Supplementary Material

Integrative epigenomic profiling by high-speed super-resolution imaging and deep learning reveals chromatin state biomarkers

Yicheng Wang[1†], Nur Syatila Ab Ghani[2†], Munmee Dutta[3], Shungo Adachi[4], Kaoru Katoh[5], Masakazu Namihira[6,7*], Toutai Mitsuyama[3], Yutaka Saito[1,2,3*]

[1] Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-0882, Japan

[2] Department of Data Science, School of Frontier Engineering, Kitasato University, 1-15-1 Kitazato, Minami-ku, Sagamihara, Kanagawa 252-0373, Japan

[3] Artificial Intelligence Research Center, National Institute of Advanced Industrial Science and Technology (AIST), 2-4-7 Aomi, Koto-ku, Tokyo 135-0064, Japan

[4] Department of Proteomics, National Cancer Center Research Institute, Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045, Japan

[5] Exploratory Research Center on Life and Living Systems, National Institutes of Natural Sciences, 5-1 Higashiyama Myodaijicho, Okazaki, Aichi 444-8787, Japan

[6] Molecular Biosystems Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan

[7] Laboratory of Neural Regeneration and Brain Repair, Division of Biological Science, Graduate School of Science and Technology, Nara Institute of Science and Technology (NAIST), 8916-5 Takayama-cho, Ikoma, Nara, 630-0192, Japan

† These authors contributed equally to this work.

* To whom correspondence should be addressed. Email: saito.yutaka@kitasato-u.ac.jp.

Correspondence may also be addressed to Masakazu Namihira, Email: m-namihira@aist.go.jp.

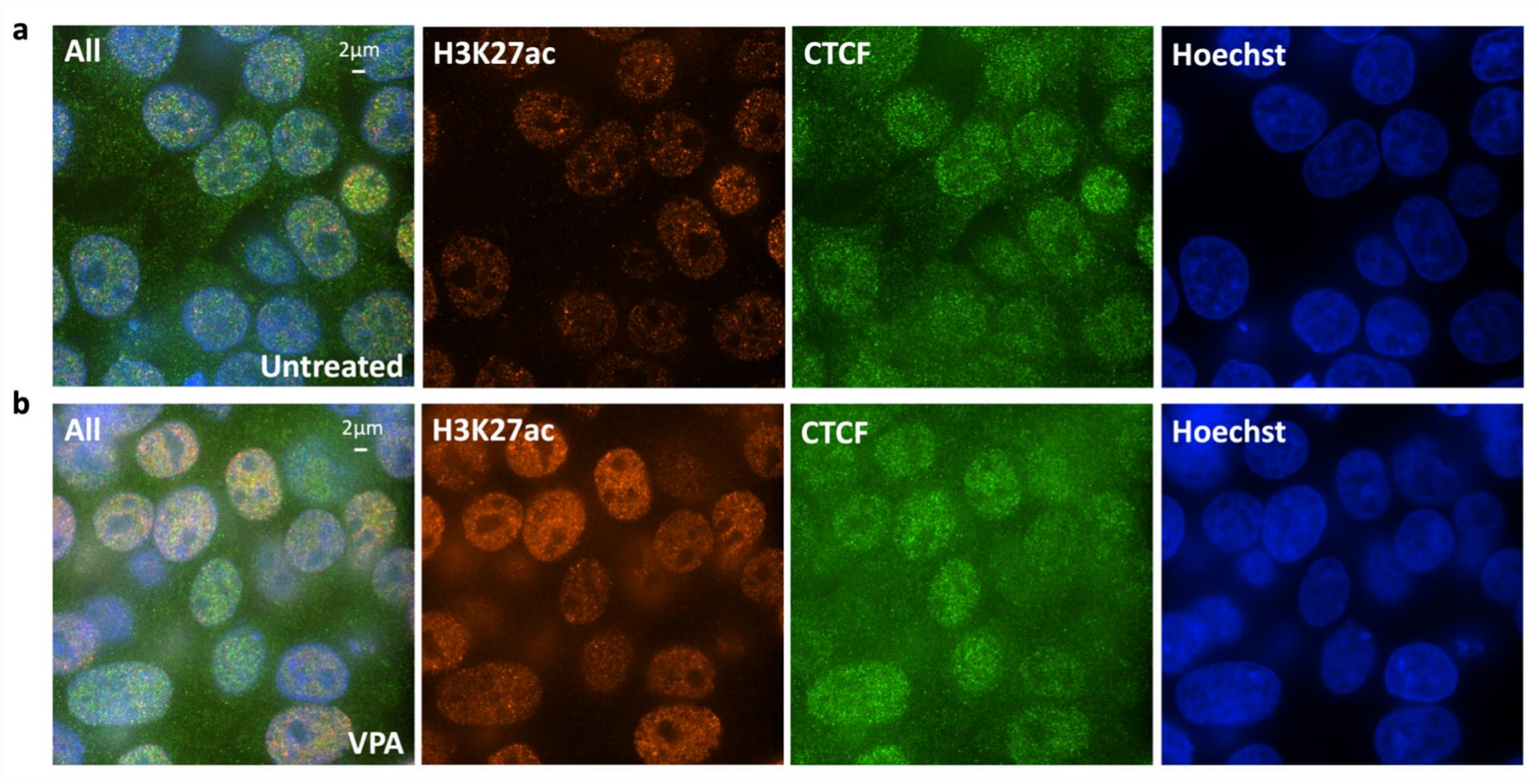

Figure S1. Example images of cells captured using the high-speed super-resolution microscopy SoRa. (a) Untreated cells under various staining conditions. **(b)** VPA-treated cells under various staining conditions.

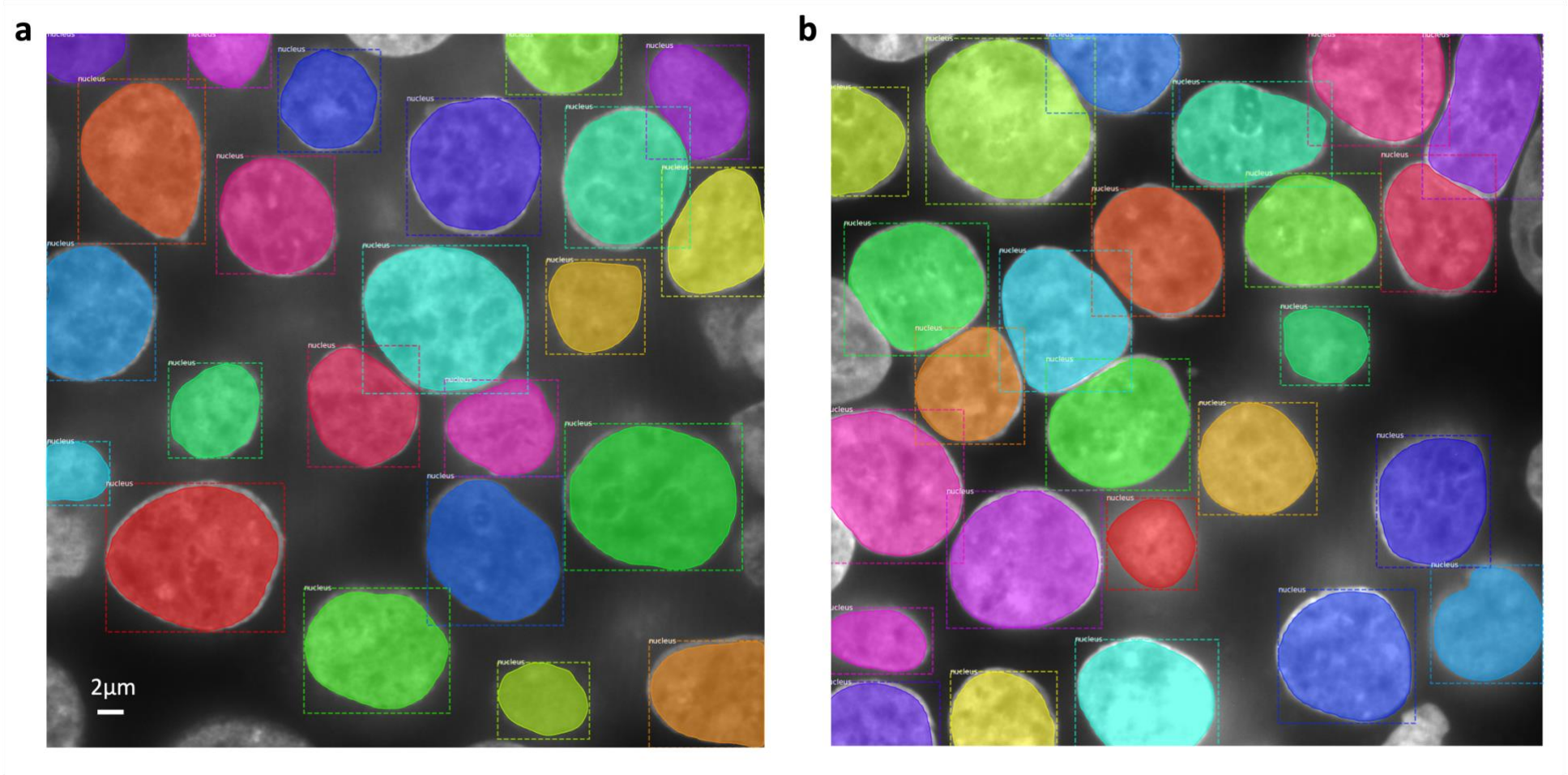

Figure S2. Examples of nucleus segmentation by Mask RCNN. The varied color schemes represent different nucleus that have been segmented. **(a)** Example segmentation result from an untreated cell image. **(b**) Example segmentation result from a VPA-treated cell image.

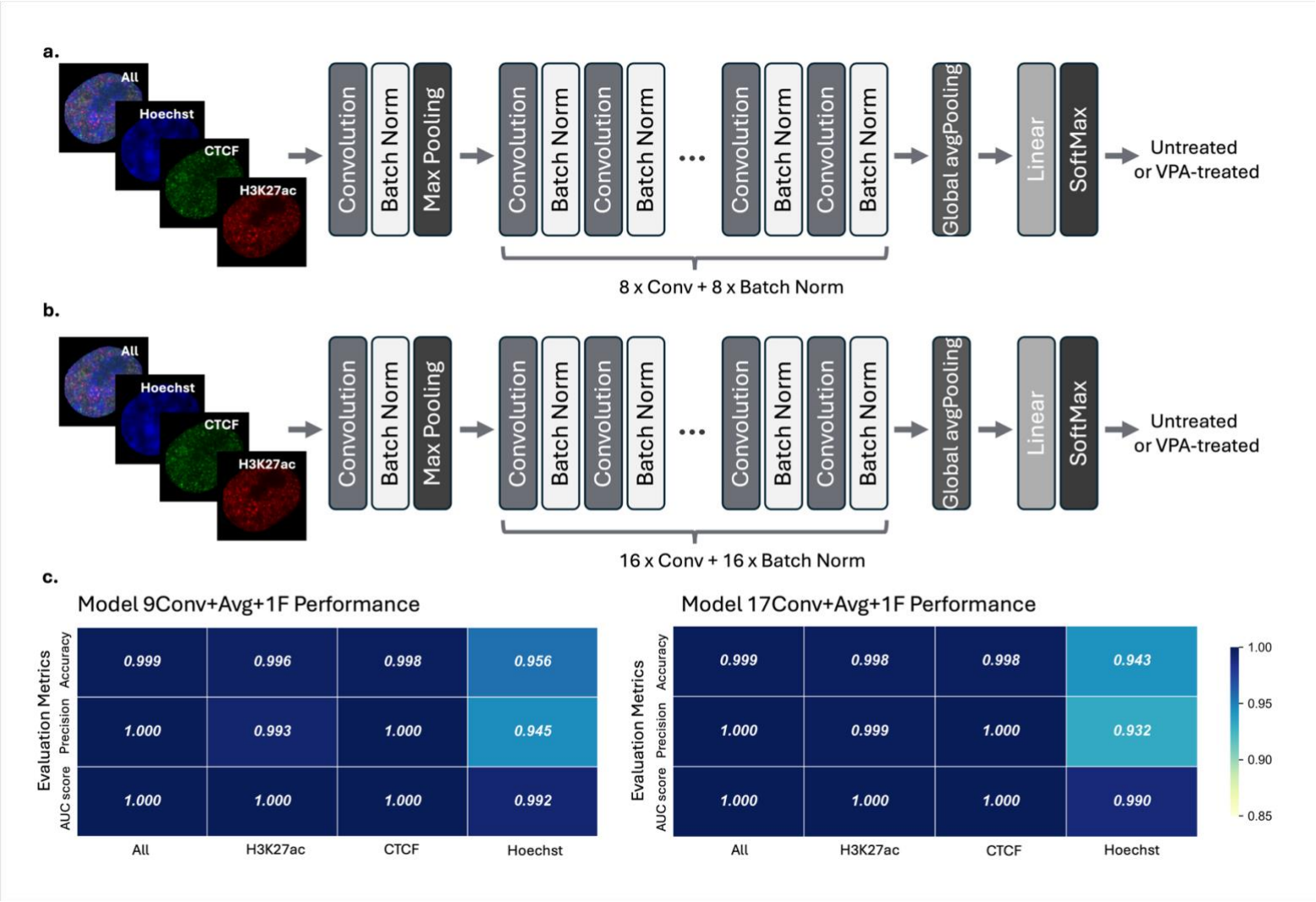

Figure S3. Two additional models employed in this study. (a) 9Conv+Avg+1F model, which includes an average pooling layer. **(b)** 17Conv+Avg+1F model, also featuring an average pooling layer. **(c)** Performance of these two models.

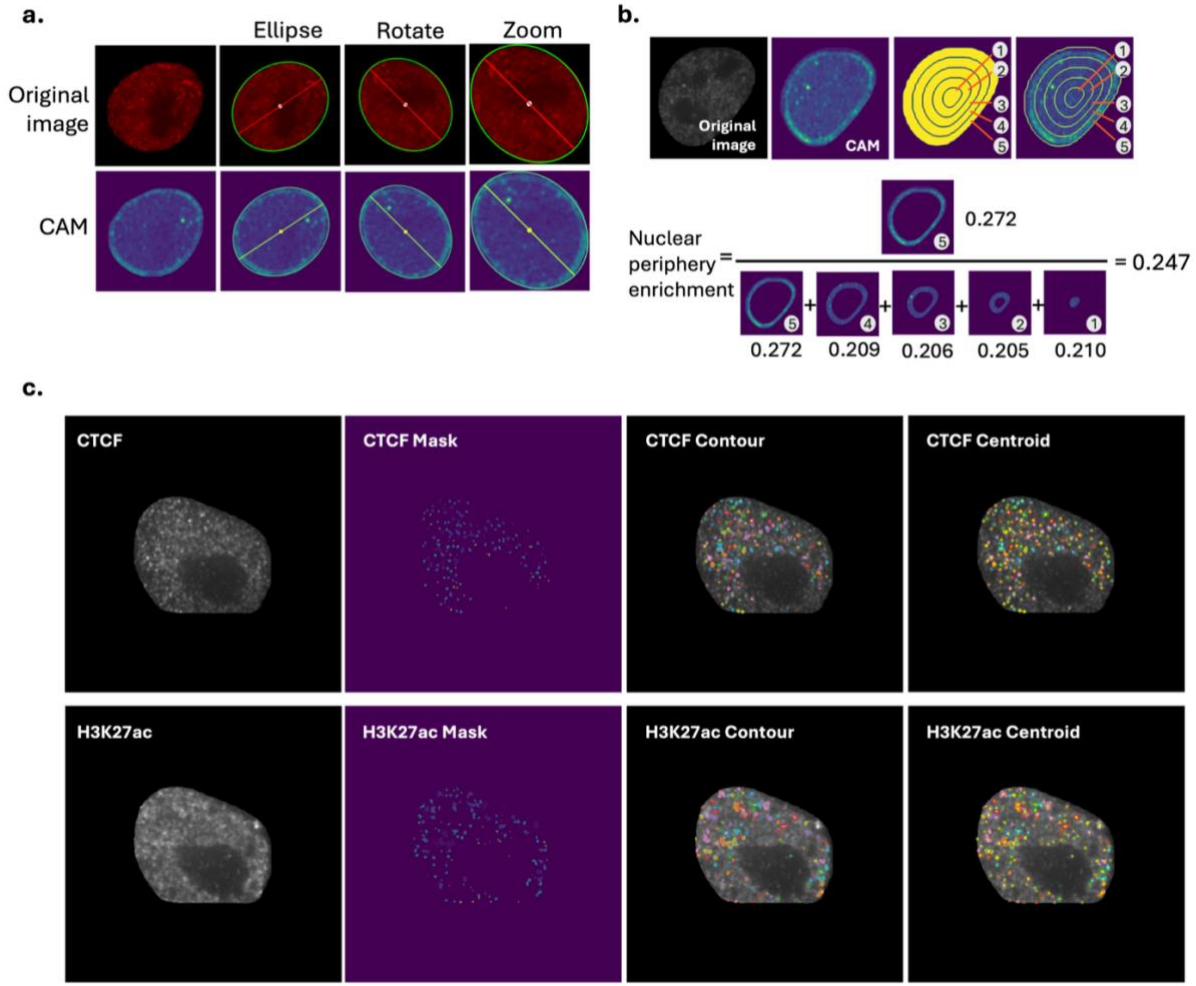

Figure S4. Overview of CAM heatmap processing and puncta segmentation methods. (a) Workflow for processing Score-CAM heatmaps, including alignment to fitted nuclear ellipses for generating average heatmaps. **(b)** Procedure for calculating nuclear periphery enrichment (Methods). **(c)** Examples of puncta segmentation results for CTCF (top) and H3K27ac (bottom), with corresponding puncta centroids indicated.

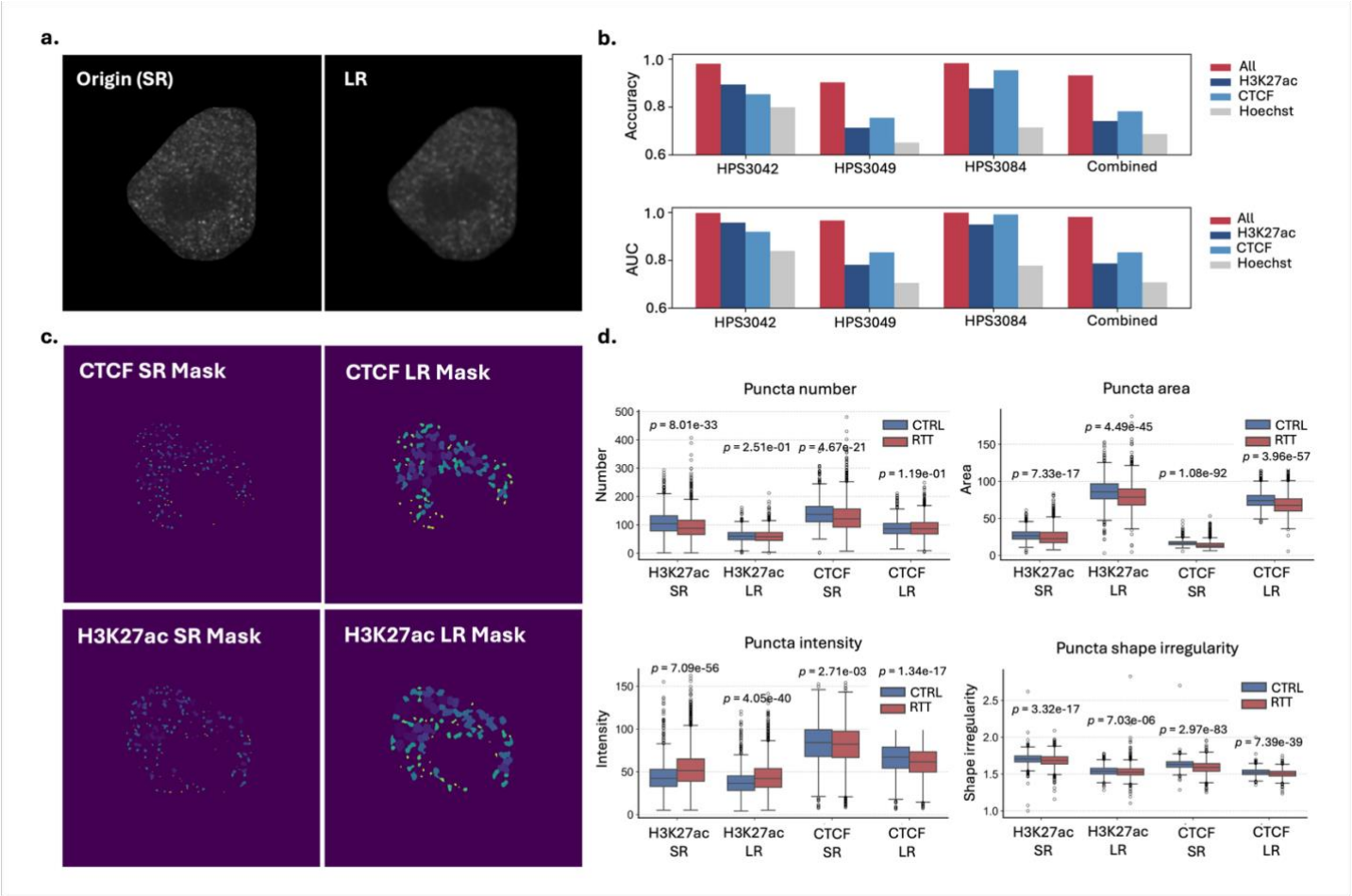

Figure S5. Analysis of simulated low-resolution image in Rett syndrome and control cells. (a) Comparison between original super-resolution (SR) images and simulated low-resolution (LR) images. **(b)** Classification performance of the same CNN model applied to Rett and control cells using LR data. Results are shown for models trained separately on each mutation line as well as for a combined model trained across all lines. **(c)** Examples of puncta segmentation in LR images. **(d)** Quantitative analysis of nuclear puncta properties in LR images compared to SR images.