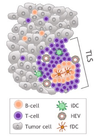
A multi-tumor machine learning model to identify tertiary lymphoid structures (TLS) in histopathological H&E images as a potential clinical biomarker

Vanessa Matos-Cruz[1], Rachel L. Sargent[1], Varsha Chinnaobireddy[2], Maryam Pouryahya[2], George Lee[1], Darren Fahy[2], Christian Kirkup[2], Kathleen Sucipto[2], Sai Chowdary Gullapally[2], Jacqueline Brosnan-Cashman[2], Archit Khosla[2], Nishant Agrawal[2], Ben Glass[2], Sergine Brutus[2], Limin Yu[2], Benjamin Chen[1], Vipul Baxi[1], Scott Ely[1]

[1] Bristol-Myers Squibb, Princeton, NJ; [2] PathAI, Boston, MA

Introduction



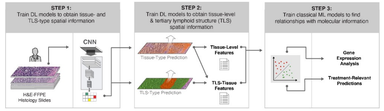
Tertiary lymphoid structures (TLS) are ectopic lymphoid structures composed of B-cells, T-cells, and supportive cells that develop in non-lymphoid organs and are often found in tumors. The criticality and functions of TLS in an adaptive anti-tumor response are still being elucidated, but studies have shown associations between TLS and IO outcomes across multiple indications. These correlations are dependent on TLS maturity and localization within the tumor microenvironment (TME). Currently, identification of TLS in tumors by pathologists is not routine or standardized.

Objectives:
Develop a machine-learning algorithm based on H&E images to score TLS as a clinical biomarker to:
1) Accurately and reproducibly identify TLS regions within the TME
2) Predict TLS subregions and maturity state
3) Extract TLS model-derived features

This algorithm can be deployed in an exploratory manner to score TLS features in research and trial cohorts to assess its utility as a predictive biomarker and complement immune response measurements.

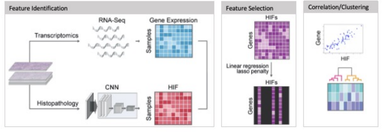
Methods

Figure 1: Model development overview



An AI-based model was trained using pathologist-derived annotations of slides from the TCGA database to analyze H&E images and extract human interpretable features (HIFs) at object and tissue level. HIFs capture specific and biologically-relevant characteristics across multiple indications.

Figure 2: Validation of TLS HIFs by RNAseq analysis correlation



Extracted HIFs were validated by using a published TLS GES ([2]; Table 1).

Table 1. Chemokines encoded by TLS gene expression signature

12-Chemokine gene signature[2]		
CCL2	CCL8	CXCL9
CCL3	CCL18	CXCL10
CCL4	CCL19	CXCL11
CCL5	CCL21	CXCL13

Results

Table 2. TCGA H&E images used for AI model development

Indication	Train	Validation	Test	Total
NSCLC-AD	334	124	57	515
NSCLC-SQ	307	89	50	446
Breast	204	58	863	1125
Bladder	204	59	27	290
Stomach	202	58	141	401

Table 3. Selected AI-model extracted TLS features at slide and object level

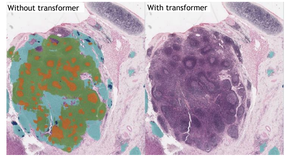
Histological features	Number of annotations
Mature TLS (mTLS)	1631
Germinal center (GC)	1725
Immature TLS (iTLS)	6229
Lymphoid aggregate (LA)	2280
Lymphoid infiltrate	7515
Dense plasma cell infiltrate	2240
Other tissue	6589
Cancer epithelium	3503
Total:	31712

Table 4. Comparison of AI-based vs. manual TLS analyses

Feature	Pathologists	AI-model
Number of TLS	Yes	Yes
TLS stage	Yes	Yes
TLS location	Yes	Yes
TLS coordinates	No	Yes
Pixel-by-pixel TLS area	No	Yes

TLS AI-model extracts similar features that are identifiable by a pathologist. Model extracted HIFs are consistent and scalable.

Figure 3. Development of novel TLS transformer: rules-based post-processing of TLS predictions



TLS AI-model performs post-processing QC to identify and correct TLS predictions based on a series of rules provided by pathologist. Application of a TLS transformer allows for the removal of false-positive TLS predictions on lymph nodes by application of a numerical germinal center cutoff.

H&E TLS AI model qualitative assessment: Algorithm identifies TLS regions within the TME

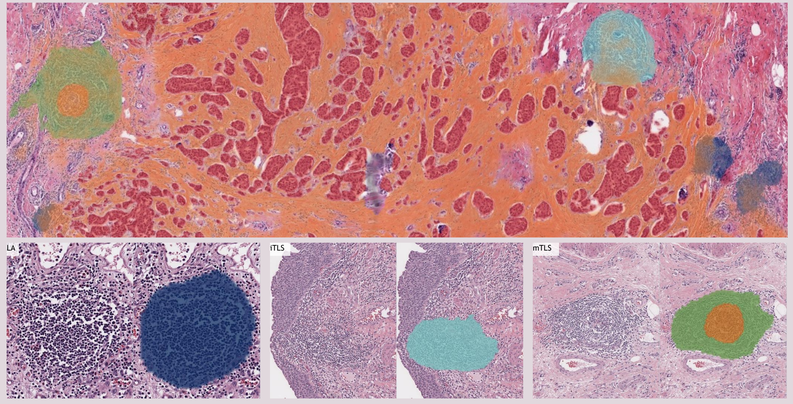
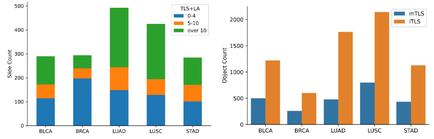


Figure 4. AI-powered algorithm identifies and classifies TLS

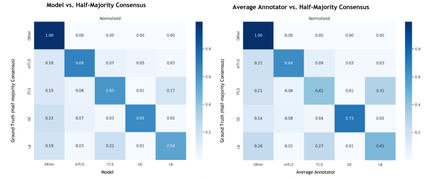
TLS algorithm pseudo-colours tumor (**red**) and stromal (**orange**) microenvironment on a representative H&E image (Top panel). Algorithm classifies TLS in: lymphoid aggregate (**navy**), immature (**cyan**) and mature (**green**) with germinal center (**dark orange**) on a scanned H&E image (bottom panel).

Figure 5. Prevalence of TLS in TCGA datasets



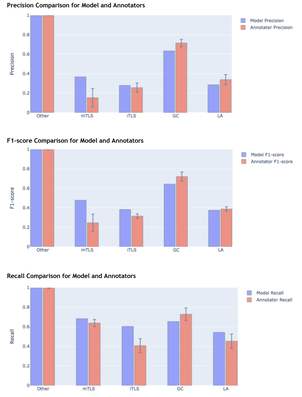
TCGA H&E images utilized for AI-model training displayed a dynamic range of number of TLS and maturity states.

Figure 6. Confusion matrices of model



Exhaustive annotations were collected from five expert pathologists and compared to model predictions.

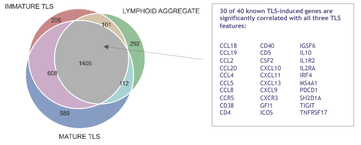
Figure 7. TLS model performance metrics



Pathologist-trained algorithm performs comparably to pathologists in identifying and classifying TLS in H&E images.

Figure 8. TLS correlate with gene expression and prognosis in BRCA

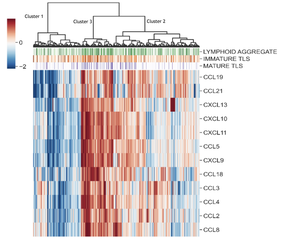
Differential gene expression analysis in BRCA



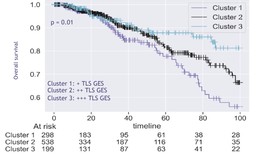
30 of 40 known TLS-induced genes are significantly correlated with all three TLS features:

CCL18	CD40	IGSF6
CCL19	CD5	IL10
CCL2	CSF2	IL18Z
CCL20	CXCL10	IL2RA
CCL4	CXCL11	IRF4
CCL5	CXCL13	MS4A1
CCL8	CXCL9	PDCD1
CCR5	CXCR3	SH2D1A
CD38	GFI1	TIGIT
CD4	ICOS	TNFRSF17

HIFs correlate with TLS GES in BRCA



High TLS-GES correlates with improved prognosis



Differential gene expression analysis in BRCA: Identification of genes that are differentially expressed based on TLS stage. 30 of 40 published TLS-associated gene are upregulated in all groups.

HIFs correlate with TLS GES in BRCA: Hierarchical clustering analysis using a 12-chemokine GES correlated with HIFs (LA, iTLS, mTLS).

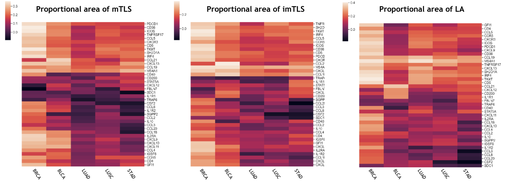
High TLS-GES correlates with improved prognosis: Higher TLS-GES levels correlate with better overall survival.

Table 5. Model-derived features correlated with a published TLS gene signature[2]

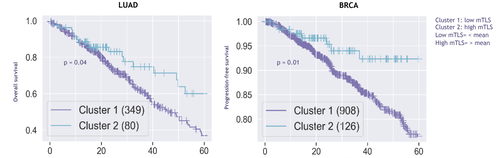
Calculated features
Area proportion immature TLS
Mean perimeter immature TLS
Area proportion mature TLS
Area prop lymphoid aggregate

Figure 9. Validation of TLS-extracted HIFs by correlation analysis to known TLS-GES

AI-model features correlate to TLS-associated genes across cancer types



mTLS proportional area demonstrates prognostic value in LUAD and BRCA



AI-model features correlate to TLS-associated genes across cancer types: Proportional area of of different TLS stages correlate with a subset of TLS-associated genes.

mTLS proportional area demonstrates prognostic value in LUAD and BRCA: High mTLS correlated with better overall survival in LUAD and progression-free survival in BRCA.

Conclusions and Future Directions

- We developed a model that accurately detects TLS regions in multiple tumor types, both within tumor and within the surrounding tissue.
- Model-derived TLS features were associated with expression of a well-accepted published 12-chemokine TLS gene signature[(2)].
- Clustering by level of expression of the 12-chemokine gene panel is concordant with the TLS model-derived features.
- TLS model-derived features showed prognostic value in multiple tumor types.
- Moving forward, we plan to expand this analysis into clinical cohorts in order to assess the value of TLS as a prognostic and predictive biomarker in the immuno-oncology setting.

References

1. Sautès-Fridman C, Petitprez F, Calderaro J, Fridman WH. Tertiary lymphoid structures in the era of cancer immunotherapy. Nat Rev Cancer. 2019 Jun;19(6):307-325.

2. Zhu G, Falahat R, Wang K, Maillous A, Artzi N, Mulé JJ. Tumor-Associated Tertiary Lymphoid Structures: Gene-Expression Profiling and Their Bioengineering. Front Immunol. 2017 Jun 30;8:767.

Declaration of Interests

Bristol Myers Squibb employees: VMC, RLS, GL, BC, VB, SE

PathAI employees; VC, MP, DF, CK, KS, SCG, JBC, AK, NA, BG, SB, LY

Scientific Content on Demand

To request a copy of this poster:

Scan QR code
via a barcode reader application

QR codes are valid for 30 days after the congress presentation date.

Email: vanessa.matos-cruz@bms.com