer 2018 3 31. doi: 10.1016/j.cllc.2018.02.008. Epub ahead of print.

36. Munari E, Zamboni G, Marconi M, et al. PD-L1 expression heterogeneity in non-small cell lung cancer: evaluation of small biopsies reliability. Oncotarget 2017;8:90123–90131. [PubMed: 29163815]

37. Prasad V, Kaestner V. Nivolumab and pembrolizumab: Monoclonal antibodies against programmed cell death-1 (PD-1) that are interchangeable. Semin Oncol. 2017;44:132–135. [PubMed: 28923211]
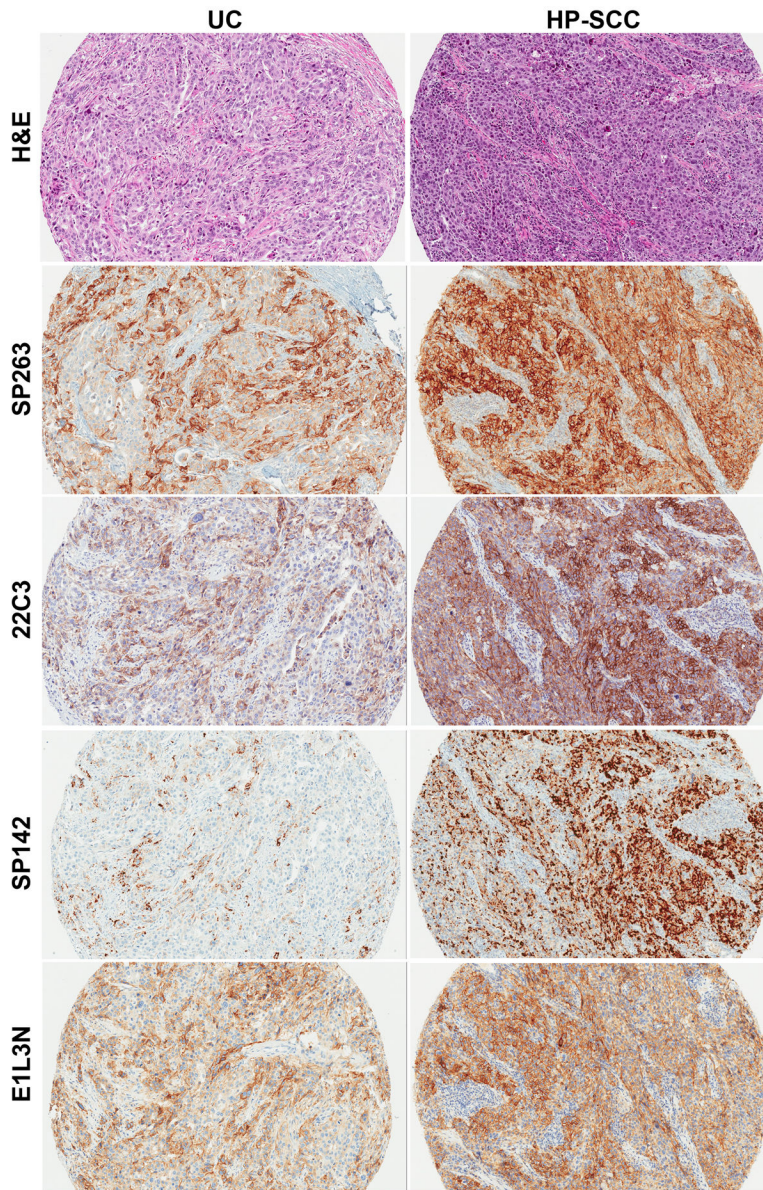
Figure 1.
Representative H&E of urothelial carcinoma (UC) and hypopharyngeal squamous cell carcinoma (HP-SCC) with corresponding PD-L1 staining with each antibody clone.
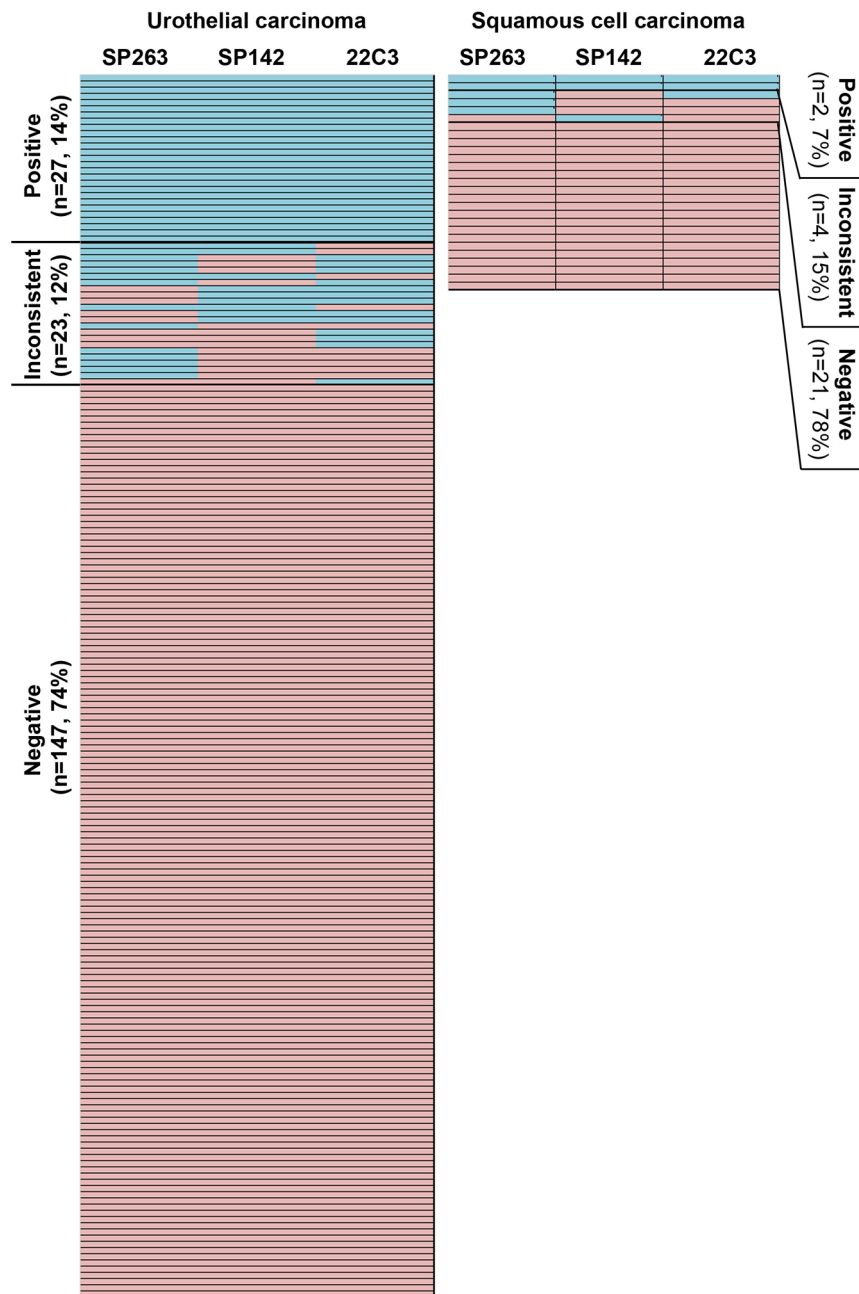
Figure 2. Heat map comparing PD-L1 SP263, SP142 and 22C3 antibody clones in urothelial carcinoma and hypopharyngeal squamous cell carcinoma.

Blue: PD-L1-positive tumors using established clinical algorithms as summarized in table 1.

Orange: PD-L1-negative tumors.

Urothelial carcinoma

Squamous cell carcinoma

		E1L3N			
		UC		HP-SCC	
		Negative	Positive	Negative	Positive
SP263	Negative	149	7	22	0
	Positive	9	32	1	4
SP142	Negative	151	10	23	1
	Positive	12	24	0	3
22C3	Negative	140	17	24	0
	Positive	2	38	1	2

Figure 3. Heat map comparing the performance betweeen E1L3N clone and other antibody clones.

Blue: PD-L1-positive tumors. Orange: PD-L1 negative tumors.
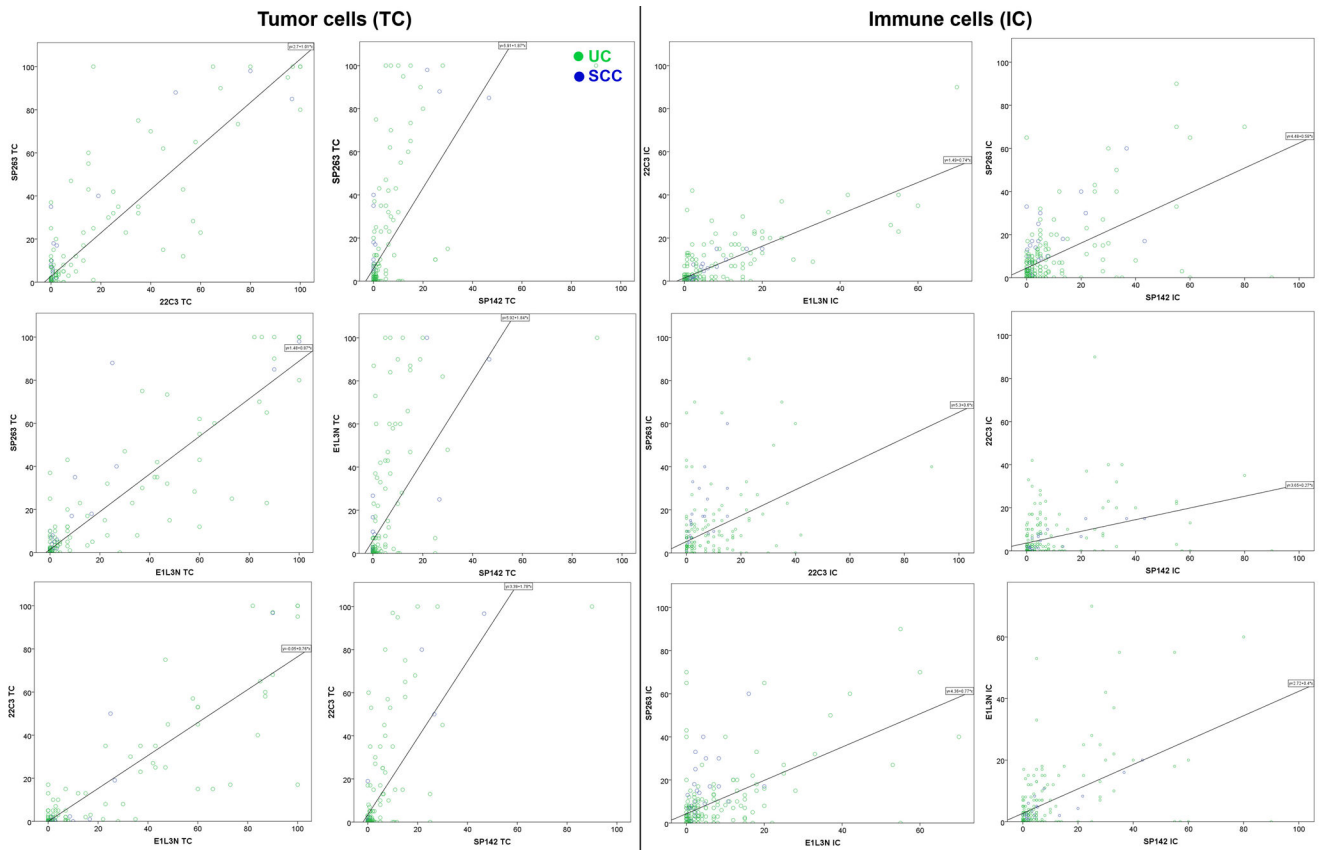
Figure 4. Scatterplots of pairwise comparison of tumor cell (TC, left) and immune cell (IC, right) percentage scoring between different antibody clones.

Blue: HP-SCC; green: UC. A 45-degree diagonal regression line indicated perfect

correlation of TC scoring between the two antibody clones compared.

Table 1.

Clinicopathological features for cases of urothelial carcinoma (UC) of the bladder and hypopharyngeal squamous cell carcinomas (HP-SCC).

	UC (n=197)	HP-SCC (n=27)
Age, years, median (range)	71 (33–93)	64 (49–79)
Sex[a]		
Male	143 (73%)	20 (74%)
Female	54 (27%)	7 (26%)
Histotype (UC)		
Usual	146 (74%)	N/A
Variant/divergent	51 (26%)	N/A
Differentiation (HP-SCC)		
Moderately differentiated	N/A	20 (74%)
Poorly differentiated	N/A	7 (26%)
AJCC pT category		
pT1	6 (3%)	1 (4%)[b]
pT2	27 (14%)	5 (19%)[b]
pT3	108 (55%)	4 (15%)
pT4	56 (28%)	17 (63%)
Lymph node metastases		
Present	79 (40%)	19 (70%)
Absent	109 (55%)	8 (30%)
No nodes	9 (5%)	0 (0%)

[a]Values are expressed as number of patients (percentage of column total) unless otherwise specified.

[b]One patient in pT1 and one patient in pT2 categories were staged post-neoadjuvant therapy.

N/A, not applicable.

Table 2.

Performance comparison among PD-L1 antibodies using Fleiss' kappa analyses to compare performance and intraclass correlation coefficient (ICC) to compare tumor cell (TC) and immune cell (IC) percentages.

Kappa analyses for antibody performance			
	Kappa	**95% confidence interval**	**P Value**
SP142 vs. 22C3	0.791	0.660–0.922	<0.001
SP243 vs. E1L3N	0.762	0.631–0.893	<0.001
22C3 vs. E1L3N	0.743	0.612–0.873	<0.001
SP263 vs. SP142	0.724	0.593–0.855	<0.001
SP263 vs. 22C3	0.706	0.575–0.837	<0.001
SP142 vs. E1L3N	0.639	0.508–0.770	<0.001
ICC analyses for TC percentage			
	ICC	**95% confidence interval**	**P value**
All four clones	0.900	0.870–0.923	<0.001
SP263 vs. E1L3N	0.949	0.933–0.961	<0.001
22C3 vs. E1L3N	0.939	0.918–0.955	<0.001
SP263 vs. 22C3	0.929	0.905–0.947	<0.001
SP142 vs. 22C3	0.619	0.481–0.717	<0.001
SP263 vs. SP142	0.525	0.340–0.652	<0.001
SP142 vs E1L3N	0.500	0.320–0.629	<0.001
ICC analyses for IC percentage			
	ICC	**95% confidence interval**	**P value**
All four clones	0.805	0.76–0.844	<0.001
22C3 vs. E1L3N	0.866	0.826–0.897	<0.001
SP263 vs. SP142	0.722	0.638–0.786	<0.001
SP263 vs. E1L3N	0.695	0.595–0.769	<0.001
SP142 vs. E1L3N	0.667	0.567–0.744	<0.001
SP263 vs. 22C3	0.567	0.435–0.668	<0.001
SP142 vs. 22C3	0.519	0.376–0.630	<0.001