Cross-Species Morphology Learning Enables Nucleic Acid-Independent Detection of Live Mutant Blood Cells

Sarmad Ahmad Khan[1], Dominik Faerber[2], Danielle Kirkey[3,4], Simon Raffel[5], Brandon Hadland[3,4], Michael Deininger[6]*, Florian Buettner[1,7,8,9]*, Helong Gary Zhao[1,3]*

[1] German Cancer Research Center (DKFZ), Heidelberg, Germany.

[2] Department of Information Technology and Electrical Engineering, ETH Zurich, Zurich, Switzerland.

[3] Translational Science and Therapeutics Division, Fred Hutchinson Cancer Center, Seattle, USA.

[4] Department of Pediatrics, University of Washington, Seattle, USA.

[5] Department of Medicine, Hematology, Oncology and Rheumatology, University Hospital Heidelberg, Heidelberg, Germany.

[6] Versiti Blood Research Institute, Milwaukee, USA.

[7] German Cancer Consortium (DKTK), partner site Frankfurt/Mainz, a partnership between DKFZ and UCT Frankfurt-Marburg, Germany.

[8] Frankfurt Cancer Institute, Frankfurt am Main, Germany.

[9] Department of Medicine and Institute of Informatics, Goethe University Frankfurt, Frankfurt am Main, Germany.

* These are co-corresponding authors.

Lead contact, HGZ: hzhao2@fredhutch.org

Abstract

In hematology/oncology clinics, molecular diagnostics based on nucleic acid sequencing or hybridization are routinely employed to detect malignancy-associated genetic mutations and are instrumental in therapeutic stratification and prognostication. However, their limited cost-efficiency constrains their use in pre-malignant screening—specifically, the detection of rare circulating mutant blood cells in asymptomatic individuals. In both neonates and adults, the presence of malignancy-associated mutations in peripheral blood correlates with an elevated risk of future neoplastic transformation, with certain mutations, such as *KMT2A* rearrangements, exhibiting near-complete penetrance. If feasible, pre-malignant screening could enable early intervention and even disease prevention. Here, we introduce a high-throughput, single-cell computer vision platform capable of identifying mutant peripheral blood cells by recognizing mutation-specific morphological features. The morphology recognition module was developed through cross-species learning from murine to human datasets, enabling a generalizable and cost-effective approach for detecting mutations in live blood cells. The platform holds promise for translation into pre-malignant screening applications in asymptomatic neonates and adults as well as measurable residual disease monitoring in malignancies. Furthermore, it provides a novel single-cell morphological data modality that complements existing molecular layers, including genomics, epigenomics, transcriptomics, and proteomics.

Introduction

Mutation detection is central to the molecular diagnostics of hematologic malignancies[1] and is critical for tracing measurable residual disease (MRD)[2,3]. In pre-malignant conditions, the presence of malignancy-associated mutations in asymptomatic adults—termed clonal hematopoiesis—is associated with increased risks of hematologic malignancies and other chronic diseases, such as cardiovascular diseases[4,5]. Similarly, oncogenic mutations can be detected in peripheral blood (PB) samples of newborns who develop pediatric acute leukemia later in life[6]. Notably, the presence of *KMT2A*-rearrangment mutations in cord blood is associated with near 100% progression to pediatric acute leukemia[6]. These pre-malignant conditions underscore the potential for detecting early mutant lesions months to years (in pediatric cases) or even decades (in adults) before disease onset, allowing potential preventative interventions. Therefore, cost-efficient detection of rare mutant cells in both MRD and pre-malignant states holds significant clinical value.

Currently, oncogenic mutations are primarily detected using nucleic acid-based techniques, such as DNA sequencing and *in situ* hybridization. In hematologic malignancies, the standard-of-practice for high-sensitivity detection of common mutations is next-generation sequencing (NGS)[1]. However, the high cost and technical limitations of NGS[7] have hindered its application in screening asymptomatic newborns and adults. To enable screening, a novel approach is needed.

Machine learning (ML)-based image recognition has been shown to match or even out-perform medical professionals in cancer diagnosis[8]. In hematopathology, ML models trained on pathologist-annotated blood smears have shown high accuracy in cell lineage classification[9] and drug response prediction in adult acute myeloid leukemia (AML)[10,11]. With the goal of enabling pre-malignant screening, we sought to detect mutant blood cells by their native morphology, through high-throughput imaging and supervised ML of single-cell images. Given the impracticality of acquiring large numbers of fresh patient samples for supervised learning—due to inter-individual heterogeneity and the challenge of obtaining mutation ground truth without compromising cellular morphology—we established a cross-species learning platform composed of a genetic modeling module and an imaging / learning module. In this approach, mutation-specific morphological features were learned from orthologous genetic mouse models and "humanized" using one pair of mutant and normal human samples (Fig. 1A). We tested the platform with two representative mutations with relevance to both pre-malignant screening and MRD: *KMT2A-MLLT3* (also known as *MLL*-AF9, a frequent mutation in pediatric and adult leukemias[12] and detectable in newborn PB[13,14]) and *Jak2*^{V617F} (ortholog of *JAK2*^{V617F}, a frequent mutation in myeloproliferative neoplasms[15,16] and clonal hematopoiesis[17,18]).

Methods

Genetic mouse models

For generating CH-like chimeric hematopoiesis with *KMT2A-MLLT3* or GFP-only mock control, C57BL/6J mouse BM cells were lentivirally transduced with *KMT2A-MLLT3*/GFP or GFP-only construct at a multiplicity of infection (MOI) of ~0.1 and transplanted into lethally irradiated isogenic littermates. For *Jak2*^{V617F} mutation, B6N.129S6(SJL)-*Jak2*$^{tm1.1Ble}$/AmlyJ mouse[19] BM cells were transduced with *cre*-GFP construct at a MOI of ~0.1 to activate the V617F mutation before isogenic transplantation. PB samples with detectable GFP^{+} cells (>1%) were used for imaging. Mouse studies were approved by institutional IACUC committees.

Human samples

For young healthy donors, PB was collected by lancet finger prick or by phlebotomy and stored in EDTA-K tubes for < 24 hours. For patients with *JAK2*^{V617F}, PB was collected by phlebotomy, and stored in EDTA-K tubes for < 24 hours. For AML patients with *KMT2A-MLLT3*, diagnostic BM samples were engrafted into irradiated NSG-SGM3 mice (NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl} Tg(CMV-IL3,CSF2,KITLG)1Eav/MloySzJ) as patient-derived xenografts (PDXs). PB samples from PDX-carrying mice with detectable hCD45^{+} cells (>1%) were used for imaging. Blood and BM donors have consented to participation in medical research. All human samples were deidentified prior to being used for experiments in this study. Use of deidentified human samples has been approved by institutional ethics committees.

Single-cell imaging and ML

Single-cell images were obtained in a high-throughput fashion using Amnis ImageStream® imaging flow cytometer (Cytek Biosciences, Fremont, CA). Before imaging, 10 µL of PB sample was transferred to a sterile Eppendorf tube, and 100 µL of PBS (with 5 µM nuclear dye – DRAQ5 and 1 µg/mL dead cell dye - 7-AAD) was added to the tube to dilute the sample. 5 minutes later, diluted samples were loaded on to the Amnis ImageStream to obtain single-cell images in five channels: brightfield, side scatter (SCC / darkfield), GFP, 7AAD, and DRAQ5 (Suppl. Fig. 1A). Recorded events were gated to remove multi-cell clusters, cells out of focus, cells without nuclei, and dead cells. GFP (mutant cell label) and 7AAD (dead cell label) channels were used in cytometer gating but not in ML. ML was limited to channels of brightfield, SCC, and DRAQ5 (nucleus). Brightfield (BF), side scatter (SSC), and nuclear images of selected cells were exported as 16-bit raw pixel value tif files or as cif files using IDEAS software (Cytek Biosciences) for ML training and testing. Morphological and textural features of each cell were extracted using three open-source tools, namely *Pyfeats*[20], *scikit-image*[21], and *CellProfiler*[22], and the best-performing tool was selected for each model. For CellProfiler, cif files were used and a pipeline was implemented to process cif files and stitch single-cell images into 30 × 30 grids for systematic feature extraction. ML carried out with XGBoost (Suppl. Fig. 1B). Model performance was evaluated by balanced accuracy (*BA*) in "Leave-One-Out" (LOO) cross-validations (one biological sample insulated from training is used in blinded external validation), area under the curve of receiver operating characteristic (AUC-ROC), area under the curve of precision recall (AUC-PR). The complete code of the ML analysis is deposited on our GitHub (https://github.com/MLO-lab/leukemia-morphology-analysis). Additional description of feature extraction and ML is provided in the supplementary information.

Results and Discussion

The cross-species learning platform consists of two main modules: (1) mouse genetic / PDX models, and (2) single-cell imaging and ML-based image recognition (Fig. 1A). Module 1 generates materials for ML training and testing. In each genetic model, a single mutation, together with GFP marker, was introduced to form chimeric hematopoiesis that contains both mutant cells (GFP^{+}) and normal cells (GFP^{-}). As mock treatment controls, cells with GFP only were included in testing to ensure specificity. Module 2 captures and learns the native morphologies of fresh PB cells. For better interpretability in ML, we employed an explainable AI (XAI) approach of image feature extraction followed by decision tree learning (XGBoost[23]). ML models were initially developed using mouse genetic models, and then optimized and "humanized" by incorporating one pair of mutant and normal human samples in training. To ensure robustness and generalizability of our ML models, we used human samples collected from three different hospitals in external model testing (Fig. 1B). Moreover, to eliminate "batch effects", images were recorded on four different ImageStream® cytometers with variable hardware configurations at four different locations (Fig. 1B).

KMT2A-MLLT3^{+} leukemic cells typically present myelocytic to monocytic phenotype. In our mouse model, however, *KMT2A-MLLT3*$^{+}$ PB cells showed a variety of morphologies with different sizes and opacities (Fig. 1C). Despite such morphological diversity, feature-aware decision tree learning (XGBoost) achieved appreciable accuracies in mutant cell detection, with the best performing feature extraction tool being

CellProfiler[22] (Fig. 1D, Suppl. Fig. 1C, 2A). Importantly, our XAI approach demonstrated superior accuracies over other popular image classification approaches, including deep learning tools (ConvNeXt[24] and BEiT[25]) (Suppl. Fig. 1D). In comparison, three anonymous hematologists who had not been trained on unstained cell images were not able to distinguish between *KMT2A-MLLT3*$^{+}$ and normal cells (balanced accuracy, *BA*: 0.30 ~ 0.60) (Fig. 1E). One anonymous hematology professional was voluntarily trained on annotated images and showed improved *BA* in a blinded test (Fig. 1E). This suggests that *KMT2A-MLLT3* causes subtle changes in cellular native morphology that are detectable by ML and are at least partially recognizable by trained human.

However, XGBoost models trained on mouse datasets exhibited inconsistent performance in distinguishing human *KMT2A-MLLT3*$^{+}$ cells vs. normal counterparts (Fig. 2A), suggesting that XGBoost models trained solely on mouse datasets can be confounded by species-specific morphological features that are not conserved in human cells. To overcome this limitation, one pair of randomly selected human patient sample (from PDX) and normal human donor sample were incorporated into the training datasets. This hybrid training approach markedly improved model performance in identifying mutant human cells from previously unseen samples (Fig. 2B). So far, we have assessed the performance of each ML model by computing the *BA* for each testing sample, which reflects how often a correct prediction is made on a single-cell level. However, for future translational applications, our primary interest lies not in cell-level predictions, but in assessing the likelihood that a given sample carries a mutation. To this end, we performed analysis at the sample level (*i.e.*, donor/patient level), assigning a score to each donor/patient. This score is calculated as the mean of the confidence values across all cell-level predictions associated with that donor/patient. Subsequently, we evaluate performance metrics across all donors/patients within a single testing cohort (Fig. 2C-D). The drastically improved sample-wise ROC curve (AUC=1.0) and PR curve (AUC=1.0) of the mouse+human training method underscored the importance of supplementing a pair of human samples in the cross-species learning (Fig. 2D). Additionally, this improved training method demonstrated high confidence in cell-level predictions across novel human samples (Fig. 2E). To further evaluate the system, we analyzed synthetic mixtures of healthy donor cells and *KMT2A-MLLT3*$^{+}$ patient cells at predefined ratios. The models' predictions showed a strong linear correlation with the actual proportions of mutant cells (Fig. 2F, Suppl. Fig. 3), indicating robust quantitative fidelity. Such linearity enables further statistical optimization for large-scale retrospective clinical studies. Collectively, these findings demonstrated that supplementing mouse datasets with a single pair of mutant and normal human samples is sufficient to ensure high accuracy in cross-species ML models for analyzing novel human specimens.

To elucidate the specific morphological alterations induced by *KMT2A-MLLT3* via XAI, we used SHAP (SHapley Additive exPlanations[26]) values to identify the features most influential in predicting mutant cells. In mutant mouse cells, the top features include altered nucleus texture and increased heterogeneity in cell size (*e.g.* bounding box minimum) (Fig. 3A, Suppl. Fig. 4A). These are consistent with studies showing that *KMT2A-MLLT3* and AML are associated with upregulation of anabolic activity, increased nuclear dynamics, and upregulation of genes controlling cell size[27-31]. Among the features shared between mouse and human cells, several stood out: increased granularity/membrane ruffles (*e.g.* SSC intensity and SSC weighted moments[32]), altered cellular texture (*e.g.* BF weighted moments), and opacity (*e.g.* BF intensity) (Fig. 3B, Suppl. Fig. 4B). These features align with the characteristic intracellular granules[33] and the monocytic morphology[34,35] of *KMT2A-MLLT3*$^{+}$ AML. Furthermore, previous reports have shown that *KMT2A-MLLT3* expression increases both cell size and granularity in the myeloid cell line U937[36], further validating our observations. Taken together, these results reveal substantial overlap in *KMT2A-MLLT3*-associated morphological features between mouse and human cells, supporting effective cross-species learning. Notably, since all cell lineages were included in the decision tree learning, the identified mutation-associated features were cell type-agnostic and contribute to ML predictions in a unified, lineage-independent manner.

To explore the generalizability of our platform, we applied the same approach to *Jak2*^{V617F}/*JAK2*^{V617F} mutation. XGBoost models trained on *Jak2*^{V617F} mutant and normal mouse PB cells yielded *BA* of 0.68 ~ 0.91 in LOO cross-validations (Fig. 3C-D), with cell-wise ROC-AUC of 0.73 ~ 0.92 (Suppl. Fig. 5). To extend the model to human samples, we incorporated one normal human PB sample into the training dataset (uniformed mutant human samples were unavailable). In one of three fresh PB samples from patients with myelofibrosis and MDS/MPN, the model detected frequencies of *JAK2*^{V617F} -mutant cells consistent with estimates derived from diagnostic NGS (Fig. 3E, Suppl. Table 1). Due to the lack of ground truth of mutation status at the single-cell level in fresh human PB samples, model performance metrics were not determined. We postulate that the variability in model accuracy across patient samples reflects underlying genetic and disease heterogeneity (Suppl. Table 1). In future studies, we propose to use CRISPR-edited *JAK2*^{V617F} human HSPC xenografts in addition to mouse models as training materials, to overcome the experimental challenges associated with using patient samples and PDXs of *JAK2*^{V617F}-associated chronic diseases.

In summary, we developed a cross-species learning platform capable of detecting mutant human cells in fresh PB samples (Suppl. Table 2). We demonstrated that two myeloid malignancy-associated mutations induce morphological changes in mouse PB cells that are detectable by ML independent of nucleic acids. These morphological features reflect the biological impact of the oncogenic mutations, such as metabolic adaptations to elevated anabolic processes. Importantly, the inclusion of even a single mutant human sample during training significantly enhanced model accuracy in analyzing unseen human specimens. Our ongoing efforts aim to extend this approach to additional somatic mutations for applications in early diagnosis and disease monitoring.

Acknowledgements

We thank the members of the Hadland lab and the Meshinchi lab at FHCC, the Raffel lab at UKHD, the Deininger lab at VBRI for experimental assistance. We thank the members of the Buettner lab and the Schulz lab at DKFZ / Goethe University Frankfurt for helpful discussions. We thank the core facilities at UW-Madison, VBRI, DKFZ, and BwNW for instrumental support. We thank the FHCC/UW Leukemia Repository Team, especially Tim Monahan, for their help in patient sample collection and coordination. This work was supported by a Helmholtz International Researcher Fellowship, Fred Hutch IIRC Innovation Award, Cancer Consortium Bench-to-Bedside Adult Oncology Award (sponsored by Bezos Family Foundation) to HGZ, and NIH/NHLBI R01 grant (HL168110) to BH. This work was co-funded by the European Union (ERC, TAIPO, 101088594 to FB). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. SR's research is funded by German Research Foundation (DFG) grant RA 3611/1-1.

Authorship Contributions

HGZ and FB conceptualized the study in its current form. SK, DF, and HGZ performed ML works. HGZ and DF performed experiments, collected data, and analyzed data. MD and BH provided facilities and materials for mouse experiments. DK, SR, and HGZ provided human samples. SK, HGZ, and FB wrote the manuscript with inputs from co-authors. FB, HGZ, BH, and MD provided research funding. HGZ and FB jointly supervised the study.

Declaration of Interests

The authors declare no current conflict of interest. A provisional US patent related to this work was filed and later lapsed.

References

1. Taylor, J., Xiao, W., and Abdel-Wahab, O. (2017). Diagnosis and classification of hematologic malignancies on the basis of genetics. Blood *130*, 410-423. 10.1182/blood-2017-02-734541.
2. Heuser, M., Freeman, S.D., Ossenkoppele, G.J., Buccisano, F., Hourigan, C.S., Ngai, L.L., Tettero, J.M., Bachas, C., Baer, C., Bene, M.C., et al. (2021). 2021 Update on MRD in acute myeloid leukemia: a consensus document from the European LeukemiaNet MRD Working Party. Blood *138*, 2753-2767. 10.1182/blood.2021013626.
3. Dekker, S.E., Rea, D., Cayuela, J.M., Arnhardt, I., Leonard, J., and Heuser, M. (2023). Using Measurable Residual Disease to Optimize Management of AML, ALL, and Chronic Myeloid Leukemia. Am Soc Clin Oncol Educ Book *43*, e390010. 10.1200/EDBK_390010.
4. Beeler, J.S., Bick, A.G., and Bolton, K.L. (2023). Genetic causes and cardiovascular consequences of clonal hematopoiesis in the UK Biobank. Nat Cardiovasc Res *2*, 13-15. 10.1038/s44161-022-00198-3.
5. Walsh, K. (2024). The emergence of clonal hematopoiesis as a disease determinant. J Clin Invest *134*. 10.1172/JCI180063.
6. Marcotte, E.L., Spector, L.G., Mendes-de-Almeida, D.P., and Nelson, H.H. (2021). The Prenatal Origin of Childhood Leukemia: Potential Applications for Epidemiology and Newborn Screening. Front Pediatr *9*, 639479. 10.3389/fped.2021.639479.
7. Bacher, U., Shumilov, E., Flach, J., Porret, N., Joncourt, R., Wiedemann, G., Fiedler, M., Novak, U., Amstutz, U., and Pabst, T. (2018). Challenges in the introduction of next-generation sequencing (NGS) for diagnostics of myeloid malignancies into clinical routine use. Blood Cancer J *8*, 113. 10.1038/s41408-018-0148-6.
8. Tran, K.A., Kondrashova, O., Bradley, A., Williams, E.D., Pearson, J.V., and Waddell, N. (2021). Deep learning in cancer diagnosis, prognosis and treatment selection. Genome Med *13*, 152. 10.1186/s13073-021-00968-x.
9. Matek, C., Krappe, S., Munzenmayer, C., Haferlach, T., and Marr, C. (2021). Highly accurate differentiation of bone marrow cell morphologies using deep neural networks on a large image data set. Blood *138*, 1917-1927. 10.1182/blood.2020010568.
10. Heinemann, T., Kornauth, C., Severin, Y., Vladimer, G.I., Pemovska, T., Hadzijusufovic, E., Agis, H., Krauth, M.T., Sperr, W.R., Valent, P., et al. (2022). Deep Morphology Learning Enhances Ex Vivo Drug Profiling-Based Precision Medicine. Blood Cancer Discov *3*, 502-515. 10.1158/2643-3230.BCD-21-0219.
11. Eckardt, J.N., Middeke, J.M., Riechert, S., Schmittmann, T., Sulaiman, A.S., Kramer, M., Sockel, K., Kroschinsky, F., Schuler, U., Schetelig, J., et al. (2022). Deep learning detects acute myeloid leukemia and predicts NPM1 mutation status from bone marrow smears. Leukemia *36*, 111-118. 10.1038/s41375-021-01408-w.
12. Meyer, C., Larghero, P., Almeida Lopes, B., Burmeister, T., Groger, D., Sutton, R., Venn, N.C., Cazzaniga, G., Corral Abascal, L., Tsaur, G., et al. (2023). The KMT2A recombinome of acute leukemias in 2023. Leukemia *37*, 988-1005. 10.1038/s41375-023-01877-1.
13. Brown, P., Pieters, R., and Biondi, A. (2019). How I treat infant leukemia. Blood *133*, 205-214. 10.1182/blood-2018-04-785980.
14. Ford, A.M., Ridge, S.A., Cabrera, M.E., Mahmoud, H., Steel, C.M., Chan, L.C., and Greaves, M. (1993). In utero rearrangements in the trithorax-related oncogene in infant leukaemias. Nature *363*, 358-360. 10.1038/363358a0.
15. Levine, R.L., Loriaux, M., Huntly, B.J., Loh, M.L., Beran, M., Stoffregen, E., Berger, R., Clark, J.J., Willis, S.G., Nguyen, K.T., et al. (2005). The JAK2V617F activating mutation occurs in chronic myelomonocytic leukemia and acute myeloid leukemia, but not in acute lymphoblastic leukemia or chronic lymphocytic leukemia. Blood *106*, 3377-3379. 10.1182/blood-2005-05-1898.
16. Jelinek, J., Oki, Y., Gharibyan, V., Bueso-Ramos, C., Prchal, J.T., Verstovsek, S., Beran, M., Estey, E., Kantarjian, H.M., and Issa, J.P. (2005). JAK2 mutation 1849G>T is rare in acute leukemias but can be found in CMML, Philadelphia chromosome-negative CML, and megakaryocytic leukemia. Blood *106*, 3370-3373. 10.1182/blood-2005-05-1800.
17. Zink, F., Stacey, S.N., Norddahl, G.L., Frigge, M.L., Magnusson, O.T., Jonsdottir, I., Thorgeirsson, T.E., Sigurdsson, A., Gudjonsson, S.A., Gudmundsson, J., et al. (2017). Clonal hematopoiesis, with and without candidate driver mutations, is common in the elderly. Blood *130*, 742-752. 10.1182/blood-2017-02-769869.
18. Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P.V., Mar, B.G., Lindsley, R.C., Mermel, C.H., Burtt, N., Chavez, A., et al. (2014). Age-related clonal hematopoiesis associated with adverse outcomes. N Engl J Med *371*, 2488-2498. 10.1056/NEJMoa1408617.
19. Mullally, A., Lane, S.W., Ball, B., Megerdichian, C., Okabe, R., Al-Shahrour, F., Paktinat, M., Haydu, J.E., Housman, E., Lord, A.M., et al. (2010). Physiological Jak2V617F expression causes a lethal myeloproliferative neoplasm with differential effects on hematopoietic stem and progenitor cells. Cancer Cell *17*, 584-596. 10.1016/j.ccr.2010.05.015.
20. Giakoumoglou, N. (2021). PyFeats: Open-source software for image feature extraction.

21. van der Walt, S., Schonberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Gouillart, E., Yu, T., and scikit-image, c. (2014). scikit-image: image processing in Python. PeerJ *2*, e453. 10.7717/peerj.453.
22. Stirling, D.R., Swain-Bowden, M.J., Lucas, A.M., Carpenter, A.E., Cimini, B.A., and Goodman, A. (2021). CellProfiler 4: improvements in speed, utility and usability. BMC Bioinformatics *22*, 433. 10.1186/s12859-021-04344-9.
23. Chen, T., and Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. (ACM), pp. 785–794.
24. Liu, Z., Mao, H.Z., Wu, C.Y., Feichtenhofer, C., Darrell, T., and Xie, S.N. (2022). A ConvNet for the 2020s. Proc Cvpr Ieee, 11966-11976. 10.1109/Cvpr52688.2022.01167.
25. Bao, H., Dong, L., and Wei, F. (2021). BEiT: {BERT} Pre-Training of Image Transformers. arXiv, 2106.08254.
26. Lundberg, S.M., and Lee, S.I. (2017). A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems 30 (Nips 2017) *30*.
27. Zhu, Y., Murtadha, M., Liu, M., Caserta, E., Napolitano, O., Nguyen, L.X.T., Wang, H., Moloudizargari, M., Nigam, L., Tandoh, T., et al. (2025). Identification of CD84 as a potent survival factor in acute myeloid leukemia. J Clin Invest *135*. 10.1172/JCI176818.
28. Patnana, P.K., Liu, L., Frank, D., Nimmagadda, S.C., Behrens, M., Ahmed, H., Xie, X., Liebmann, M., Wei, L., Gerdemann, A., et al. (2023). Dose-dependent expression of GFI1 alters metabolism in the haematopoietic progenitors and MLL::AF9-induced leukaemic cells. Br J Haematol *202*, 1033-1048. 10.1111/bjh.18939.
29. Ni, F., Yu, W.M., Li, Z., Graham, D.K., Jin, L., Kang, S., Rossi, M.R., Li, S., Broxmeyer, H.E., and Qu, C.K. (2019). Critical role of ASCT2-mediated amino acid metabolism in promoting leukaemia development and progression. Nat Metab *1*, 390-403. 10.1038/s42255-019-0039-6.
30. Cadart, C., and Heald, R. (2022). Scaling of biosynthesis and metabolism with cell size. Mol Biol Cell *33*. 10.1091/mbc.E21-12-0627.
31. Mesbahi, Y., Trahair, T.N., Lock, R.B., and Connerty, P. (2022). Exploring the Metabolic Landscape of AML: From Haematopoietic Stem Cells to Myeloblasts and Leukaemic Stem Cells. Front Oncol *12*, 807266. 10.3389/fonc.2022.807266.
32. Jähne, B. (2005). Digital Image Processing, 6th Edition (Springer-Verlag).
33. Goasguen, J.E., Bennett, J.M., Bain, B.J., Vallespi, T., Brunning, R., Mufti, G.J., and International Working Group on Morphology of Myelodysplastic, S. (2009). Morphological evaluation of monocytes and their precursors. Haematologica *94*, 994-997. 10.3324/haematol.2008.005421.
34. Zeng, A.G.X., Iacobucci, I., Shah, S., Mitchell, A., Wong, G., Bansal, S., Chen, D., Gao, Q., Kim, H., Kennedy, J.A., et al. (2025). Single-cell Transcriptional Atlas of Human Hematopoiesis Reveals Genetic and Hierarchy-Based Determinants of Aberrant AML Differentiation. Blood Cancer Discov, OF1-OF18. 10.1158/2643-3230.BCD-24-0342.
35. Stavropoulou, V., Kaspar, S., Brault, L., Sanders, M.A., Juge, S., Morettini, S., Tzankov, A., Iacovino, M., Lau, I.J., Milne, T.A., et al. (2016). MLL-AF9 Expression in Hematopoietic Stem Cells Drives a Highly Invasive AML Expressing EMT-Related Genes Linked to Poor Outcome. Cancer Cell *30*, 43-58. 10.1016/j.ccell.2016.05.011.
36. Caslini, C., Shilatifard, A., Yang, L., and Hess, J.L. (2000). The amino terminus of the mixed lineage leukemia protein (MLL) promotes cell cycle arrest and monocytic differentiation. Proc Natl Acad Sci U S A *97*, 2797-2802. 10.1073/pnas.040574897.

Figure Legends

Figure 1. Study design and general performance of ML models in detecting PB cells with a single mutation. (A) Schematic illustration of the cross-species learning platform and the study design. CNN: convolutional neural network, MPN: myeloproliferative neoplasm. NSG: severely immunodeficient mouse strain that allows engraftment of human myeloid cells. **(B)** Study sites for human sample collection and cell imaging. FHCC: Fred Hutchinson Cancer Center, BwNW: Bloodworks Northwest, UW: University of Wisconsin-Madison, MCW: Medical College of Wisconsin, VBRI: Versiti Blood Research Institute, UKHD: Heidelberg University Hospital, DKFZ: German Cancer Research Center. **(C)** Representative images of chimeric PB samples from wildtype mouse with *KMT2A-MLLT3* mutation. Mutant cells are marked by GFP expression. Images were recorded on an ImageStream® device. **(D)** Performance of XGBoost models with different combinations of cell features and different mouse datasets with *KMT2A-MLLT3*

mutation. **(E)** Comparison between trained ML models and hematologists in recognizing mutant cells by native morphology.

Figure 2. ML models were "humanized" by adding just one pair of human samples to the bulk of mouse datasets. (A) BA of predicting mutation status in unseen human samples and controls by XGBoost models trained on mouse-only datasets with *KMT2A-MLLT3* mutation. **(B)** BA of predicting mutation status in unseen human samples and controls by XGBoost models trained on mouse datasets with *KMT2A-MLLT3* mutation plus one normal human donor sample and one *KMT2A-MLLT3*$^{+}$ AML sample. Note that one pair of random samples (FHCC-2 and AML-1 PDX-1) from Fig. 2A were used in training. **(C, D)** Sample-wise ROC curve and PR curve of different models when applied to predicting human cells in AML PDX samples. The sample-wise median values of each parameter were used in plotting. **(E)** Average probability of assigning "mutant" label to cells in each unseen human sample, by XGBoost models trained on mouse+human datasets with *KMT2A-MLLT3* mutation. **(F)** Performance in analyzing unseen human pseudo-samples with preset proportions of mutant cells by XGBoost models trained on mouse+human datasets with *KMT2A-MLLT3* mutation. Upper panel: XGBoost models used different confidence thresholds in calling a cell "mutant". Lower panel: XGBoost models used different random seeds during training and LOO cross-validation.

Figure 3. Explanation of ML model's working mechanism with feature analysis and application of the platform in detecting *Jak2*^{V617F}/*JAK2*^{V617F} mutation. (A) Top features with impact (SHAP value) on the output of XGBoost models trained on mouse-only datasets with *KMT2A-MLLT3* mutation. Features were extracted from merged three-channel cell images by CellProfiler. **(B)** Top features with impact (SHAP value) on the output of XGBoost models trained on mouse+human datasets with *KMT2A-MLLT3* mutation. Features were extracted from individual channels of cell images by scikit-image. BF: brightfield, SCC: side scatter, also known as darkfield / DF. **(C, D)** Performance comparison among XGBoost models that use different channel combinations and different feature extraction methods, when tested on different *Jak2*^{V617F} mouse datasets during LOO cross-validation. **(E)** Test of XGBoost models (trained on mouse and healthy human samples) in predicting mutant frequency in unseen human *JAK2*^{V617F} patient PB samples. Estimated % of mutant cells are determined by NGS results, assuming euploidy and absence of copy number variations.

Figure 1

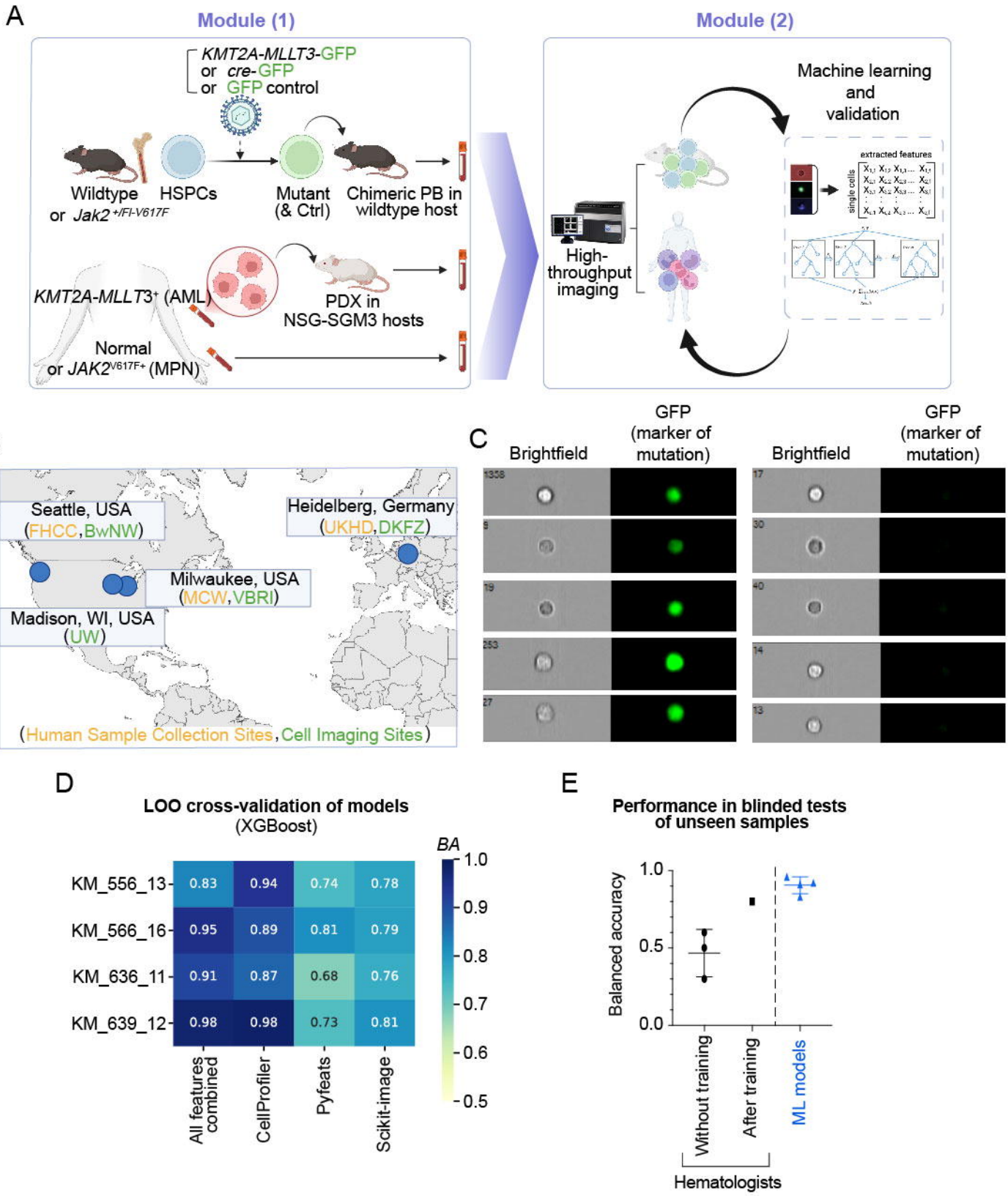

A
Module (1)
Module (2)
KMT2A-MLLT3-GFP
or cre-GFP
or GFP control
Wildtype
or Jak2 +/FI-V617F
HSPCs
Mutant
(& Ctrl)
Chimeric PB in
wildtype host
KMT2A-MLLT3+ (AML)
PDX in
NSG-SGM3 hosts
Normal
or JAK2V617F+ (MPN)
Machine learning
and
validation
extracted features
single cells
High-
throughput
imaging
B
Seattle, USA
(FHCC, BwNW)
Heidelberg, Germany
(UKHD, DKFZ)
Milwaukee, USA
(MCW, VBRI)
Madison, WI, USA
(UW)
(Human Sample Collection Sites, Cell Imaging Sites)
C
Brightfield
GFP
(marker of
mutation)
Brightfield
GFP
(marker of
mutation)
D
LOO cross-validation of models
(XGBoost)
BA
KM_556_13
KM_566_16
KM_636_11
KM_639_12
0.83 0.94 0.74 0.78
0.95 0.89 0.81 0.79
0.91 0.87 0.68 0.76
0.98 0.98 0.73 0.81
All features combined
CellProfiler
Pyfeats
Scikit-image
1.0
0.9
0.8
0.7
0.6
0.5
E
Performance in blinded tests
of unseen samples
Balanced accuracy
1.0
0.5
0.0
Without training
After training
ML models
Hematologists

Figure 2

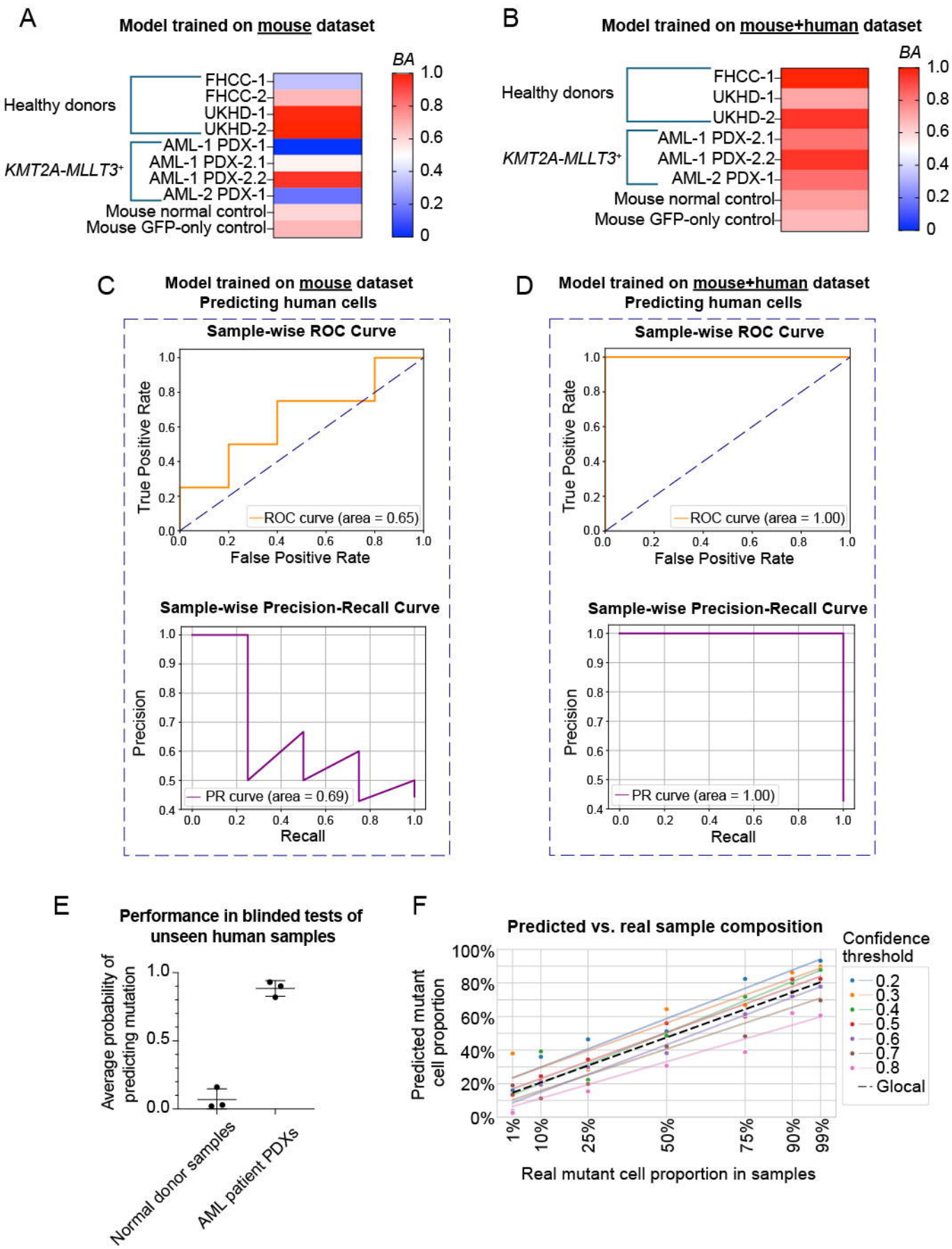
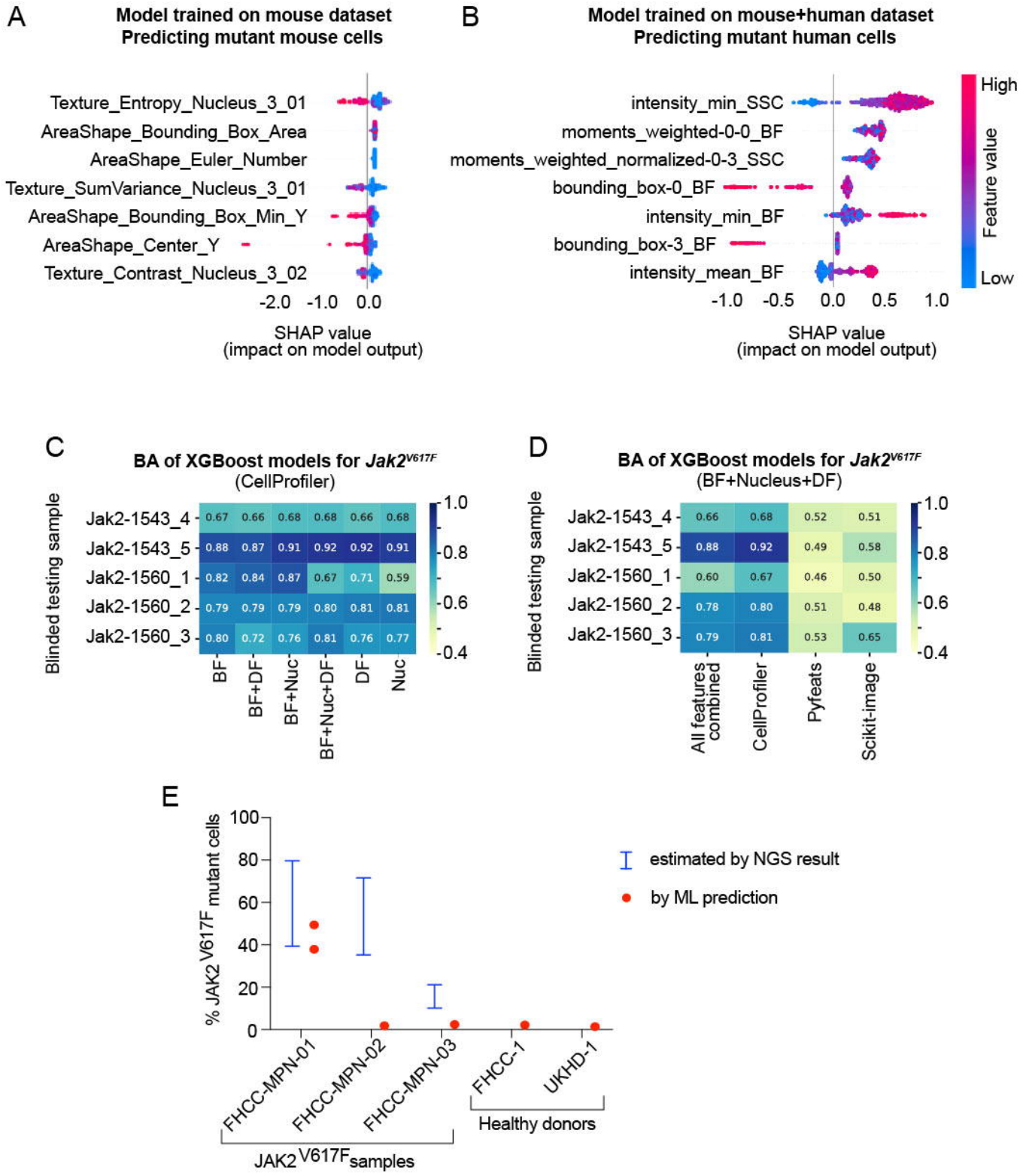

A
Model trained on mouse dataset
Healthy donors
KMT2A-MLLT3+
FHCC-1
FHCC-2
UKHD-1
UKHD-2
AML-1 PDX-1
AML-1 PDX-2.1
AML-1 PDX-2.2
AML-2 PDX-1
Mouse normal control
Mouse GFP-only control
BA
B
Model trained on mouse+human dataset
Healthy donors
KMT2A-MLLT3+
FHCC-1
UKHD-1
UKHD-2
AML-1 PDX-2.1
AML-1 PDX-2.2
AML-2 PDX-1
Mouse normal control
Mouse GFP-only control
BA
C
Model trained on mouse dataset
Predicting human cells
Sample-wise ROC Curve
True Positive Rate
False Positive Rate
ROC curve (area = 0.65)
Sample-wise Precision-Recall Curve
Precision
Recall
PR curve (area = 0.69)
D
Model trained on mouse+human dataset
Predicting human cells
Sample-wise ROC Curve
True Positive Rate
False Positive Rate
ROC curve (area = 1.00)
Sample-wise Precision-Recall Curve
Precision
Recall
PR curve (area = 1.00)
E
Performance in blinded tests of unseen human samples
Average probability of predicting mutation
Normal donor samples
AML patient PDXs
F
Predicted vs. real sample composition
Predicted mutant cell proportion
Real mutant cell proportion in samples
Confidence threshold
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Glocal

Figure 3


A
Model trained on mouse dataset
Predicting mutant mouse cells
Texture_Entropy_Nucleus_3_01
AreaShape_Bounding_Box_Area
AreaShape_Euler_Number
Texture_SumVariance_Nucleus_3_01
AreaShape_Bounding_Box_Min_Y
AreaShape_Center_Y
Texture_Contrast_Nucleus_3_02
-2.0 -1.0 0.0
SHAP value
(impact on model output)
B
Model trained on mouse+human dataset
Predicting mutant human cells
intensity_min_SSC
moments_weighted-0-0_BF
moments_weighted_normalized-0-3_SSC
bounding_box-0_BF
intensity_min_BF
bounding_box-3_BF
intensity_mean_BF
-1.0 -0.5 0.0 0.5 1.0
SHAP value
(impact on model output)
High
Feature value
Low
C
BA of XGBoost models for Jak2V617F
(CellProfiler)
Blinded testing sample
Jak2-1543_4 0.67 0.66 0.68 0.68 0.66 0.68
Jak2-1543_5 0.88 0.87 0.91 0.92 0.92 0.91
Jak2-1560_1 0.82 0.84 0.87 0.67 0.71 0.59
Jak2-1560_2 0.79 0.79 0.79 0.80 0.81 0.81
Jak2-1560_3 0.80 0.72 0.76 0.81 0.76 0.77
BF
BF+DF
BF+Nuc
BF+Nuc+DF
DF
Nuc
1.0
0.8
0.6
0.4
D
BA of XGBoost models for Jak2V617F
(BF+Nucleus+DF)
Blinded testing sample
Jak2-1543_4 0.66 0.68 0.52 0.51
Jak2-1543_5 0.88 0.92 0.49 0.58
Jak2-1560_1 0.60 0.67 0.46 0.50
Jak2-1560_2 0.78 0.80 0.51 0.48
Jak2-1560_3 0.79 0.81 0.53 0.65
All features combined
CellProfiler
Pyfeats
Scikit-image
1.0
0.8
0.6
0.4
E
% JAK2 V617F mutant cells
100
80
60
40
20
0
FHCC-MPN-01
FHCC-MPN-02
FHCC-MPN-03
FHCC-1
UKHD-1
Healthy donors
JAK2 V617F samples
estimated by NGS result
by ML prediction