

Basic Study

Prediction of genetic alterations from gastric cancer histopathology images using a fully automated deep learning approach

Hyun-Jong Jang, Ahwon Lee, Jun Kang, In Hye Song, Sung Hak Lee

ORCID number: Hyun-Jong Jang 0000-0003-4535-1560; Ahwon Lee 0000-0002-2523-9531; Jun Kang 0000-0002-7967-0917; In Hye Song 0000-0001-6325-3548; Sung Hak Lee 0000-0003-1020-5838.

Author contributions: Jang HJ and Lee SH were responsible for the study concept and design; Lee SH enrolled the cohorts and collected clinicopathological data from patients; Lee A, Kang J, Song IH and Lee SH performed the assays; Jang HJ, Lee A, Kang J, Song IH and Lee SH analyzed data; Jang HJ and Lee SH wrote the manuscript.

Institutional review board

statement: The study was reviewed and approved by the Institutional Review Board of the College of Medicine at the Catholic University of Korea (KC19SESI0787).

Conflict-of-interest statement: The authors declare that they have no conflicts of interest.

Data sharing statement: No additional data are available.

ARRIVE guidelines statement: The authors have read the ARRIVE guidelines, and the manuscript was prepared and revised according to the ARRIVE guidelines.

Hyun-Jong Jang, Catholic Big Data Integration Center, Department of Physiology, College of Medicine, The Catholic University of Korea, Seoul 06591, South Korea

Ahwon Lee, Jun Kang, In Hye Song, Sung Hak Lee, Department of Hospital Pathology, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul 06591, South Korea

Corresponding author: Sung Hak Lee, MD, PhD, Associate Professor, Department of Hospital Pathology, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, 222, Banpo-daero, Seocho-gu, Seoul 06591, South Korea. hakjjang@catholic.ac.kr

Abstract

BACKGROUND

Studies correlating specific genetic mutations and treatment response are ongoing to establish an effective treatment strategy for gastric cancer (GC). To facilitate this research, a cost- and time-effective method to analyze the mutational status is necessary. Deep learning (DL) has been successfully applied to analyze hematoxylin and eosin (H and E)-stained tissue slide images.

AIM

To test the feasibility of DL-based classifiers for the frequently occurring mutations from the H and E-stained GC tissue whole slide images (WSIs).

METHODS

From the GC dataset of The Cancer Genome Atlas (TCGA-STAD), wild-type/mutation classifiers for CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes were trained on 360 × 360-pixel patches of tissue images.

RESULTS

The area under the curve (AUC) for the receiver operating characteristic (ROC) curves ranged from 0.727 to 0.862 for the TCGA frozen WSIs and 0.661 to 0.858 for the TCGA formalin-fixed paraffin-embedded (FFPE) WSIs. The performance of the classifier can be improved by adding new FFPE WSI training dataset from our institute. The classifiers trained for mutation prediction in colorectal cancer completely failed to predict the mutational status in GC, indicating that DL-based mutation classifiers are incompatible between different cancers.

CONCLUSION

Supported by the National Research Foundation of Korea funded by the Korea government (Ministry of Science and ICT), No. 2019R1F1A1062367.

Country/Territory of origin: South Korea

Specialty type: Pathology

Provenance and peer review: Unsolicited article; Externally peer reviewed.

Peer-review report's scientific quality classification

Grade A (Excellent): 0
 Grade B (Very good): B, B
 Grade C (Good): C
 Grade D (Fair): 0
 Grade E (Poor): 0

Open-Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <https://creativecommons.org/licenses/by-nc/4.0/>

Received: July 29, 2021
Peer-review started: July 29, 2021
First decision: August 19, 2021
Revised: September 5, 2021
Accepted: November 13, 2021
Article in press: November 13, 2021
Published online: November 28, 2021

P-Reviewer: Kumar I, Nearchou IP, Salvi M
S-Editor: Wang LL
L-Editor: A
P-Editor: Wang LL



This study concluded that DL could predict genetic mutations in H and E-stained tissue slides when they are trained with appropriate tissue data.

Key Words: Gastric cancer; Mutation; Deep learning; Digital pathology; Formalin-fixed paraffin-embedded

©The Author(s) 2021. Published by Baishideng Publishing Group Inc. All rights reserved.

Core Tip: Recently, deep learning approach has been implemented to predict the mutational status from hematoxylin and eosin (H and E)-stained tissue images of diverse tumors. The aim of our study was to evaluate the feasibility of classifiers for mutations in the CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes in gastric cancer tissues. The area under the curves for receiver operating characteristic curves ranged from 0.727 to 0.862 for the The Cancer Genome Atlas (TCGA) frozen tissues and 0.661 to 0.858 for the TCGA formalin-fixed paraffin-embedded tissues. This study confirmed that deep learning-based classifiers can predict major mutations from the H and E-stained gastric cancer whole slide images when they are trained with appropriate data.

Citation: Jang HJ, Lee A, Kang J, Song IH, Lee SH. Prediction of genetic alterations from gastric cancer histopathology images using a fully automated deep learning approach. *World J Gastroenterol* 2021; 27(44): 7687-7704

URL: <https://www.wjgnet.com/1007-9327/full/v27/i44/7687.htm>

DOI: <https://dx.doi.org/10.3748/wjg.v27.i44.7687>

INTRODUCTION

Molecular tests to identify specific mutations in solid tumors have improved our ability to stratify cancer patients for more selective treatment regimens[1]. Therefore, molecular tests to detect various mutations are recommended for some tumors, including EGFR mutations in lung cancer, KRAS in colorectal cancer, and BRAF in melanoma. However, it is not routinely applied to cancer patients because molecular tests are not cost- and time-efficient[2]. Furthermore, the clinical significance of many mutations is still not well understood. For example, mutation profiling of gastric cancer (GC) is still proceeding, and the meaning of each mutation is not clearly understood[3]. GC is the fifth most common cancer and the third leading cause of cancer-related deaths worldwide[4]. It is important to evaluate the relationship between the mutational status and clinical characteristics of GC to improve the clinical outcomes of GC patients. Furthermore, many targeted drugs for treating various tumors are not effective in GC therapy because GC is not enriched with known driver mutations[5]. Therefore, research to characterize the roles of GC-related genes on the clinical behavior of tumors and the potential response to targeted therapies will have immense importance for the improvement of treatment response in GC[6]. A cost- and time-effective method to determine the mutational status of GC patients is necessary to promote these studies.

Recently, deep learning (DL) has been increasingly implemented to predict the mutational status from hematoxylin and eosin (H and E)-stained tissue slides of various cancers[7-11]. The H and E-stained tissue slides were made for almost all cancer patients for basic diagnostic studies by pathologists[12]. Therefore, mutation prediction from the H and E-stained tissue slide based on a computational method can be a cost- and time-effective alternative tool for conventional molecular tests[13-15]. Although it has long been recognized that the morphological features of tissue architecture reflect the underlying molecular alterations[16,17], the features are not easily identifiable by human evaluators[18,19]. DL offers an alternative solution to overcome the limitations of a visual examination of tissue morphology by pathologists. By combining feature learning and model fitting in a unified step, DL can capture the most discriminative features for a given task directly from a large set of tissue images [20]. Digitization of tissue slides has been rapidly increasing after the approval of digitized whole-slide images (WSIs) for diagnostic purposes[21]. Digitized tissue data

are rapidly accumulating with their associated mutational profiles. Therefore, the DL-based analysis of tissue slides for the mutational status of cancer tissues has immense potential as an alternative or complementary method for conventional molecular tests.

Based on the potential of DL for the detection of mutations from digitized tissue slides, in a previous study, we successfully built DL-based classifiers for the prediction of mutational status of APC, KRAS, PIK3CA, SMAD4, and TP53 genes in colorectal cancer tissue slides[11]. This study investigated the feasibility of classifiers for mutations in the CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes in GC tissues. First, the classifiers were trained and tested for GC tissue slides from The Cancer Genome Atlas (TCGA). The generalizability of the classifiers was tested using an external dataset. Then, new classifiers were trained for combined datasets from TCGA and external datasets to investigate the effect of the extended datasets. The results suggest that it is feasible to predict mutational status directly from tissue slides with deep learning-based classifiers. Finally, as the classifiers for KRAS, PIK3CA, and TP53 mutations for both colorectal and GC were available, we also analyzed the generalizability of the DL-based mutation classifiers trained for different cancer types.

MATERIALS AND METHODS

Part I: Tests with The Cancer Genome Atlas whole-slide image datasets

Patient cohort: The Cancer Genome Atlas (TCGA) provides extensive archives of digital pathology slides with multi-omics test results to test the possibility of tissue-based mutation detection[22]. After a carefully review of all the WSIs in the TCGA GC dataset (TCGA-STAD), we eliminated WSIs with poor scan quality and very small tumor contents. We selected slides from 25, 19, 34, 64, and 160 patients, which were confirmed to have mutations in CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes, respectively. There were more than two slides for many patients in the TCGA dataset, with a maximum of four slides for some patients. However, in many cases, one or two slides contained only normal tissues. We excluded normal slides and selected a maximum of two tumor-containing slides per patient. The final number of frozen tissue slides was 34, 26, 50, 94, and 221 and that of formalin-fixed paraffin-embedded (FFPE) tissue slides was 27, 19, 34, 66, and 174 for CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes, respectively. We selected 183 patients with wild-type CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes. Therefore, the same patients with wild-type genes for CDH1, ERBB2, KRAS, PIK3CA, and TP53 can be involved in the training of every classifier as a non-mutated group. This may help the comparison of the different classifiers more standardized because they all have the same group of patients as the wild-type group. The TCGA IDs of the patients in each group are listed in [Supplementary Table 1](#). Our previous studies recognized that a DL model cannot perform optimally for both training and testing unless the dataset is forced to have similar amounts of data between classes[23]. Therefore, we limited the difference in patient numbers between the mutation and wild-type groups to less than 1.4 fold by random sampling. For example, only 35 of the 183 wild-type patients were randomly selected as the CDH1 wild-type group because there were only 25 CDH1 mutated patients. Ten-fold cross-validation was performed based on these randomly sampled wild-type patients. However, the classifiers yielded better results when the tumor patches from all wild-type patients other than the test sets were randomly sampled to match the 1.4 fold data ratio of wild-type/mutation groups for training, as this strategy could include a greater variety of tissue images. Therefore, we included all wild-type patients other than the test sets during training and randomly selected patients during testing.

Deep learning model: In general, a WSI is too large to be analyzed simultaneously using a deep neural network. Therefore, the analysis results for small image patches are integrated for conclusion. We divided a WSI into non-overlapping patches of 360×360 pixel tissue images at $20\times$ magnification to detect mutational status. To make the classification process fully automated, artifacts in the WSIs such as air bubbles, compression artifacts, out-of-focus blur, pen markings, tissue folding, and white background should be removed automatically. A simple convolutional neural network (CNN), termed as tissue/non-tissue classifier, was trained to discriminate these various artifacts all at once. The structure of the tissue/non-tissue classifier was described in our previous study[11]. The tissue/non-tissue classifier could filter out almost 99.9% of the improper tissue patches. Then, tissue patches classified as “improper” by the tissue/non-tissue classifier were removed, and the remaining

“proper” tissue patches were collected. For the tumor or mutation classifiers described below, only proper tissue patches were analyzed (Figure 1).

Morphologic features reflecting mutations in specific genes might be expressed mainly in tumor tissues rather than normal tissues[24,25]. Therefore, tumor tissues should be separated from the WSI to predict the mutational status of the WSI. In a previous study, we successfully built normal/tumor classifiers for various tumors, including GC[26]. We concluded that frozen and FFPE slides should be separately analyzed using a deep neural network due to their different morphologic features. Thus, we adopted the normal/tumor classifiers for frozen and FFPE tissue slides from a previous study to delineate the normal/tumor gastric tissues for the frozen and FFPE slides of the TCGA-STAD dataset in the present study. Mutation classifiers were trained separately for the selected tumor patches for frozen and FFPE tissues. We selected tumor patches with a tumor probability higher than 0.9 to collect tissue patches with evident tumor features. We adopted a patient-level ten-fold cross-validation to completely characterize the TCGA-STAD dataset. Therefore, patients in each mutation/wild-type group for the five genes were separated into ten different folds, and one of the ten folds was used to test the classifiers trained with data from the other nine folds. Therefore, ten different classifiers were trained and tested for each group. The same label for all tumor tissue patches in a WSI as either ‘wild-type’ or ‘mutated’ were assigned based on the mutational status of the patient. Thereafter, the Inception-v3 model, a widely used CNN architecture, was trained to classify the tumor patches into ‘wild-type’ or ‘mutated’ tissues, as in our previous study on mutation prediction in colorectal cancer[11]. We fully trained the network from the beginning and did not adopt a transfer-learning scheme. The average probability of all tumor patches in a WSI was calculated to determine the slide-level mutation probability of a WSI. The Inception-v3 model was implemented using the TensorFlow DL library (<http://tensorflow.org>), and the network was trained with a mini-batch size of 128 and cross-entropy loss function as a loss function. For training, we used the RMSProp optimizer, with an initial learning rate of 0.1, weight decay of 0.9, momentum of 0.9, and epsilon of 1.0. Ten percent of the training slides were used as the validation dataset, and training was stopped when the loss for the validation data started to increase. Data augmentation techniques, including random horizontal/vertical flipping and random rotations by 90°, were applied to the tissue patches during training. Color normalization was applied to the tissue patches to avoid the effect of stain differences[27,28]. At least five classifiers were trained on each fold of mutation for the frozen and FFPE WSIs separately. The classifier with the best area under the curve (AUC) for the receiver operating characteristic (ROC) curves on the test dataset was included in the results. The ROC curves for fold with the lowest AUC, highest AUC, and the concatenated results for data from all ten folds are shown in the figures.

In summary, a WSI is analyzed as follows: 1. The whole slide is split into non-overlapping 360 × 360 pixel tissue patches, 2. Proper tissue patches are selected by tissue/non-tissue classifier, 3. Only tumor patches with tumor probability higher than 0.9 are selected, 4. High probability tumor patches are classified by each wild-type/mutation classifier, 5. The probabilities of tumor patches are averaged to obtain the slide-level mutation probability. The number of tissue patches used for the training of all mutation prediction models is summarized in [Supplementary Table 2](#). The average number of training epochs for each classifier is summarized in [Supplementary Table 3](#).

Part II: Tests on the external cohorts

Patient cohort: GC tissue slides were collected from 96 patients who had previously undergone surgical resection at Seoul St. Mary’s Hospital between 2017 and 2020 (SSMH dataset). An Aperio slide scanner (Leica Biosystems) was used to scan the FFPE slides. The Institutional Review Board of the College of Medicine at the Catholic University of Korea approved this study (KC19SESI0787).

Mutation prediction on SSMH dataset: For CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes, 6, 6, 12, 11, and 39 patients were confirmed to have the mutations, respectively. Thirty-eight patients had wild-type genes for all five genes. For CDH1, ERBB2, KRAS, and PIK3CA genes, we selected the number of wild-type patients to be 1.4 times that of mutated patients. For TP53, all 38 patients with wild-type genes were enrolled. The normal/tumor classifier for TCGA FFPE tissues was also used to discriminate the tumor tissue patches of SSMH WSIs. Our previous study showed that the normal/tumor classifier for TCGA-STAD was valid for SSMH FFPE slides[29]. First, the mutational status of the SSMH slides was analyzed by classifiers trained on TCGA-

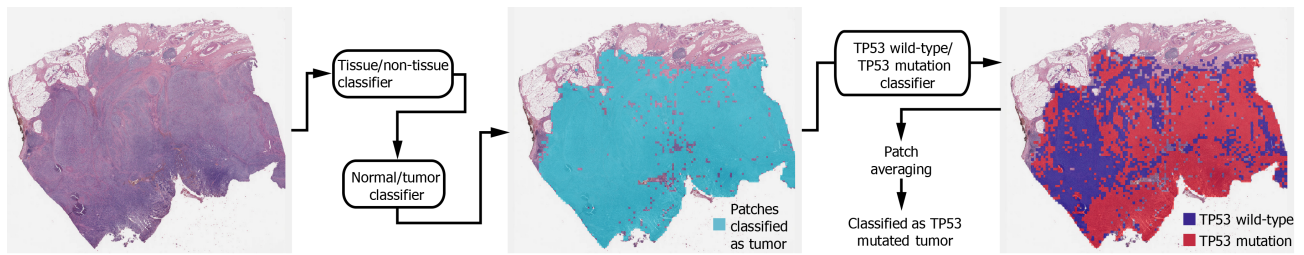


Figure 1 Workflow for the fully automated prediction of mutation. Tissue image patches with tumor probability higher than 0.9 were selected by sequential application of the tissue/non-tissue and normal/tumor classifiers. Then the tumor patches were classified into the wild-type or mutated patches. The patch-level probabilities of mutation are averaged to yield the slide-level probability.

STAD FFPE WSIs. Subsequently, new classifiers were trained using both TCGA and SSMH FFPE tissues. Patient-level three-fold cross validation was applied to the SSMH datasets because the number of mutated patients was not sufficient for ten-fold cross-validation.

Statistical analysis

To demonstrate the performance of each classifier, the ROC curves and their AUCs are presented in the figures. For the concatenated results from all ten folds, 95% confidence intervals (CIs) were also presented using the percentile bootstrap method. In addition, the accuracy, sensitivity, specificity, and F1 score of the classification results of mutation prediction models with cutoff values for maximal Youden index (sensitivity + specificity - 1) were presented. We used a permutation test with 1000 iterations to compare the differences between the two paired or unpaired ROC curves when a comparison was necessary[30]. Statistical significance was set at $P < 0.05$.

RESULTS

Tissue patches with high tumor probability were automatically collected from a WSI by sequentially applying the tissue/non-tissue and normal/tumor classifiers to 360×360 pixels tissue image patches (Figure 1). Then, classifiers to distinguish the mutational status of CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes in the tumor tissue patches from the TCGA-STAD frozen and FFPE WSI datasets were separately trained with a patient-level ten-fold cross-validation scheme.

The classification results of the TCGA-STAD WSIs are presented in Figures 2 to 6 for CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes. Results for the frozen and FFPE tissues are presented in the upper and lower part of each figure, respectively. Panels A and C demonstrated the representative binary heatmaps of tissue patches classified as wild-type or mutated tissues. The WSIs with gene mutation correctly classified as mutation, with wild-type gene correctly classified as wild-type, with gene mutation falsely classified as wild-type, and with wild-type gene falsely classified as mutation are presented from left to right for panels A and C. The binary heatmaps were drawn with the wild-type/mutation discrimination threshold set to 0.5. We simply set the threshold to 0.5, because every classifier for different folds had different optimal thresholds. Slide-level ROC curves for folds with the lowest and highest AUCs are presented to demonstrate the differences in the performance between folds (left and middle ROC curves in each figure). Finally, the slide-level ROC curves for the concatenated results from all ten folds were used to infer the overall performance (right ROC curves). The results for the CDH1 gene are shown in Figure 2. The AUCs per fold ranged from 0.833 to 1.000 for frozen WSIs and from 0.833 to 1.000 for FFPE WSIs. The AUCs for the concatenated results were 0.842 (95%CI: 0.749-0.936) and 0.781 (95%CI: 0.645-0.917) for frozen and FFPE WSIs, respectively. For ERBB2 (Figure 3), the lowest and highest AUCs per fold were 0.667 and 1.000, respectively, for both frozen and FFPE WSIs. The concatenated AUCs were 0.751 (95%CI: 0.631-0.871) and 0.661 (95%CI: 0.501-0.821), respectively. For the KRAS gene (Figure 4), the AUCs per fold were between 0.775 and 1.000 for frozen WSIs and between 0.750 and 1.000 for FFPE WSIs. The concatenated AUCs were 0.793 (95%CI: 0.706-0.879) and 0.858 (95%CI: 0.738-0.979) for frozen and FFPE WSIs, respectively. For the PIK3CA gene (Figure 5), the concatenated AUC for the frozen WSIs was 0.862 (95%CI: 0.809-0.916), with a range of 0.705

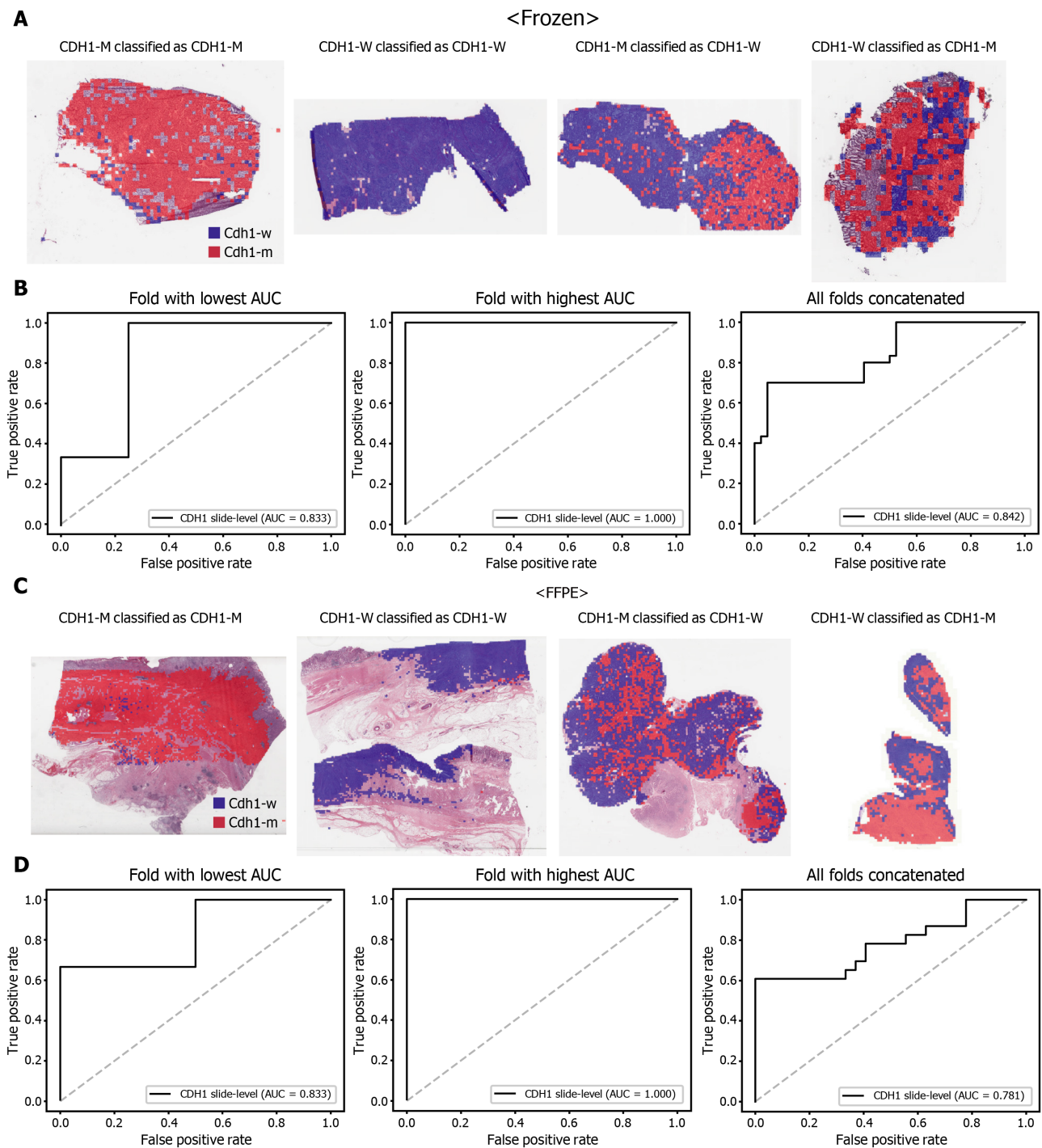


Figure 2 Classification results of CDH1 gene in the The Cancer Genome Atlas gastric cancer dataset. A: Representative whole slide images (WSIs) of the frozen slides with CDH1 gene mutation correctly classified as mutation, with wild-type gene correctly classified as wild-type, with gene mutation falsely classified as wild-type, and with wild-type gene falsely classified as mutation, from left to right; B: Receiver operating characteristic (ROC) curves for the fold with lowest area under the curve (AUC), for the fold with highest AUC, and for the concatenated results of all ten folds, from left to right, obtained with the classifiers trained with the frozen tissues; C and D: Same as A and B but the results were for the formalin-fixed paraffin-embedded WSIs. CDH1-M: CDH1 mutated, CDH1-W: CDH1 wild-type.

to 0.990. For FFPE WSIs, the lowest and highest AUCs per fold were 0.675 and 1.000, respectively, yielding a concatenated AUC of 0.828 (95% CI: 0.750-0.907). Lastly, the results for the TP53 gene are presented in Figure 6. The AUCs per fold were between 0.666 to 0.810 for frozen WSIs and between 0.702 to 0.847 for FFPE WSIs. The concatenated AUCs were 0.727 (95% CI: 0.683-0.771) and 0.727 (95% CI: 0.671-0.784) for frozen and FFPE WSIs, respectively. For the colorectal cancer dataset from TCGA, mutation classification results for frozen tissues were better than those for FFPE tissues in some genes[11]. However, there were no significant differences between the frozen and FFPE tissues in the TCGA-STAD dataset ($P = 0.491, 0.431, 0.187, 0.321, \text{ and } 0.613$

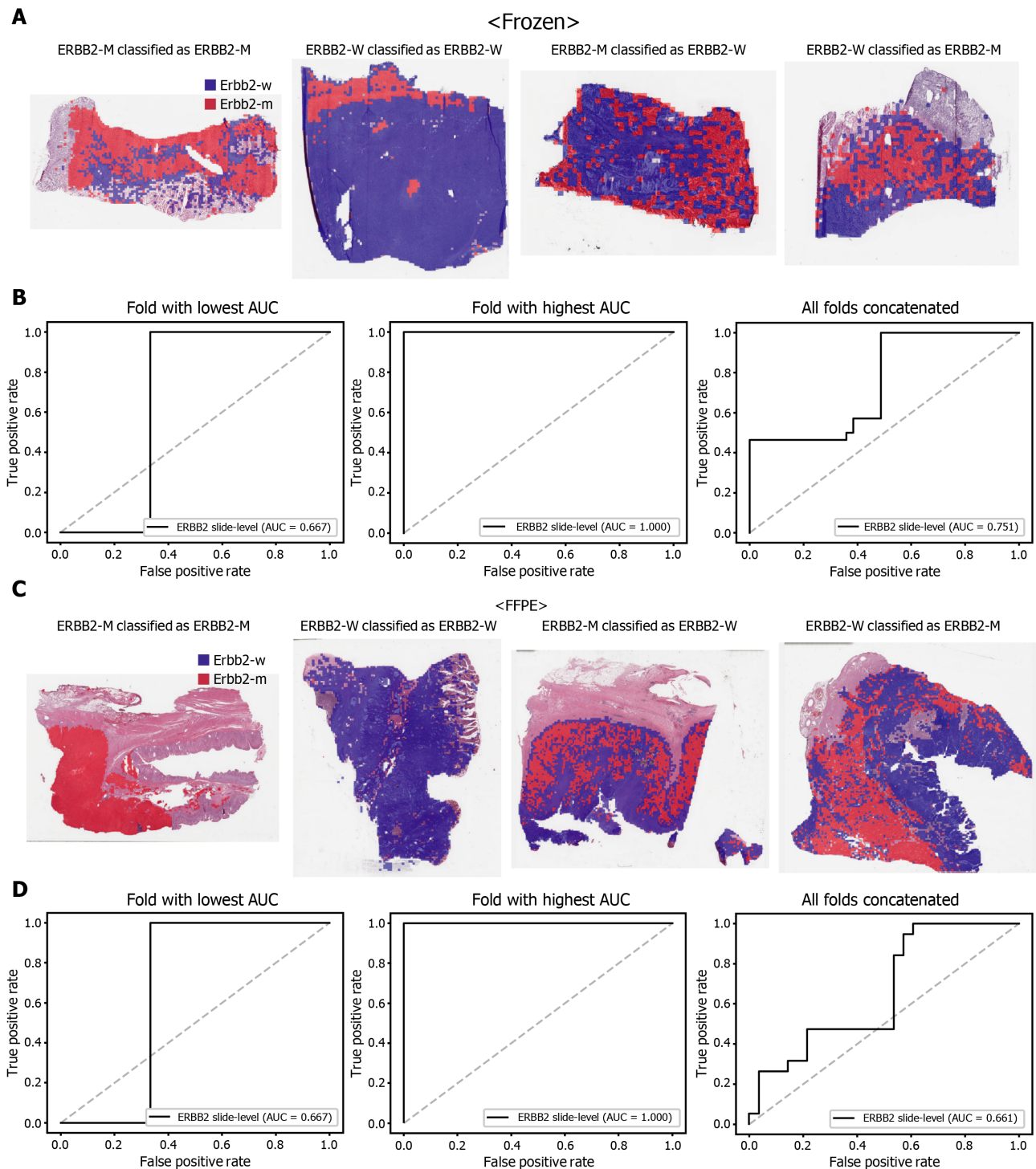


Figure 3 Classification results of ERBB2 gene in the The Cancer Genome Atlas gastric cancer dataset. The configuration of the figure is the same as in Figure 2. A and B: Upper panels are results for the frozen tissue and lower panels; C and D: Results for the formalin-fixed paraffin-embedded tissues. ERBB2-M: ERBB2 mutated, ERBB2-W: ERBB2 wild-type.

between the concatenated AUCs for the frozen and FFPE tissues by Venkatraman’s permutation test for unpaired ROC curves for CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes, respectively. For a clearer assessment of the performance of each model, the accuracy, sensitivity, specificity, and F1 score of the classification results are presented in Table 1.

The performance of a DL model on an external dataset should be tested to validate the generalizability of the trained model. Therefore, we collected GC FFPE WSIs with matching mutation data from Seoul St. Mary’s Hospital (SSMH dataset). The normal/tumor classifier for TCGA-STAD FFPE tissues was also applied to select tissue patches with high tumor probabilities. Thereafter, the mutation classifier for each gene

Table 1 Accuracy, sensitivity, specificity, and F1 score of the classification results of mutation prediction models for the The Cancer Genome Atlas datasets

	Accuracy	Sensitivity	Specificity	F1 score
TCGA Frozen Tissue Slides				
CDH1	0.847	0.700	0.952	0.792
ERBB2	0.716	1.000	0.512	0.746
KRAS	0.773	0.883	0.708	0.745
PIK3CA	0.834	0.771	0.884	0.806
TP53	0.667	0.743	0.602	0.673
TCGA FFPE Tissue Slides				
CDH1	0.820	0.608	1.000	0.756
ERBB2	0.574	1.000	0.285	0.655
KRAS	0.894	0.727	1.000	0.842
PIK3CA	0.803	0.629	0.923	0.723
TP53	0.673	0.678	0.668	0.659

TCGA: The Cancer Genome Atlas; FFPE: Formalin-fixed paraffin-embedded.

trained on the TCGA-STAD FFPE tissues was tested on the SSMH dataset. The slide-level ROC curves for the CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes are presented in [Supplementary Figure 1](#). The AUCs for CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes were 0.667, 0.630, 0.657, 0.688, and 0.572, respectively. For the KRAS, PIK3CA, and TP53 genes, the performance of the TCGA-trained mutation classifiers on the SSMH dataset were worse than that of the TCGA dataset ($P = 0.389$, $P = 0.849$, $P < 0.05$, $P < 0.05$, and $P < 0.05$ for CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes, respectively, by Venkatraman's permutation test for unpaired ROC curves). These results demonstrate that the mutation classifiers trained with TCGA-STAD WSI datasets had limited generalizability. It is of interest if the performance can be enhanced by training the classifiers with expanded datasets, including both TCGA and SSMH datasets. Cancer tissues from different ethnic groups can show different features[16,19]; therefore, the performance of the classifiers can be improved by mixing the datasets. When the classifiers trained with the mixed datasets were used, the performance on the SSMH dataset was generally improved because the SSMH data were included in the training data in this setting (Figures 7 and 8). The AUCs became 0.778, 0.833, 0.838, 0.761, and 0.775 for CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes, respectively ($P = 0.234$, $P < 0.05$, $P < 0.05$, $P = 0.217$, and $P < 0.05$ between the ROCs of classification results by classifiers trained on the TCGA-STAD dataset and mixed dataset for CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes, respectively, by Venkatraman's permutation test for paired ROC curves). Furthermore, the performance on the TCGA-STAD FFPE dataset was also generally improved by the new classifiers trained on both datasets, except for the PIK3CA gene, which showed worse results ([Supplementary Figure 2](#)). The AUCs were 0.918, 0.872, 0.885, 0.766, and 0.820 for the CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes, respectively ($P < 0.05$, $P < 0.05$, $P = 0.216$, $P < 0.05$, and $P < 0.05$ compared with the TCGA-trained classifiers by Venkatraman's permutation test for paired ROC curves). The accuracy, sensitivity, specificity, and F1 score of the classification results of mutation prediction models trained with both SSMH and TCGA datasets are presented in [Supplementary Table 4](#).

Another interesting question is whether the DL-based classifiers for mutational status can be compatible with other types of cancers. We already built the mutation classifiers for KRAS, PIK3CA, and TP53 genes in the colorectal cancer dataset of TCGA in a previous study[11]. Therefore, we tested whether the classifiers trained on colorectal cancer can distinguish the mutational status in GC. As shown in [Figure 9](#), the classifiers trained to discriminate the mutational status of KRAS, PIK3CA, and TP53 genes in the FFPE tissues of colorectal cancer approximately failed to distinguish the mutational status in the FFPE tissues of the TCGA-STAD dataset, with AUCs of 0.458, 0.550, and 0.538 for the KRAS, PIK3CA, and TP53 genes, respectively. The results indicate that the tissue morphologic features reflecting the wild-type and

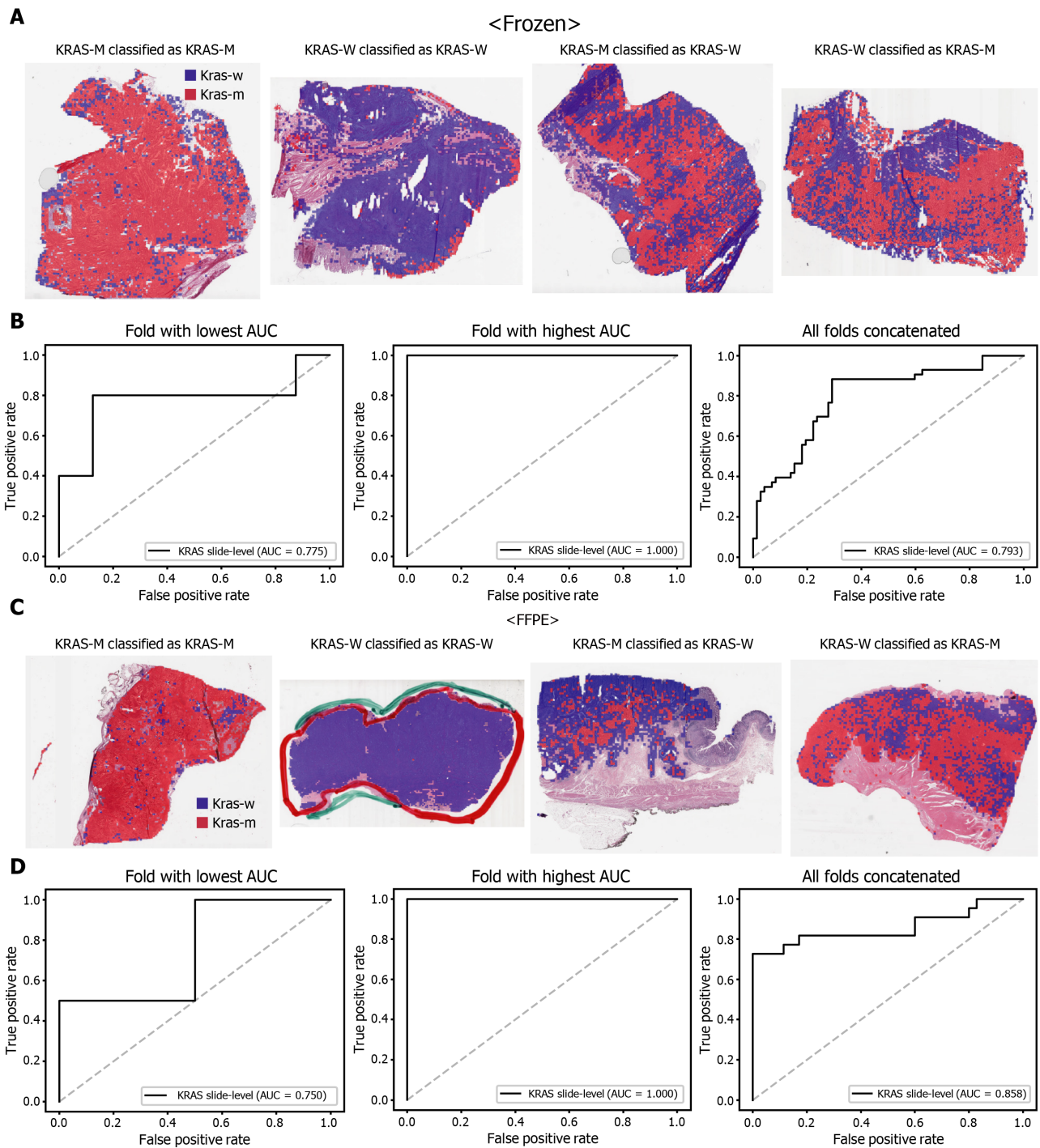


Figure 4 Classification results of KRAS gene in the The Cancer Genome Atlas gastric cancer dataset. The configuration of the figure is the same as in Figure 2. A and B: Upper panels are results for the frozen tissue and lower panels; C and D: Results for the formalin-fixed paraffin-embedded tissues. KRAS-M: KRAS mutated, KRAS-W: KRAS wild-type.

mutated genes are relatively different between cancers originating from different organs.

DISCUSSION

Recently, many drugs targeting specific biological molecules have been introduced to improve the survival of patients with advanced GC[31]. However, patient stratification strategies to maximize the treatment response of these new drugs are not yet well established. Targeted therapies can yield different responses depending on the mutational status of genes in patients with cancer[32]. To overcome this complexity,

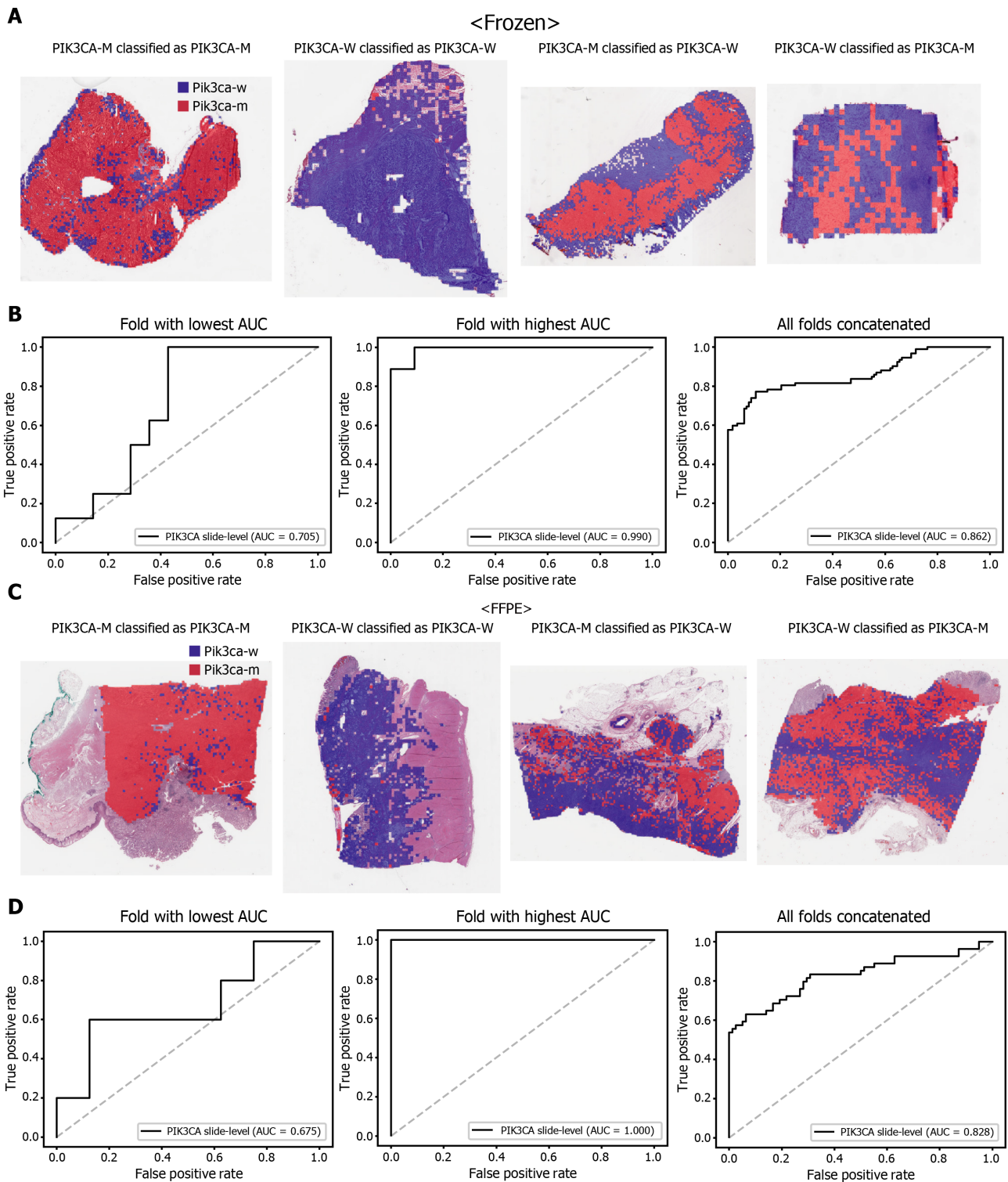


Figure 5 Classification results of PIK3CA gene in the The Cancer Genome Atlas gastric cancer dataset. The configuration of the figure is the same as in Figure 2. A and B: Upper panels are results for the frozen tissue and lower panels; C and D: Results for the formalin-fixed paraffin-embedded tissues. PIK3CA-M: PIK3CA mutated, PIK3CA-W: PIK3CA wild-type.

clinical trials for new drugs have begun to adopt the umbrella platform strategy, which assigns treatment arms based on the mutational status of cancer patients[6,33]. Therefore, data regarding the mutational status of cancer patients is essential for patient stratification in modern-day medicine. However, molecular tests to detect gene mutations are still not affordable for all cancer patients. If cost- and time-effective alternative methods for mutation detection can be introduced, it will promote prospective clinical trials and retrospective studies to correlate the treatment response with the mutational profiles of cancer patients, which can be retrospectively obtained from clinical data and stored tissue samples. Therefore, the new cost- and time-

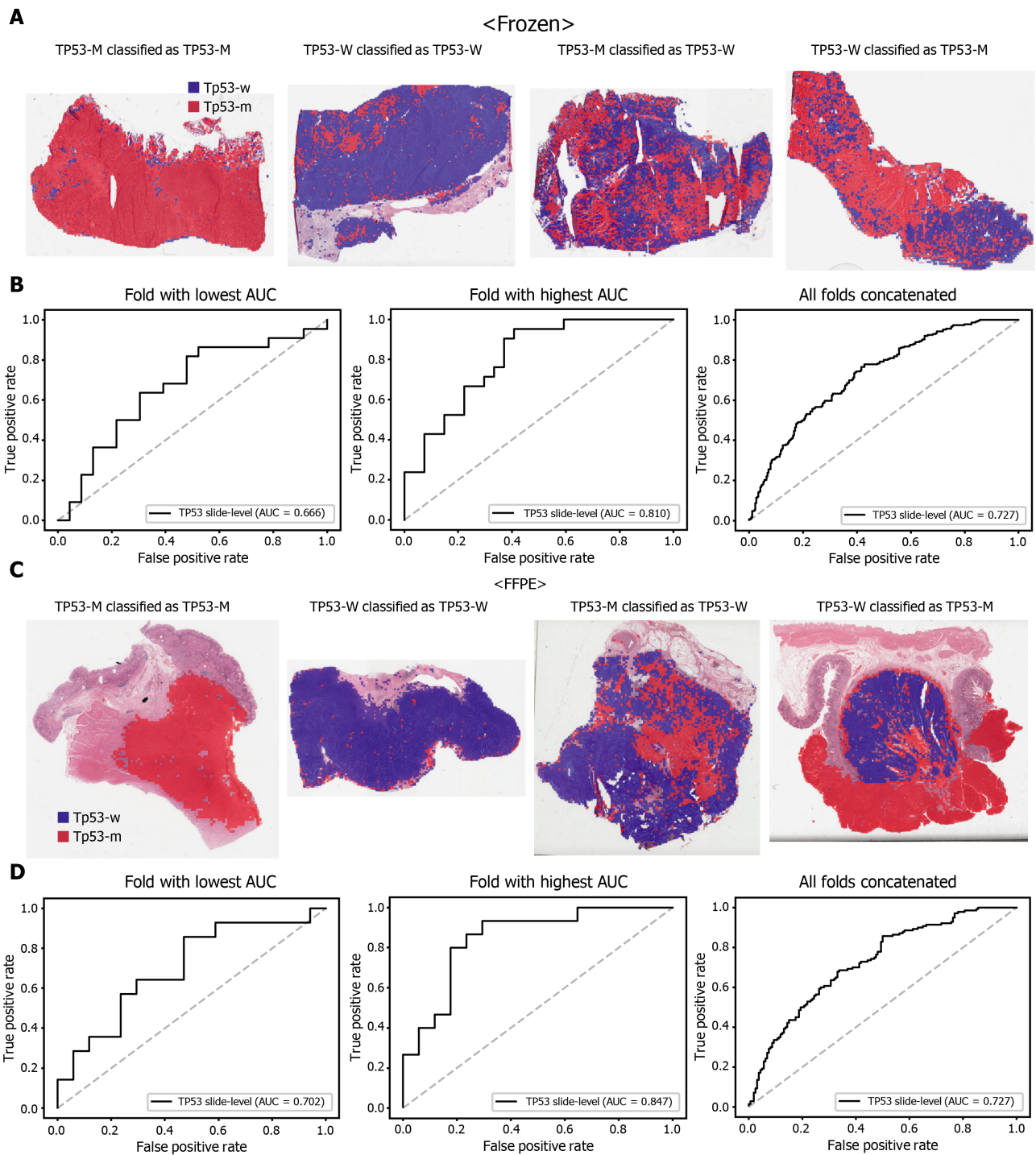
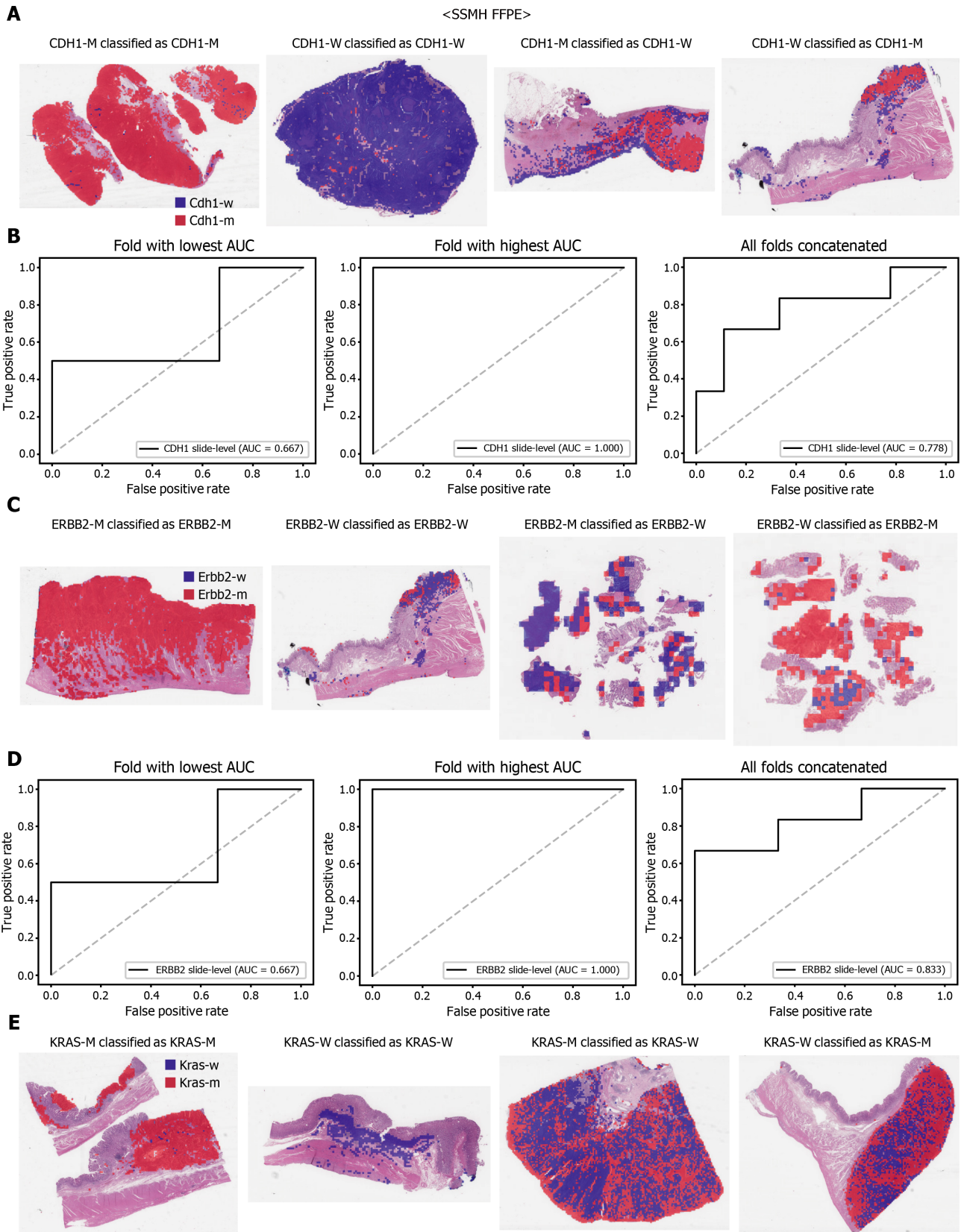


Figure 6 Classification results of TP53 gene in the The Cancer Genome Atlas gastric cancer dataset. The configuration of the figure is the same as in Figure 2. A and B: Upper panels are results for the frozen tissue and lower panels; C and D: Results for the formalin-fixed paraffin-embedded tissues. TP53-M: TP53 mutated, TP53-W: TP53 wild-type.

effective methods will help to establish molecular stratification of cancer patients that can be used to determine effective treatment and improve clinical outcomes[34].

Cancer tissue slides are made and stored for most cancer patients. As a result, DL-based mutation prediction from the tissue slides can be a good candidate for alternative methods. It has been well recognized that the molecular alterations are manifested as morphologic changes in tissue architecture[35]. For example, some morphological features in GC tissues have been associated with specific mutations, including CDH1 and KRAS genes[36,37]. Although it is impractical to quantitatively assess these features for the detection of mutations by visual inspection, DL can learn and distinguish subtle discriminative features for mutation detection in various cancer tissues[7-11]. This study demonstrated the feasibility of DL-based prediction of



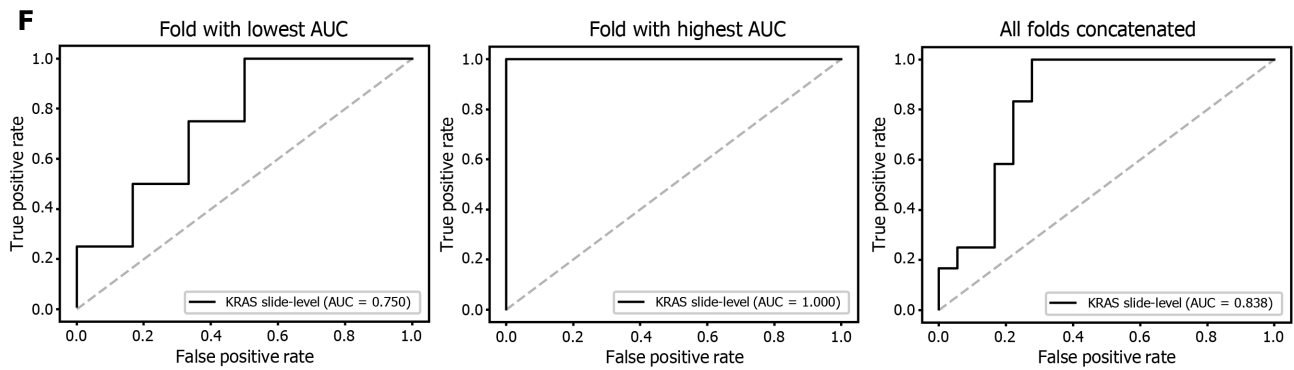


Figure 7 The classifiers trained with both The Cancer Genome Atlas and SSMH data were used to predict the mutation of CDH1 (A and B), ERBB2 (C and D), and KRAS (E and F) genes. Representative binary heatmaps of the whole slide images (WSIs) correctly classified as mutation, correctly classified as wild-type, falsely classified as wild-type, and falsely classified as mutation were presented. Receiver operating characteristic curves for the folds with the lowest and highest area under the curve and the concatenated ten folds were also presented for each gene. CDH1-M: CDH1 mutated, CDH1-W: CDH1 wild-type, ERBB2-M: ERBB2 mutated, ERBB2-W: ERBB2 wild-type, KRAS-M: KRAS mutated, KRAS-W: KRAS wild-type.

mutations in CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes, which are prevalent in both the TCGA and SSMH GC datasets, from tissue slide images of GC. Other studies have also shown that mutations in these genes are frequently observed in GC[3,5]. Furthermore, many studies have attempted to evaluate the prognostic value of these mutations[3,5,38]. However, the clinical relevance of these mutations for prognosis and treatment response has not been completely determined because the studies often presented discordant results. Various factors, including the relatively low incidence of mutation, small study size, and ethnicity of the studied groups, may have contributed to the inconsistent study results. Although it is still unclear how specific mutations are involved in the prognosis and treatment response in GC patients, further studies for the fine molecular stratification of patients based on mutational status are ongoing[6]. DL-based mutation prediction from the tissue slides could provide valuable tools to support these efforts because the mutational status can be promptly obtained with minimal cost from the existing H and E-stained tissue slides.

Furthermore, DL-based classifiers can provide important information for the study of tumor heterogeneity[39]. The heatmaps of classification results overlaid on the tissue images in figures showed that mutated and wild-type regions are aggregated into separated regions. For example, the rightmost tissues in Figure 6C showed clear demarcation between TP53-mutated and wild-type regions. These results indicated that a tumor tissue can contain molecularly heterogeneous regions which can be easily visualized with the help of DL-based classifiers. The clear demarcation of molecularly heterogeneous regions in a tissue slide is an important advantage of a DL-based system and it can help the studies for the understanding of the prognostic and therapeutic values of the tumor heterogeneity without the application of the very expensive molecular tests such as multi-point single-cell sequencing.

However, further studies are needed to build practical DL classifiers for mutation prediction. Our data showed that the performance was still unsatisfactory for verifying mutational status. The AUCs ranged from 0.661 to 0.862 for the TCGA dataset. The frequency of mutation in GC TCGA dataset was relatively low. The average mutation rate for CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes was 8.28%. In our previous study for the mutation prediction in colorectal cancer TCGA dataset, the average mutation rate for APC, KRAS, PIK3CA, SMAD4, and TP53 genes was 39.18%[11]. Furthermore, the classifiers showed limited generalizability to the external dataset. Because DL critically depends on data for learning prominent features, it is generally recommended to build a large multinational dataset[1,2]. Therefore, to test whether the expanded dataset can improve the performance of the classifiers, new classifiers were trained using mixed data from the TCGA and SSMH datasets. As a result, the AUCs generally increased with the larger multinational datasets. These results suggest that we could improve the performance of DL-based mutation classifiers if a large multi-national and multi-institutional dataset can be built. One exception was the PIK3CA gene, which showed worse performance for the TCGA FFPE slides by a classifier trained with the mixed dataset. Although the reason for the decreased performance is unclear, we speculate that there are some different tissue features for the wild-type and mutated PIK3CA gene between the two datasets due to different ethnicities, which could negatively affect the feature learning process for the

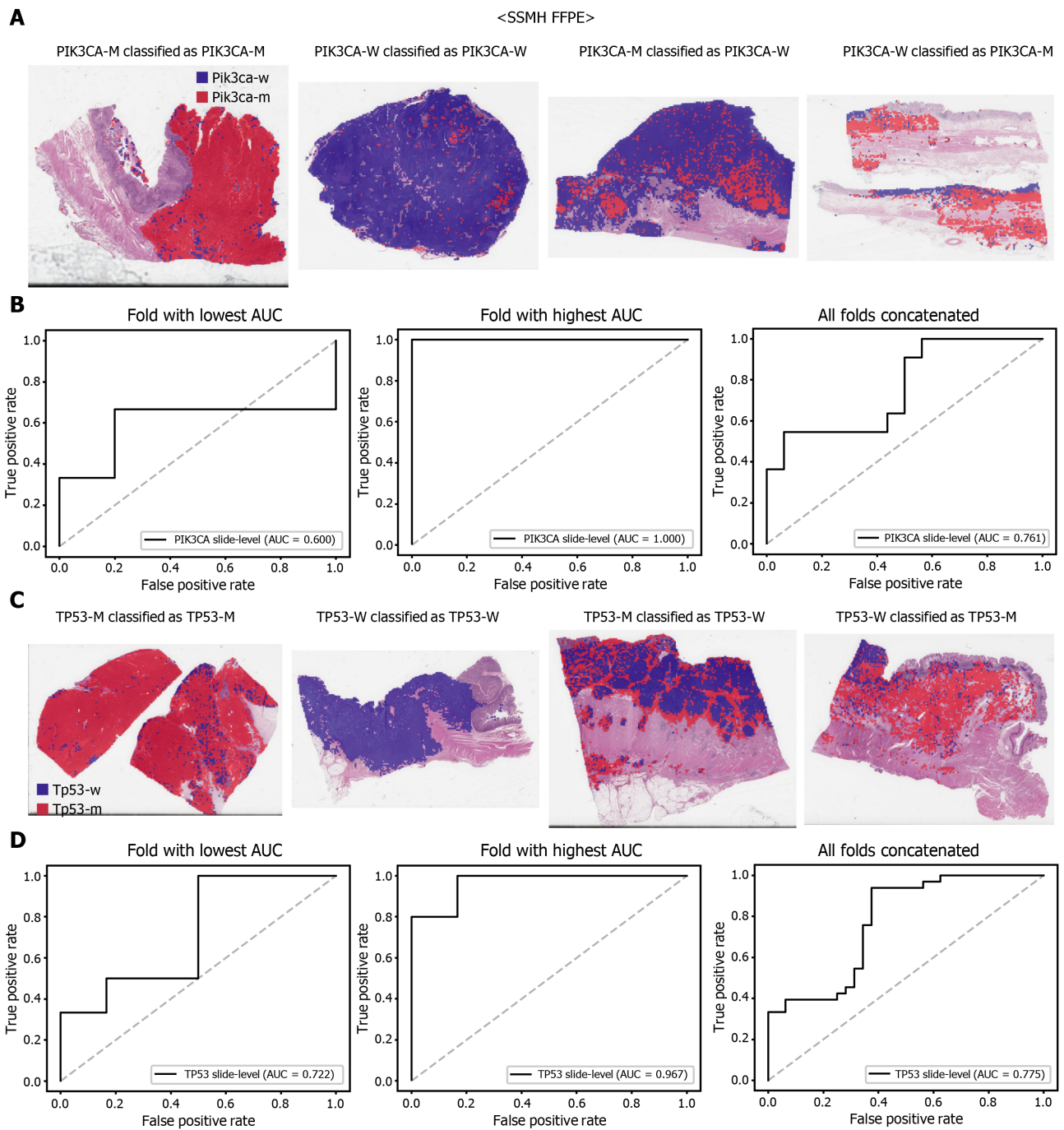


Figure 8 Mutation prediction of PIK3CA (A and B) and TP53 (C and D) genes for the SSMH gastric cancer tissue slides by the classifiers trained with both The Cancer Genome Atlas and SSMH data. The configuration of the figure is the same as in figure 7. PIK3CA-M: PIK3CA mutated, PIK3CA-W: PIK3CA wild-type, TP53-M: TP53 mutated, TP53-W: TP53 wild-type.

TCGA dataset. In addition, the numbers of patients with PIK3CA mutations were different; 64 and 11 for the TCGA and SSMH datasets, respectively. The different numbers of patients also hamper proper feature learning for the mixed dataset because data imbalance usually negatively affects the learning process. Furthermore, the studied tissues carry many additional mutations other than CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes. Because every tissue presented a different combination of mutations, the confounding effect of a mixture of different mutations on the tissue morphology would hamper the effective learning of features for the selected mutation. This factor also necessitates larger tissue datasets for proper learning of morphological features of specific mutations, irrespective of coexisting mutations. In our opinion, the datasets are still immature for building a prominent classifier for mutation prediction. Therefore, efforts to establish a larger tissue dataset with a mutation profile will help to understand the potential of DL-based mutation prediction systems. Recently, many

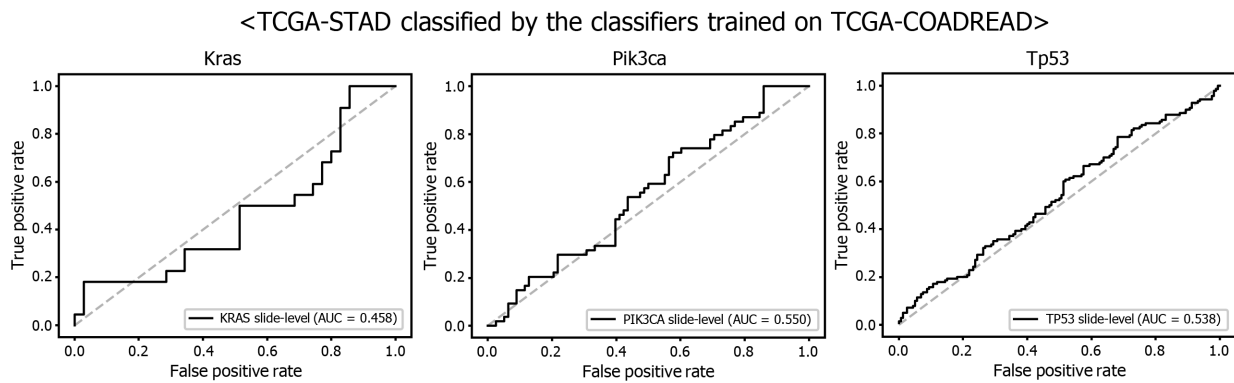


Figure 9 Mutation prediction of KRAS, PIK3CA, and TP53 genes for the The Cancer Genome Atlas gastric cancer tissue slides by the classifiers trained with The Cancer Genome Atlas colorectal cancer tissues. Receiver operating characteristic curves of the classification results for each gene were presented.

countries have started to build nationwide datasets of pathologic tissue WSIs with genomic information. Therefore, we expect that the performance of DL-based mutation prediction can be greatly improved.

Although we argued for the potential of DL-based mutation classifiers, there are important barriers to the adoption of DL-based assistant systems. First, the ‘black box’ nature of DL limits the interpretability of DL models and remains a significant barrier in their validation and adoption in clinics[18,40,41]. We could not trust a decision made by a DL model before we could clearly understand the basis of the decision. Therefore, a method for visualizing the features that determine the behavior of a DL model should be developed. Another barrier is the need for an individual DL system for an individual task. As described, separate systems should be built for tasks such as the classification of normal/tumor tissues for frozen and FFPE tissues. The classifier for each mutation should also be built separately. Furthermore, as shown in Figure 9, there was no compatibility between the different cancer types for the classification of genetic mutations. Therefore, many classifiers should be built to achieve optimal performance. It requires time to build many necessary classifiers to renovate current pathology workflows.

CONCLUSION

Despite these limitations, DL has enormous potential for innovative medical practice. It can help capture important information by learning features automatically from the data that are waiting to be explored in the vast database of modern hospital information systems. This information will be used to determine the best medical practice and improve patient outcomes. The tissue slides of cancer patients contain important information on the prognosis of patients[42]; therefore, DL-based analysis of tissue slides has enormous potential for fine patient stratification in the era of precision medicine. Furthermore, its cost- and time-effective nature could help save the medical cost and decision time for patient care.

ARTICLE HIGHLIGHTS

Research background

Studies correlating specific genetic mutations and treatment response are ongoing to establish an effective treatment strategy for gastric cancer (GC). With the increased digitization of pathologic tissue slides, deep learning (DL) can be a cost- and time-effective method to analyze the mutational status directly from the hematoxylin and eosin (H and E)-stained tissue whole slide images (WSIs).

Research motivation

Recent studies suggested that mutational status can be predicted directly from the H and E-stained WSIs with DL-based methods. Motivated by these studies, we invest-

igated the feasibility of DL-based mutation prediction for the frequently occurring mutations from H and E-stained WSIs of GC tissues.

Research objectives

To predict the mutational status of CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes from the H and E-stained WSIs of GC tissues with DL-based methods.

Research methods

DL-based classifiers for the CDH1, ERBB2, KRAS, PIK3CA, and TP53 mutations were trained for the The Cancer Genome Atlas (TCGA) datasets. Then, the classifiers were validated with our own dataset. Finally, TCGA and our own dataset were combined to train a new classifier to test the effect of extended data on the performance of the classifiers.

Research results

The area under the curve (AUC) for receiver operating characteristic (ROC) curves were between 0.727 and 0.862 for the TCGA frozen WSIs and between 0.661 and 0.858 for the TCGA formalin-fixed paraffin-embedded WSIs. Furthermore, the results could be improved with the classifiers trained with both TCGA and our own dataset.

Research conclusions

This study demonstrated that mutational status could be predicted directly from the H and E-stained WSIs of GC tissues with DL-based methods. The performance of the classifiers could be improved if more data can be used to train the classifiers.

Research perspectives

Current molecular tests for the mutational status are not feasible for all cancer patients because of technical barriers and high costs. Although there is still room for much improvement, the DL-based method can be a reasonable alternative for molecular tests. It could help to stratify patients based on their mutational status for retrospective studies or prospective clinical trials with very low cost. Furthermore, it could support the decision-making process for the management of patients with GCs.

REFERENCES

- 1 Serag A, Ion-Margineanu A, Qureshi H, McMillan R, Saint Martin MJ, Diamond J, O'Reilly P, Hamilton P. Translational AI and Deep Learning in Diagnostic Pathology. *Front Med (Lausanne)* 2019; **6**: 185 [PMID: 31632973 DOI: 10.3389/fmed.2019.00185]
- 2 Djuric U, Zadeh G, Aldape K, Diamandis P. Precision histology: how deep learning is poised to revitalize histomorphology for personalized cancer care. *NPJ Precis Oncol* 2017; **1**: 22 [PMID: 29872706 DOI: 10.1038/s41698-017-0022-1]
- 3 Ito T, Matoba R, Maekawa H, Sakurada M, Kushida T, Orita H, Wada R, Sato K. Detection of gene mutations in gastric cancer tissues using a commercial sequencing panel. *Mol Clin Oncol* 2019; **11**: 455-460 [PMID: 31620276 DOI: 10.3892/mco.2019.1926]
- 4 Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; **68**: 394-424 [PMID: 30207593 DOI: 10.3322/caac.21492]
- 5 Nemtsova MV, Kalinkin AI, Kuznetsova EB, Bure IV, Alekseeva EA, Bykov II, Khorobrykh TV, Mikhaylenko DS, Tanas AS, Kutsev SI, Zaletaev DV, Strelnikov VV. Clinical relevance of somatic mutations in main driver genes detected in gastric cancer patients by next-generation DNA sequencing. *Sci Rep* 2020; **10**: 504 [PMID: 31949278 DOI: 10.1038/s41598-020-57544-3]
- 6 Lee J, Kim ST, Kim K, Lee H, Kozarewa I, Mortimer PGS, Odegaard JI, Harrington EA, Lee J, Lee T, Oh SY, Kang JH, Kim JH, Kim Y, Ji JH, Kim YS, Lee KE, Kim J, Sohn TS, An JY, Choi MG, Lee JH, Bae JM, Kim S, Kim JJ, Min YW, Min BH, Kim NKD, Luke S, Kim YH, Hong JY, Park SH, Park JO, Park YS, Lim HY, Talasaz A, Hollingsworth SJ, Kim KM, Kang WK. Tumor Genomic Profiling Guides Patients with Metastatic Gastric Cancer to Targeted Treatment: The VIKTORY Umbrella Trial. *Cancer Discov* 2019; **9**: 1388-1405 [PMID: 31315834 DOI: 10.1158/2159-8290.CD-19-0442]
- 7 Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, Moreira AL, Razavian N, Tsirigos A. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med* 2018; **24**: 1559-1567 [PMID: 30224757 DOI: 10.1038/s41591-018-0177-5]
- 8 Schaumberg AJ, Rubin MA, Fuchs TJ. H and E-stained Whole Slide Image Deep Learning Predicts SPOP Mutation State in Prostate Cancer. *bioRxiv* 2018 [DOI: 10.1101/064279]
- 9 Kather JN, Heij LR, Grabsch HI, Loeffler C, Echle A, Muti HS, Krause J, Niehues JM, Sommer

- KAJ, Bankhead P, Kooreman LFS, Schulte JJ, Cipriani NA, Buelow RD, Boor P, Ortiz-Brüchle NN, Hanby AM, Speirs V, Kochanny S, Patnaik A, Srisuwananukorn A, Brenner H, Hoffmeister M, van den Brandt PA, Jäger D, Trautwein C, Pearson AT, Luedde T. Pan-cancer image-based detection of clinically actionable genetic alterations. *Nat Cancer* 2020; **1**: 789-799 [PMID: [33763651](#) DOI: [10.1038/s43018-020-0087-6](#)]
- 10 **Chen M**, Zhang B, Topatana W, Cao J, Zhu H, Juengpanich S, Mao Q, Yu H, Cai X. Classification and mutation prediction based on histopathology H and E images in liver cancer using deep learning. *NPJ Precis Oncol* 2020; **4**: 14 [PMID: [32550270](#) DOI: [10.1038/s41698-020-0120-3](#)]
 - 11 **Jang HJ**, Lee A, Kang J, Song IH, Lee SH. Prediction of clinically actionable genetic alterations from colorectal cancer histopathology images using deep learning. *World J Gastroenterol* 2020; **26**: 6207-6223 [DOI: [10.3748/wjg.v26.i40.6207](#)]
 - 12 **Saltz J**, Gupta R, Hou L, Kure T, Singh P, Nguyen V, Samaras D, Shroyer KR, Zhao T, Batiste R, Van Arnam J; Cancer Genome Atlas Research Network, Shmulevich I, Rao AUK, Lazar AJ, Sharma A, Thorsson V. Spatial Organization and Molecular Correlation of Tumor-Infiltrating Lymphocytes Using Deep Learning on Pathology Images. *Cell Rep* 2018; **23**: 181-193.e7 [PMID: [29617659](#) DOI: [10.1016/j.celrep.2018.03.086](#)]
 - 13 **Echle A**, Rindtorff NT, Brinker TJ, Luedde T, Pearson AT, Kather JN. Deep learning in cancer pathology: a new generation of clinical biomarkers. *Br J Cancer* 2021; **124**: 686-696 [PMID: [33204028](#) DOI: [10.1038/s41416-020-01122-x](#)]
 - 14 **Calderaro J**, Kather JN. Artificial intelligence-based pathology for gastrointestinal and hepatobiliary cancers. *Gut* 2021; **70**: 1183-1193 [PMID: [33214163](#) DOI: [10.1136/gutjnl-2020-322880](#)]
 - 15 **Coudray N**, Tsigos A. Deep learning links histology, molecular signatures and prognosis in cancer. *Nature Cancer* 2020; **1**: 755-757 [DOI: [10.1038/s43018-020-0099-2](#)]
 - 16 **Cao R**, Yang F, Ma S, Liu L, Li Y, Wu D, Zhao Y, Wang T, Lu W, Cai W, Zhu H, Guo X, Lu Y, Kuang J, Huan W, Tang W, Huang J, Yao J, Dong Z. A transferrable and interpretable multiple instance learning model for microsatellite instability prediction based on histopathology images. *bioRxiv* 2020 [DOI: [10.1101/2020.02.29.971150](#)]
 - 17 **Kather JN**, Calderaro J. Development of AI-based pathology biomarkers in gastrointestinal and liver cancer. *Nat Rev Gastroenterol Hepatol* 2020; **17**: 591-592 [PMID: [32620817](#) DOI: [10.1038/s41575-020-0343-3](#)]
 - 18 **Mobadersany P**, Yousefi S, Amgad M, Gutman DA, Barnholtz-Sloan JS, Velázquez Vega JE, Brat DJ, Cooper LAD. Predicting cancer outcomes from histology and genomics using convolutional networks. *Proc Natl Acad Sci U S A* 2018; **115**: E2970-E2979 [PMID: [29531073](#) DOI: [10.1073/pnas.1717139115](#)]
 - 19 **Yu KH**, Berry GJ, Rubin DL, Ré C, Altman RB, Snyder M. Association of Omics Features with Histopathology Patterns in Lung Adenocarcinoma. *Cell Syst* 2017; **5**: 620-627.e3 [PMID: [29153840](#) DOI: [10.1016/j.cels.2017.10.014](#)]
 - 20 **LeCun Y**, Bengio Y, Hinton G. Deep learning. *Nature* 2015; **521**: 436-444 [PMID: [26017442](#) DOI: [10.1038/nature14539](#)]
 - 21 **Nam S**, Chong Y, Jung CK, Kwak TY, Lee JY, Park J, Rho MJ, Go H. Introduction to digital pathology and computer-aided pathology. *J Pathol Transl Med* 2020; **54**: 125-134 [PMID: [32045965](#) DOI: [10.4132/jptm.2019.12.31](#)]
 - 22 **Cooper LA**, Demicco EG, Saltz JH, Powell RT, Rao A, Lazar AJ. PanCancer insights from The Cancer Genome Atlas: the pathologist's perspective. *J Pathol* 2018; **244**: 512-524 [PMID: [29288495](#) DOI: [10.1002/path.5028](#)]
 - 23 **Cho KO**, Lee SH, Jang HJ. Feasibility of fully automated classification of whole slide images based on deep learning. *Korean J Physiol Pharmacol* 2020; **24**: 89-99 [PMID: [31908578](#) DOI: [10.4196/kjpp.2020.24.1.89](#)]
 - 24 **Pon JR**, Marra MA. Driver and passenger mutations in cancer. *Annu Rev Pathol* 2015; **10**: 25-50 [PMID: [25340638](#) DOI: [10.1146/annurev-pathol-012414-040312](#)]
 - 25 **Bailey MH**, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, Colaprico A, Wendl MC, Kim J, Reardon B, Ng PK, Jeong KJ, Cao S, Wang Z, Gao J, Gao Q, Wang F, Liu EM, Mularoni L, Rubio-Perez C, Nagarajan N, Cortés-Ciriano I, Zhou DC, Liang WW, Hess JM, Yellapantula VD, Tamborero D, Gonzalez-Perez A, Suphavilai C, Ko JY, Khurana E, Park PJ, Van Allen EM, Liang H; MC3 Working Group; Cancer Genome Atlas Research Network, Lawrence MS, Godzik A, Lopez-Bigas N, Stuart J, Wheeler D, Getz G, Chen K, Lazar AJ, Mills GB, Karchin R, Ding L. Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* 2018; **173**: 371-385.e18 [PMID: [29625053](#) DOI: [10.1016/j.cell.2018.02.060](#)]
 - 26 **Jang HJ**, Song IH, Lee SH. Generalizability of Deep Learning System for the Pathologic Diagnosis of Various Cancers. *Appl Sci-Basel* 2021; **11**: 808 [DOI: [10.3390/app11020808](#)]
 - 27 **Salvi M**, Acharya UR, Molinari F, Meiburger KM. The impact of pre- and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis. *Comput Biol Med* 2021; **128**: 104129 [PMID: [33254082](#) DOI: [10.1016/j.combiomed.2020.104129](#)]
 - 28 **Swiderska-Chadaj Z**, de Bel T, Blanchet L, Baidoshvili A, Vossen D, van der Laak J, Litjens G. Impact of rescanning and normalization on convolutional neural network performance in multi-center, whole-slide classification of prostate cancer. *Sci Rep* 2020; **10**: 14398 [PMID: [32873856](#) DOI: [10.1038/s41598-020-71420-0](#)]
 - 29 **Jang H-J**, Song IH, Lee SH. Generalizability of Deep Learning System for the Pathologic Diagnosis

- of Various Cancers. *Applied Sciences* 2021; **11**: 808
- 30 **Venkatraman ES.** A permutation test to compare receiver operating characteristic curves. *Biometrics* 2000; **56**: 1134-1138 [PMID: 11129471 DOI: 10.1111/j.0006-341x.2000.01134.x]
 - 31 **Hashimoto T,** Kurokawa Y, Mori M, Doki Y. Update on the Treatment of Gastric Cancer. *JMA J* 2018; **1**: 40-49 [PMID: 33748521 DOI: 10.31662/jmaj.2018-0006]
 - 32 **Wang M,** Yu L, Wei X, Wei Y. Role of tumor gene mutations in treatment response to immune checkpoint blockades. *Precis Clin Med* 2019; **2**: 100-109 [DOI: 10.1093/pemedi/pbz006]
 - 33 **Freidlin B,** Korn EL. Biomarker enrichment strategies: matching trial design to biomarker credentials. *Nat Rev Clin Oncol* 2014; **11**: 81-90 [PMID: 24281059 DOI: 10.1038/nrclinonc.2013.218]
 - 34 **Hamilton PW,** Bankhead P, Wang Y, Hutchinson R, Kieran D, McArt DG, James J, Salto-Tellez M. Digital pathology and image analysis in tissue biomarker research. *Methods* 2014; **70**: 59-73 [PMID: 25034370 DOI: 10.1016/j.ymeth.2014.06.015]
 - 35 **Madabhushi A,** Lee G. Image analysis and machine learning in digital pathology: Challenges and opportunities. *Med Image Anal* 2016; **33**: 170-175 [PMID: 27423409 DOI: 10.1016/j.media.2016.06.037]
 - 36 **Luo W,** Fedda F, Lynch P, Tan D. CDH1 Gene and Hereditary Diffuse Gastric Cancer Syndrome: Molecular and Histological Alterations and Implications for Diagnosis And Treatment. *Front Pharmacol* 2018; **9**: 1421 [PMID: 30568591 DOI: 10.3389/fphar.2018.01421]
 - 37 **Hewitt LC,** Saito Y, Wang T, Matsuda Y, Oosting J, Silva ANS, Slaney HL, Melotte V, Hutchins G, Tan P, Yoshikawa T, Arai T, Grabsch HI. KRAS status is related to histological phenotype in gastric cancer: results from a large multicentre study. *Gastric Cancer* 2019; **22**: 1193-1203 [PMID: 31111275 DOI: 10.1007/s10120-019-00972-6]
 - 38 **Takahashi N,** Yamada Y, Taniguchi H, Fukahori M, Sasaki Y, Shoji H, Honma Y, Iwasa S, Takashima A, Kato K, Hamaguchi T, Shimada Y. Clinicopathological features and prognostic roles of KRAS, BRAF, PIK3CA and NRAS mutations in advanced gastric cancer. *BMC Res Notes* 2014; **7**: 271 [PMID: 24774510 DOI: 10.1186/1756-0500-7-271]
 - 39 **Bera K,** Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol* 2019; **16**: 703-715 [PMID: 31399699 DOI: 10.1038/s41571-019-0252-y]
 - 40 **Colling R,** Pitman H, Oien K, Rajpoot N, Macklin P; CM-Path AI in Histopathology Working Group, Snead D, Sackville T, Verrill C. Artificial intelligence in digital pathology: a roadmap to routine use in clinical practice. *J Pathol* 2019; **249**: 143-150 [PMID: 31144302 DOI: 10.1002/path.5310]
 - 41 **Bhatt C,** Kumar I, Vijayakumar V, Singh U, Kumar A. The state of the art of deep learning models in medical science and their challenges. *Multimedia Syst* 2020; **12** [DOI: 10.1007/s00530-020-00694-1]
 - 42 **Fu Y,** Jung AW, Torne RV, Gonzalez S, Vöhringer H, Jimenez-Linan M, Moore L, Gerstung M. Pan-cancer computational histopathology reveals mutations, tumor composition and prognosis. *bioRxiv* 2019 [DOI: 10.1101/813543]



Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: <https://www.f6publishing.com/helpdesk>
<https://www.wjgnet.com>

