Prompting Whole Slide Image Based Genetic Biomarker Prediction

Ling Zhang[1], Boxiang Yun[1], Xingran Xie[1], Qingli Li[1], Xinxing Li[2], and Yan Wang[1(✉)]

[1] Shanghai Key Laboratory of Multidimensional Information Processing,
East China Normal University, Shanghai, China
[2] Department of Gastrointestinal Surgery,
Tongji Hospital Medical College of Tongji University
51265904017@stu.ecnu.edu.cn, 52265904012@stu.ecnu.edu.cn,
51215904112@stu.ecnu.edu.cn, qlli@cs.ecnu.edu.cn, ahtxxxx2015@163.com,
ywang@cee.ecnu.edu.cn

Abstract. Prediction of genetic biomarkers, *e.g.*, microsatellite instability and BRAF in colorectal cancer is crucial for clinical decision making. In this paper, we propose a whole slide image (WSI) based genetic biomarker prediction method via prompting techniques. Our work aims at addressing the following challenges: (1) extracting foreground instances related to genetic biomarkers from gigapixel WSIs, and (2) the interaction among the fine-grained pathological components in WSIs. Specifically, we leverage large language models to generate medical prompts that serve as prior knowledge in extracting instances associated with genetic biomarkers. We adopt a coarse-to-fine approach to mine biomarker information within the tumor microenvironment. This involves extracting instances related to genetic biomarkers using coarse medical prior knowledge, grouping pathology instances into fine-grained pathological components and mining their interactions. Experimental results on two colorectal cancer datasets show the superiority of our method, achieving 91.49% in AUC for MSI classification. The analysis further shows the clinical interpretability of our method. Code is publicly available at https://github.com/DeepMed-Lab-ECNU/PromptBio.

Keywords: Genetic biomarker prediction · Whole slide image · Text prompt.

1 Introduction

Colorectal cancer (CRC) is the second leading cause of cancer-related deaths worldwide [25]. The evaluation of genetic biomarkers is crucial for CRC diagnosis and prognosis, such as microsatellite instability (MSI) and mutations in the BRAF, since these biomarkers can identify CRC patients with different treatment response and prognosis [13, 6, 7]. MSI has been officially adopted as global classification standards for the diagnosis and treatment of CRC in the National Comprehensive Cancer Network guidelines [1]. However, the method of

testing genetic biomarkers is time-consuming and expensive, such as immuno-histochemistry, polymerase chain reaction and next-generation-sequencing [8]. Due to diagnostic needs, the pathology whole slide images (WSIs) stained with hematoxylin and eosin (H&E) are routinely available for all CRC patients and previous works [12, 24] suggested that genetic alterations are expressed in digital pathology WSIs. Therefore, automatically predicting genetic biomakers from WSIs is feasible and highly demanded in clinical practice.

Genetic biomarker prediction based on gigapixel WSIs can be cast as a multiple instance learning (MIL) problem, where WSIs are bags and patches cropped from WSIs are instances, aiming to predict a bag label. However, learning genetic biomarker representations based on bags is challenging, because it's hard to extract foreground instances related to genetic biomarkers from a WSI. Moreover, the tumor microenvironment is a complex network. Various pathological component interactions exist between immune cells and non-cellular components such as the extracellular matrix, exosomes and interleukins [3], which may be relevant to genetic biomarkers. Therefore, how to extract foreground instances related to genetic biomarkers and mine the interaction among pathological components in the the tumor microenvironment is crucial for genetic biomarker prediction.

Specifically, to our knowledge, most WSI based genetic biomarker prediction methods, such as MOMA [27] and MSIntuit [22], indiscriminately input all instances into the network, which bring a large number of irrelevant instances. Many other WSI classification methods can be extended to predict genetic biomarkers. Among these, some methods have implemented learnable or non-learnable instance selection. For example, VIBFT [16] introduces a variational information bottleneck to find the minimal sufficient statistics of WSIs, and MILBooster [21] proposes a bag filter that increases the positive instance ratio of positive bags using K-means algorithm. But these selection methods lack interpretable medical justification. Even with a reduced number of instances, they may still lose key tissue regions relevant to genetic biomarker. Moreover, most methods [11, 14] treat pathology instances as independent entities and overlook their interaction. Recent studies [23, 15] have used Vision Transformer frameworks to model the relationships between instances, but they neglect the interaction among different pathological components. HGT [9] attempts to explore the interaction among different pathological components in WSIs, but their construction of pathological components relies on superpixel rather than actual pathological component distribution. TOP [20] models different prototypes guided by pathology language prior knowledge and the pathology language prior knowledge is flexible and adaptable, but it aggregates instance features into a bag feature through a simple weighting method without exploring the interaction among different prototypes.

To this end, we propose a task-specific framework dubbed Prompting whole slide image-based genetic Biomarker prediction (PromptBio), consisting of three modules, which leverages the capabilities of the large language model (LLM) to generate pathology text prompts, and uses biomarker-associated medical prior knowledge to guide the pathological component extraction and feature fusion.
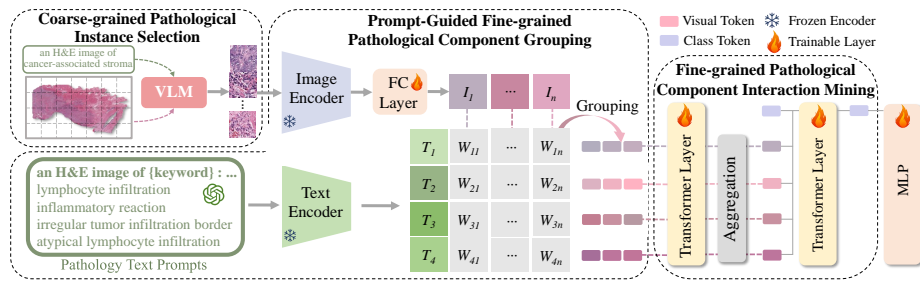
Fig. 1. Illustration of PromptBio model. The overall framework consists of three parts: 1) coarse-grained pathological instance selection module, 2) prompt-guided fine-grained pathological component grouping module, 3) fine-grained pathological component interaction mining module. Given a dataset $\mathcal{D}$ consisting of N pathology WSIs, our PromptBio first performs coarse-grained pathological instance selection to extract the instances belonging to cancer-associated stroma. Then our PromptBio performs fine-grained pathological component grouping on extracted instances. The grouping is guided by pathology text prompts of cancer-associated stroma with MSI. Finally, our PromptBio performs fine-grained pathological component interaction mining.

Pathological components of diverse microenvironments range from fine-grained (*e.g.*, lymphocyte infiltration, inflammatory reaction, *etc.*) to coarse-grained (*e.g.*, cancer-associated stroma, epithelium tissue, *etc.*). Among various coarse-grained pathological components, cancer-associated stroma is proven as the **key contributor** to the tumor microenvironment [5, 4] and is highly associated with consensus molecular subtypes [2]. We regard cancer-associated stroma as the tissue of our interest and use it to discard irrelevant information in a WSI. Thus, we first propose a coarse-grained pathological instance selection module to extract foreground instances belonging to cancer-associated stroma from a gigapixel WSI. This process is guided by coarse-grained medical prompts. Next, we focus on learning the distribution and interaction among fine-grained pathological components in cancer-associated stroma, to extract comprehensive and in-depth biomarker-associated features. Two modules, *i.e.*, fine-grained pathological component grouping and interaction mining, are proposed. In the grouping module, we prompt LLM to generate biomarker-associated descriptions for fine-grained pathological components, which identify multiple semantic types from the WSIs. We design a grouping strategy to aggregate instances of each type and gather their distribution. Prompting techniques ensure the accuracy of the tumor microenvironment modeling. Furthermore, the hierarchical contextual interaction of intra- and inter-pathological components are extracted and fused via our interaction mining module, enhancing the interpretability of the tumor microenvironment modeling. Extensive experiments show our method outperforms all state-of-the-arts by a large margin, with over 5% improvement in AUC for CRC MSI prediction on TCGA dataset. Additionally, our method shows clinical interpretability, providing potential for future clinical applications.

2   Method

Given a dataset $\mathcal{D}$ consisting of N pathology WSIs, pathological components of diverse microenvironments in WSIs range from fine-grained (*e.g.*, lymphocyte infiltration, *etc.*) to coarse-grained (cancer-associated stroma, *etc.*). Prompting Whole Slice Image Based Genetic Biomarker Prediction (PromptBio) first performs coarse-grained pathological instance selection based on medical prior knowledge to extract the instances belonging to cancer-associated stroma. Considering that the tumor microenvironment is a complex network of various pathological components, PromptBio performs fine-grained pathological component grouping on extracted instances. The grouping is guided by pathology text prompts of cancer-associated stroma with MSI. Finally, PromptBio performs fine-grained pathological component interaction mining. As shown in Fig.1, the overall framework consists of three modules: 1) coarse-grained pathological instance selection (Sec. 2.1), 2) prompt-guided fine-grained pathological component grouping (Sec. 2.2) and 3) fine-grained pathological component interaction mining (Sec. 2.3). Next, we will introduce each part in detail.

2.1   Coarse-grained Pathological Instance Selection

To discard abundant irrelevant information in a WSI, we select the instances related to genetic biomarkers based on medical prior knowledge. Among various coarse-grained pathological components, cancer-associated stroma is proven as the key contributor to the tumor microenvironment [5, 4, 2]. We regard cancer-associated stroma as the tissue of our interest and use it to discard abundant irrelevant information in a WSI. We utilize PLIP [10] for coarse-grained pathological instance classification of CRC WSIs. PLIP is a large vision-language model (VLM) for pathology image analysis that allows zero-shot classification of tissues on WSIs, where the text encoder yields text embeddings of medical descriptions. An H&E-stained WSI contains a large number of irrelevant background regions, and we use the image preprocessing algorithm in CLAM [18] to detect the tissue regions in the WSI and crop the tissue regions into non-overlapping 512×512 pixel-sized patches. Then we input the text "an H&E image of {keywords}" into the frozen text encoder of PLIP and input the instances of the WSIs into the image encoder of PLIP. The keywords correspond to nine tissue classes [10] including cancer-associated stroma. Finally, we extract the instances belonging to cancer-associated stroma.

2.2   Prompt-Guided Fine-grained Pathological Component Grouping

To learn the distribution of fine-grained pathological components in cancer-associated stroma and extract more comprehensive and in-depth biomarker-associated features, we group selected instances into fine-grained pathological components based on pathology text prompts. We use the prompt template "If a colorectal cancer patient is in an MSI (Microsatellite Instability) status, what

are the characteristics of the cancer-associated stroma areas in the WSIs may exhibit?" for guiding GPT-4 to generate the pathology descriptions. The generated descriptions include four types: lymphocyte infiltration, inflammatory reaction, irregular tumor infiltration border, and atypical lymphocyte infiltration. Each type has a specific description. We input the four pathology descriptions into the text encoder of PLIP to obtain the text embeddings $\boldsymbol{T} \in \mathbb{R}^{4 \times d}$. After patch selection, we follow the method in IBMIL [17] to extract the instance features and then use a fully connected layer to map the instance features to d dimensions $\boldsymbol{I} \in \mathbb{R}^{n \times d}$, where n is the number of instances selected in each bag. We compute the cosine similarity between the instance features and each text embedding:

$$\boldsymbol{W}_i(\boldsymbol{I}) = \text{sim}(\boldsymbol{I}, \boldsymbol{T}_i^{\top}), i = 1, 2, 3, 4, \tag{1}$$

where $\text{sim}(\cdot\,,\,\cdot)$ refers to cosine similarity, $\boldsymbol{T}_i$ means the ith text embedding of the four components and $\boldsymbol{W}_i(\boldsymbol{I}) \in \mathbb{R}^{n \times 1}$.

Then we select the top-k image features for each pathological component based on similarity ranking, where $k = n \cdot \beta$ and β is the selection ratio. In this way, we obtain four fine-grained pathological components $\boldsymbol{C}_i \in \mathbb{R}^{k \times d}$ and corresponding instance features $\boldsymbol{C}_{ij} \in \mathbb{R}^{1 \times d}$, where $\boldsymbol{C}_i$ means the ith component in the four components with pathology semantics. After grouping, many irrelevant instances are reduced. $\boldsymbol{C}_i$ is computed by:

$$\boldsymbol{C}_i = \Phi\left(\boldsymbol{W}_i\left(\boldsymbol{I}\right), k\right), i = 1, 2, 3, 4, \tag{2}$$

where $\Phi(\boldsymbol{W}_i(\boldsymbol{I}), k)$ refers to the k out of n instances in $\boldsymbol{I}$ that maximize $\boldsymbol{W}_i\left(\boldsymbol{I}\right)$.

2.3 Fine-grained Pathological Component Interaction Mining

To mine the interaction among fine-grained pathological components, we separately input the instance features of each pathological component into the transformer layer to model the interaction of each instance within a pathological component, and separately aggregate the instance features of each pathological component. The feature $\boldsymbol{C}_i^{mean} \in \mathbb{R}^{1 \times d}$ of the ith pathological component is:

$$\boldsymbol{C}_i^{mean} = \frac{1}{k}\left(\sum_j^k \boldsymbol{C}_{ij}\right), i = 1, 2, 3, 4. \tag{3}$$

Then, we concatenate component features $\boldsymbol{C}^{mean} = \text{Concat}(\boldsymbol{C}_1^{mean}, \boldsymbol{C}_2^{mean}, \boldsymbol{C}_3^{mean}, \boldsymbol{C}_4^{mean})$. A class token $CLS \in \mathbb{R}^{1 \times d}$ is concatenated with $\boldsymbol{C}^{mean} \in \mathbb{R}^{4 \times d}$. The concatenated feature is further fed into the next transformer layer to mine the interaction among pathological components. We use a Multi-Layer Perceptron (MLP) head to map the output class token $CLS' \in \mathbb{R}^{1 \times d}$ to the final class predictions $P = \text{softmax}\left(\text{MLP}\left(CLS'\right)\right)$. So far, we design a simple yet effective method to achieve the complex biomarker prediction task.

3   Experiments

3.1   Experimental Setup

Datasets We verify the effectiveness of our method on The Caner Genome Atlas (TCGA) CRC dataset and The Clinical Proteomic Tumor Analysis Consortium (CPTAC) CRC dataset, both containing two CRC subtypes, *i.e.*, Microsatellite Stable (MSS) and Microsatellite Instability (MSI) and TCGA also contains another two CRC subtypes, *i.e.*, BRAF mutation and non-BRAF mutation. We use the image preprocessing algorithm in CLAM [18] to detect the tissue regions in WSIs and crop the tissue regions into non-overlapping 512×512 pixel-sized patches at $20\times$ magnification. Due to significant staining variations in the TCGA dataset, we employ the color normalization proposed by Macenko *et al.* [19]. Among 482 WSIs in TCGA, 420 of them belong to MSS and 62 are MSI, and 429 of them belong to BRAF mutation and 53 are non-BRAF mutation. We randomly split the TCGA dataset into training and testing sets in a 3:1 ratio, with 10% of the training set further split into a validation set. Among 105 WSIs in CPTAC, 81 of them belong to MSS and 24 cases are MSI. We randomly split the CPTAC dataset into training, validation, and test sets in a 3:1:1 ratio.

Implementation Details During training, we evaluate the model on the validation set after every epoch, and save the parameters when it performs the best. Adam is used as our optimizer and the learning rate is 5×10^{-5} with learning rate annealing. Binary Cross Entropy loss is used as our loss function. The maximum number of training epoch is 32. At last, we measure the performance on the test set. We report the area under the curve (AUC) scores. All experiments are conducted on a single NVIDIA GeForce RTX 3090.

3.2   Comparison between PromptBio and Other Methods

We compare PromptBio with the current state-of-the-art deep MIL models. As shown in Table 1, our proposed PromptBio significantly outperforms all the compared MIL baselines in AUC. In particular, *e.g.*, over 5% higher than the second best method TOP* [20] and the third best method DSMIL [14] in TCGA MSI classification. Without pathology text prompts, DSMIL [14] retains a large number of irrelevant instances, making it challenging to learn genetic biomarker representations. While our PromptBio performs fine-grained pathological component grouping based on pathology text prompts, which effectively extracts instances more related to genetic biomarkers. Though TOP* [20] models different pathlogical components guided by pathology language prior knowledge, it neglects the interactions among different components.

   We further visualize the attention of each patch in pathology images with MSI or MSS cancer and the distribution of pathological components. In Fig. 2 (a), we find that in the pathology image with MSI cancer, large regions receive high attention. The selected patches of inflammatory response and lymphatic infiltration respectively show the accumulation of neutrophils and lymphocytes.

Table 1. AUC comparison on the TCGA dataset and CPTAC dataset. * denotes that we utilize the framework of the method for our implementation. The best results are marked in bold. Improvements compared with the second best results are highlighted.

Method	TCGA(MSI)	TCGA(BRAF)	CPTAC(MSI)
ABMIL [11]	0.8419	0.7115	0.8000
DSMIL [14]	0.8629	0.7310	0.8250
CLAM-SB [18]	0.8375	0.7522	0.8750
CLAM-MB [18]	0.8432	0.6373	0.8375
TransMIL [23]	0.8571	0.6708	0.8875
DTFD [28]	0.8502	0.7467	0.8875
MHIM-MIL(TransMIL) [26]	0.7720	0.6840	0.8750
TOP* [20]	0.8632	0.7176	0.8875
PromptBio	**0.9149**↑5.17	**0.8025**↑5.02	**0.9125**↑2.50

While in the pathology image with MSS cancer, regions receiving high attention are much smaller and there are few high-similarity pathology patches of the four conponents. This indicates that prompt-guided fine-grained pathological component grouping can ensure the accuracy of the tumor microenvironment modeling and enhance clinical interpretability. Fig. 2 (c) shows the feature mixtureness of bag representations on testing set. Compared with other methods, our PromptBio yields more discriminative representations.

3.3   Ablation Study

Table 2 summarizes the results of ablation study. We evaluate the effectiveness of the proposed modules. Without prompt-guided fine-grained grouping module, the performance drops, *e.g.*, from 91.49% to 89.65% in TCGA MSI classification. Without coarse-grained instance selection, the performance also decreases. These indicate that extracted instances based on medical prior knowledge prompting are more related to genetic biomarkers and they can better reflect the tumor microenvironment, while irrelevant pathology instances impeding the learning process of the model. A performance drop without fine-grained pathological component interaction mining module can be observed, *e.g.*, from 91.49% to 85.97% in TCGA MSI classification, which shows the importance of exploring the interaction among different pathological components in the tumor microenvironment. Replacing our prompt-guided fine-grained grouping module by K-means clustering algorithm leads to an obvious performance drop. This provides evidence for the effectiveness of the pathology text prompting strategy, enabling a robust modeling of the actual distribution of pathological components. Furthermore, to show results are not sensitive *w.r.t.* GPT-4, we change the seed of inference parameters in GPT-4. Though fine-grained text prompt slightly differs, we can still observe similar results (*e.g.*, 91.56% vs 91.49% for MSI classification).

Fig. 3 shows performance changes by varying the selection ratio β. It can be seen that the performance is not sensitive within the range of [0.2,0.4]. The model's performance does not always increase as β increases, as a larger selection ratio may result in the retention of a large number of irrelevant instances. This observation demonstrates the effectiveness of our prompt-based grouping strategy in extracting instances that are more closely related to genetic biomarkers.
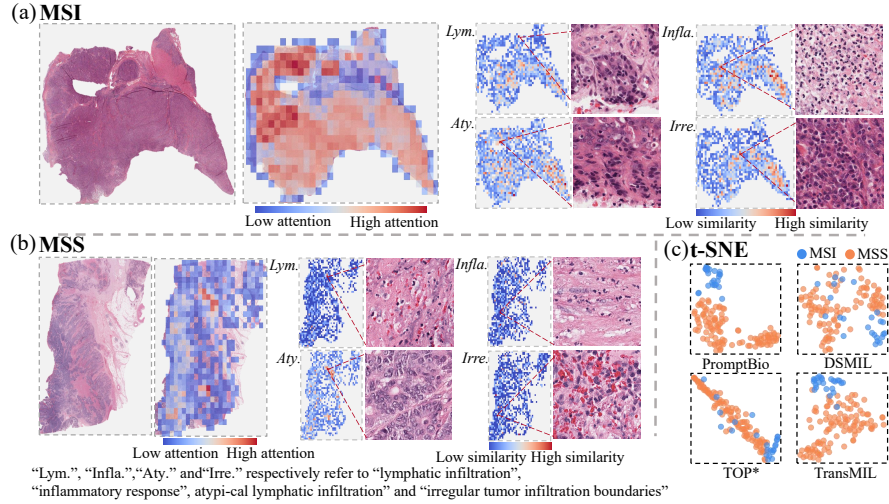
Fig. 2. Attention and t-SNE visualizations. For (a) MSI cancer and (b) MSS cancer, we show attention about contribution of each patch in a pathology image to the class token, and visualization of similarity between each patch and pathological components. "*Lym.*", "*Infla.*","*Aty.*" and "*Irre.*" respectively refer to "lymphatic infiltration", "inflammatory response", atypical lymphatic infiltration" and "irregular tumor infiltration boundaries". (c) t-SNE visualization of bag representations in different methods.
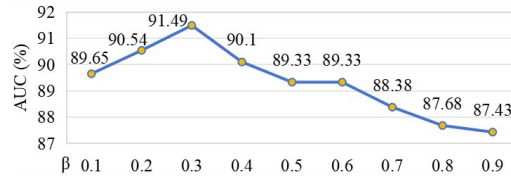

Fig. 3. Performance changes by varying the selection ratio β on TCGA MSI dataset.

4   Conclusion

In this paper, we propose a genetic biomarker prediction method via prompting techniques. We first extract instances related to genetic biomarkers based on medical prior knowledge. Then, guided by pathology text prompts, we group pathology instances from WSIs into fine-grained pathological components and mine their interaction, which enhances clinical interpretability. Experimental results on two CRC cohorts show our model achieves much better performance than other state-of-the-art methods.

Table 2. Ablation study on TCGA MSI classification, TCGA BRAF classification and CPTAC MSI classification. AUC is reported. Selection, PGG and PCIM denote our coarse-grained instance **Selection** module, **P**rompt-**G**uided fine-grained **G**rouping module and fine-grained **P**athological **C**omponent **I**nteraction **M**ining module, respectively. "K-mean" means replacing PGG with K-means algorithm. "△" means we change the seed of inference parameters in GPT-4 to generate component descriptions.

Selection	PGG	PCIM	TCGA(MSI)	TCGA(BRAF)	CPTAC(MSI)
✓			0.8965	0.7310	0.8000
	✓	✓	0.8902	0.7656	0.8875
✓	✓		0.8597	0.7690	0.8625
✓	K-means	✓	0.8267	0.7266	0.7875
✓	△	✓	**0.9156**	0.8013	0.9125
✓	✓	✓	0.9149	**0.8025**	**0.9125**

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Ajani, J.A., D'Amico, T.A., Bentrem, D.J., Cooke, D., Corvera, C., Das, P., Enzinger, P.C., Enzler, T., Farjah, F., Gerdes, H., et al.: Esophageal and esophagogastric junction cancers, version 2.2023, nccn clinical practice guidelines in oncology. Journal of the National Comprehensive Cancer Network **21**(4), 393–422 (2023)
2. Becht, E., de Reyniès, A., Giraldo, N.A., Pilati, C., Buttard, B., Lacroix, L., Selves, J., Sautès-Fridman, C., Laurent-Puig, P., Fridman, W.H.: Immune and stromal classification of colorectal cancer is associated with molecular subtypes and relevant for precision immunotherapy. Clinical cancer research **22**(16), 4057–4066 (2016)
3. Bożyk, A., Wojas-Krawczyk, K., Krawczyk, P., Milanowski, J.: Tumor microenvironment—a short review of cellular and interaction diversity. Biology **11**(6), 929 (2022)
4. Brown, K.M., Xue, A., Smith, R.C., Samra, J.S., Gill, A.J., Hugh, T.J.: Cancer-associated stroma reveals prognostic biomarkers and novel insights into the tumour microenvironment of colorectal cancer and colorectal liver metastases. Cancer Medicine **11**(2), 492–506 (2022)
5. Bussard, K.M., Mutkus, L., Stumpf, K., Gomez-Manzano, C., Marini, F.C.: Tumor-associated stromal cells as key contributors to the tumor microenvironment. Breast Cancer Research **18**, 1–11 (2016)
6. Caputo, F., Santini, C., Bardasi, C., Cerma, K., Casadei-Gardini, A., Spallanzani, A., Andrikou, K., Cascinu, S., Gelsomino, F.: Braf-mutated colorectal cancer: clinical and molecular insights. International journal of molecular sciences **20**(21), 5369 (2019)
7. Chang, L., Chang, M., Chang, H.M., Chang, F.: Microsatellite instability: a predictive biomarker for cancer immunotherapy. Applied Immunohistochemistry & Molecular Morphology **26**(2), e15–e21 (2018)
8. Dedeurwaerdere, F., Claes, K.B., Van Dorpe, J., Rottiers, I., Van der Meulen, J., Breyne, J., Swaerts, K., Martens, G.: Comparison of microsatellite instability detection by immunohistochemistry and molecular techniques in colorectal and endometrial cancer. Scientific reports **11**(1), 12880 (2021)

9. Hou, W., He, Y., Yao, B., Yu, L., Yu, R., Gao, F., Wang, L.: Multi-scope analysis driven hierarchical graph transformer for whole slide image based cancer survival prediction. In: MICCAI (2023)

10. Huang, Z., Bianchi, F., Yuksekgonul, M., Montine, T.J., Zou, J.: A visual–language foundation model for pathology image analysis using medical twitter. Nature medicine **29**(9), 2307–2316 (2023)

11. Ilse, M., Tomczak, J., Welling, M.: Attention-based deep multiple instance learning. ICML (2018)

12. Kather, J.N., Pearson, A.T., Halama, N., Jäger, D., Krause, J., Loosen, S.H., Marx, A., Boor, P., Tacke, F., Neumann, U.P., et al.: Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. Nature medicine **25**(7), 1054–1056 (2019)

13. Koncina, E., Haan, S., Rauh, S., Letellier, E.: Prognostic and predictive molecular biomarkers for colorectal cancer: updates and challenges. Cancers **12**(2), 319 (2020)

14. Li, B., Li, Y., Eliceiri, K.W.: Dual-stream multiple instance learning network for whole slide image classification with self-supervised contrastive learning. In: CVPR (2021)

15. Li, H., Yang, F., Zhao, Y., Xing, X., Zhang, J., Gao, M., Huang, J., Wang, L., Yao, J.: Dt-mil: deformable transformer for multi-instance learning on histopathological image. In: MICCAI. Springer (2021)

16. Li, H., Zhu, C., Zhang, Y., Sun, Y., Shui, Z., Kuang, W., Zheng, S., Yang, L.: Task-specific fine-tuning via variational information bottleneck for weakly-supervised pathology whole slide image classification. In: CVPR (2023)

17. Lin, T., Yu, Z., Hu, H., Xu, Y., Chen, C.W.: Interventional bag multi-instance learning on whole-slide pathological images. In: CVPR (2023)

18. Lu, M.Y., Williamson, D.F., Chen, T.Y., Chen, R.J., Barbieri, M., Mahmood, F.: Data-efficient and weakly supervised computational pathology on whole-slide images. Nature biomedical engineering **5**(6), 555–570 (2021)

19. Macenko, M., Niethammer, M., Marron, J.S., Borland, D., Woosley, J.T., Guan, X., Schmitt, C., Thomas, N.E.: A method for normalizing histology slides for quantitative analysis. In: ISBI (2009)

20. Qu, L., Fu, K., Wang, M., Song, Z., et al.: The rise of ai language pathologists: Exploring two-level prompt learning for few-shot weakly-supervised whole slide image classification. Advances in Neural Information Processing Systems **36** (2024)

21. Qu, L., Yang, Z., Duan, M., Ma, Y., Wang, S., Wang, M., Song, Z.: Boosting whole slide image classification from the perspectives of distribution, correlation and magnification. In: ICCV (2023)

22. Saillard, C., Dubois, R., Tchita, O., Loiseau, N., Garcia, T., Adriansen, A., Carpentier, S., Reyre, J., Enea, D., von Loga, K., et al.: Validation of msintuit as an ai-based pre-screening tool for msi detection from colorectal cancer histology slides. Nature Communications **14**(1), 6695 (2023)

23. Shao, Z., Bian, H., Chen, Y., Wang, Y., Zhang, J., Ji, X., et al.: Transmil: Transformer based correlated multiple instance learning for whole slide image classification. Advances in neural information processing systems **34**, 2136–2147 (2021)

24. Shimada, Y., Okuda, S., Watanabe, Y., Tajima, Y., Nagahashi, M., Ichikawa, H., Nakano, M., Sakata, J., Takii, Y., Kawasaki, T., et al.: Histopathological characteristics and artificial intelligence for predicting tumor mutational burden-high colorectal cancer. Journal of gastroenterology **56**(6), 547–559 (2021)

25. Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., Bray, F.: Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians **71**(3), 209–249 (2021)
26. Tang, W., Huang, S., Zhang, X., Zhou, F., Zhang, Y., Liu, B.: Multiple instance learning framework with masked hard instance mining for whole slide image classification. In: ICCV (2023)
27. Tsai, P.C., Lee, T.H., Kuo, K.C., Su, F.Y., Lee, T.L.M., Marostica, E., Ugai, T., Zhao, M., Lau, M.C., Väyrynen, J.P., et al.: Histopathology images predict multi-omics aberrations and prognoses in colorectal cancer patients. Nature communications **14**(1), 2102 (2023)
28. Zhang, H., Meng, Y., Zhao, Y., Qiao, Y., Yang, X., Coupland, S.E., Zheng, Y.: Dtfd-mil: Double-tier feature distillation multiple instance learning for histopathology whole slide image classification. In: CVPR (2022)