A WEAKLY SUPERVISED DEEP LEARNING FRAMEWORK TO PREDICT KRAS-STK11 AND KRAS-TP53 CO-MUTATIONS IN LUNG ADENOCARCINOMAS USING H&E TISSUE SECTIONS

1Daniel Jiménez-Sánchez*, 1Carlos De Andrea, 1Álvaro López-Jánerie, 1María D Lozano, 2Carlos Ortiz-de-Solórzano, 1Aida Romano-Martínez, 1Cristina Sairia, 2Luis Montuenga, 1Ignacio Melero, 2Rubén Pico, 3Clínica Universidad de Navarra, Pamplona, Spain; 2Cima-University of Navarra, Pamplona, Spain

Background Lung adenocarcinomas (LUAD) with co-mutations in KRAS (K-only) and the SKT11/LKB1 (KS) or TP53 (KP) genes define patient subgroups with distinct responses to anti-PD1/PD-L1 immunotherapy. In fact, multicentric studies showed that objective response rates to PD-1 blockade differed significantly among KS (7.4%), KP (35.7%), and K-only (28.6%) subgroups.1 The association of such specific genetic profiling with morphological patterns assessed on routine H&E tissue slides may contribute to a better selection for personalized immunotherapy treatments.2

Methods We developed a weakly supervised deep learning (WSDL) model to predict the mutational status of LUAD patients using routine H&E tissue slides. N=125 KRAS-mutated patients with genomic profile available were obtained from two public databases (CPTAC4 and TCGA)5 and one in-house cohort (Clínica Universidad de Navarra). 59 patients were K-only, 36 KS, and 30 KP. Our developed model was composed of two neural networks. A convolutional neural network that learns cellular features from 90x90 pixel image patches unsupervisedly, and a graph neural network that learns patient-specific patterns using only the patient mutational status. Abundances of these patterns predict the patient’s mutation type. To assess the predictive value of our WSDL model a five-fold cross-validation scheme was used. A Mann-Whitney test was applied to associate learned tissue patterns with patient mutations.

Results Figure 1(a) shows the ROC curves for the model for predicting patient co-mutations. AUC for K-only vs. KP mutations was 0.76 with a 95% CI of [0.66,0.86]. AUC for K-only vs. KS was 0.64 with a 95% CI of [0.54,0.75]. KP vs. KS was 0.78 with a 95% CI of [0.67,0.88]. Figure 1(b) shows, as an example, four WSDL-identified tissue patterns consisting of image patches containing acinar tumor, acinar tumor margin, stromal lymphocytes, and stroma. Figure 1(c) shows abundances of the total of tissue patterns learned showing the complexity and heterogeneity of LUADs. (d,c) Quantification figures representing the complexity and heterogeneity of LUADs. (d,c) Quantifications of two tissue patterns across patient types, and the nine patient-specific patterns using only the patient mutational status. Abundances of these patterns predict the patient’s mutation type. To assess the predictive value of our WSDL model a five-fold cross-validation scheme was used. A Mann-Whitney test was applied to associate learned tissue patterns with patient mutations.

Conclusions WSDL learns tissue patterns without the requirement of manual expert annotations, potentially revealing previously unappreciated or underappreciated facets of the tumor linked to specific mutation types. This model can be especially useful in complex tasks such as the determination of LUAD co-mutations from H&E tissue slides. A validation study in two independent cohorts is ongoing.

REFERENCES

Ethics Approval The study was approved by the University of Navarra Ethics Board, approval number 2019.111