

Pooled compound screening using DNA barcoding and Ghost Cytometry

SUMMARY

High content phenotypic screening is a powerful approach for identifying new genes and compounds that modulate cellular disease phenotypes. Array-based fluorescence imaging platforms are commonly used for phenotypic screening due to their high throughput capabilities. However, the approach has major limitations including well-to-well variability, resource-intensive experimental setup and image analysis, and preference for adherent fixed cells. Here, we introduce a new approach for pooled phenotypic screening by combining morphological profiling with DNA barcoding technology. We show application of the VisionSort platform, a new AI-driven cell sorter powered by Ghost Cytometry, combined with DNA barcoding for label-free high-throughput pooled phenotypic compound screening. VisionSort provides advantages over traditional screening approaches including i) scalability to large screening libraries, ii) applicability to a wide new range of morphological phenotypes, and iii) compatibility with commercial sequencing platforms for DNA barcode readouts.

INTRODUCTION

Although target-based screening has been extensively used in the pharmaceutical industry to identify new drug candidates, phenotypic screening is an emerging method to screen for first-in-class drugs. Phenotypic screening using microscopy-derived high resolution cell images is a common approach for identifying genes and compounds that induce cellular disease phenotypes. However, current microscopy-based phenotypic screening approaches are inflexible. Cell staining requires the use of multiple fluorescence markers with associated data and time intensive acquisition, storage, processing, and analysis steps required. The ideal pooled phenotypic screening platform is high-throughput, cost-effective, has minimal well-to-well batch effect, and can screen for a wide variety of cellular phenotypes.

Here we present a new approach to compound screening combining the VisionSort platform from Think-Cyte with DNA barcoding. VisionSort brings together fundamental advances in optics, microfluidics, and artificial intelligence (AI) to deliver morphological profiling and label-free cell sorting on top of the capabilities found in traditional fluorescence-only cytometers. Powered by Ghost Cytometry technology¹, VisionSort enables researchers to perform real-time, AI-driven, single cell sorting with or without labels. Flexible, easy-to-use, and user-controlled AI algorithms embedded directly in the instrument enable unbiased morphometric characterization and isolation of rare and unique cell populations.

METHODS

We conducted pooled compound screening to identify novel small molecules for the treatment of nonalcoholic steatohepatitis (NASH) using a compound library. NASH is a progressive and severe liver disease characterized by lipid accumulation, inflammation, and downstream fibrosis². We performed two types of screening assays to assess compound effects on lipid accumulation in cells, a common phenotype used for screening new drugs for NASH. The first was the conventional approach using BODIPY fluorescence intensity to detect lipid accumulation in cells by microscopy. The second used label-free cell morphological information acquired from VisionSort (Figure 1). By comparing the results of the two approaches, we demonstrate the utility of label-free screening using VisionSort to identify drug candidates with novel mechanism-of-actions.

RESULTS

Arrayed compound screening of high-content phenotypes by conventional fluorescence

Using the fluorescence-based arrayed platform, lipid accumulation was statistically detected by BODIPY intensity after induction of NASH phenotype in HepG2 cells. NASH-induced HepG2 cells were fixed and stained with BODIPY 493/503 and Hoechst 33342. Stained cells were imaged and analyzed using IN Cell analyzer 6000. Analysis was performed by "BODIPY staining area per cell (BODIPY/Nuc evaluation)" and compounds were evaluated as 100% inhibition of the stained image of HepG2 without NASH induction.

NASH-induced HepG2 cells showed a clear enhancement of BODIPY 493/503 fluorescence (Figure 2A). Cells treated with GS-0976 (Acetyl-CoA Carboxylase inhibitor), a reported hepatic steatosis reducer, showed inhibition of BODIPY 493/503 fluorescence. Under this condition, 960 compounds were added to cells at 10 μ M and

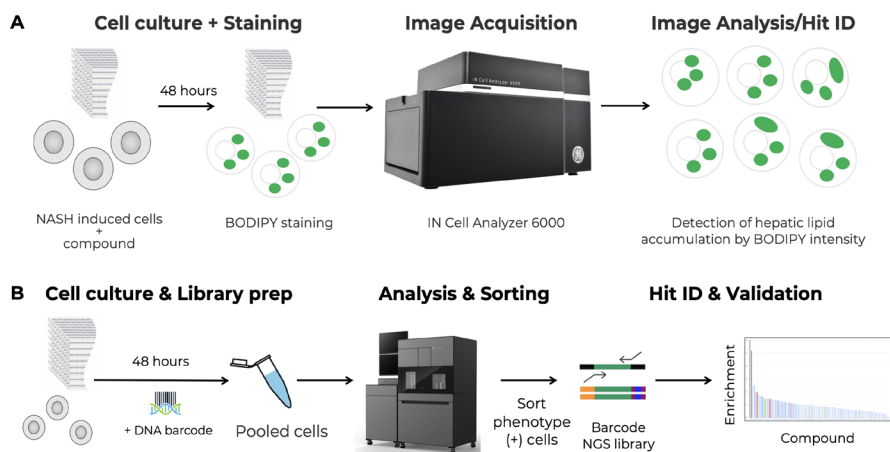


Figure 1. Workflows for high content phenotypic screening. (A) Workflow of a conventional fluorescent-based arrayed platform. NASH-induced cells are cultured with compounds in each well and stained with BODIPY fluorescent dye. After image acquisition, analysis is performed based on BODIPY intensity for detection of hepatic lipid accumulation. (B) Workflow of a novel label-free pooled platform using VisionSort. NASH induced cells are cultured with compounds in each well. Before being run on VisionSort, DNA barcodes corresponding to each well are added. Cells exhibiting the target phenotype are sorted based on pre-trained machine learning (ML) model and the DNA barcodes are amplified and sequenced to identify enriched DNA barcodes.

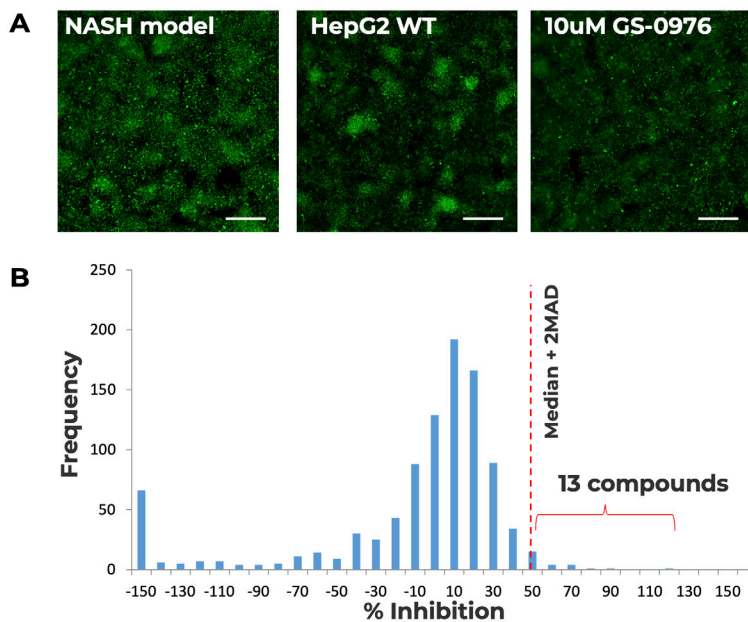


Figure 2. Conventional high-content phenotypic screening by fluorescence. (A) BODIPY-stained images of NASH-induced cells. Scale bar is 300 μ m. (B) Histograms of inhibitory activity of compounds tested by conventional fluorescence screening. Median: 1.6, MAD (median absolute deviation): 22.85, Median + 2MAD: 47.28.

evaluated for their effects on NASH-induced intracellular lipid accumulation. The screening results showed that obtained data was normally distributed, with 13 compounds showing inhibitory activity above Median + 2MAD (Median Absolute Deviation). Inhibitory activity was above 47.28% (Figure 2B).

Pooled compound screening of high-content phenotypes with VisionSort

For pooled screening using VisionSort, we first tested the ability to classify phenotypic changes between control and NASH phenotypes. HepG2 cells appeared grossly similar by morphology when evaluated by brightfield microscopy and forward and side scatter analysis using conventional FACS (Figure 3A and B). Next we trained a ML classifier to recognize these phenotypes label-free on VisionSort. To train the classifier, we separately prepared NASH-induced and uninduced HepG2 cells and combined them in equal proportions to use as the 'ground truth' training sample. We used morphological profiles from 2,000 individual cells from the ground truth sample to train a support vector machine (SVM) classifier. When the classifier model was evaluated on a test dataset of 1,000 cells each, the model showed excellent classification performance with clearly separable phenotypic populations and a high area under the ROC curve (AUC) score of 0.93 (Figure 3C).

Next, we performed compound screening with DNA barcoded cells label-free (no fluorescence stains or labels were applied to the cells). The DNA barcoding method is described previously³ and it was modified here for use in pooled screening. NASH induced cells were cultured with 960 compounds in each well. Before analysis and sorting on VisionSort, DNA barcodes corresponding to each well

were added. Cells exhibiting inhibition of the NASH phenotype (lipid accumulation) were sorted based on the pre-trained ML model. About 6 million target cells were successfully analyzed and sorted in 30 mins and processed for deep sequencing to identify hits. Analysis of the deep sequencing data revealed 32 compounds with the potential to suppress the NASH phenotype (Figure 3D).

Assessment of candidate hit compounds

Interestingly, only 2 of the 960 compounds were common hits in both conventional fluorescence screening and screening with VisionSort. One of the common hits was hispidin, a fungal polyphenol that has been reported to inhibit glycerol-3-phosphate dehydrogenase, lowering intracellular triglyceride levels in a concentration-dependent manner and increasing intracellular cAMP levels, thereby promoting lipid metabolism⁴. The other common hit compound detected by both screening approaches was dioxybenzone, a UV absorbing molecule used in sunscreen. Metabolites of dioxybenzone have been reported to act on estrogen receptor α , whose agonism is thought to increase the expression of lipid metabolism genes⁵.

By conventional fluorescence screening, compounds that increase intracellular cAMP levels, such as β -adrenergic receptor agonists and PDE4 inhibitors, were detected as unique hits. This suggests that fluorescence screening identified compounds that act to promote lipid metabolism and attenuate fluorescence intensity. Additionally, compounds showing anticancer activity, such as 10074-G5 and SU11274, were also detected as hit compounds^{6,7}. However, since it has been reported that cytotoxic compounds can appear as false positives using fluorescence screening, further validation of these results are needed.

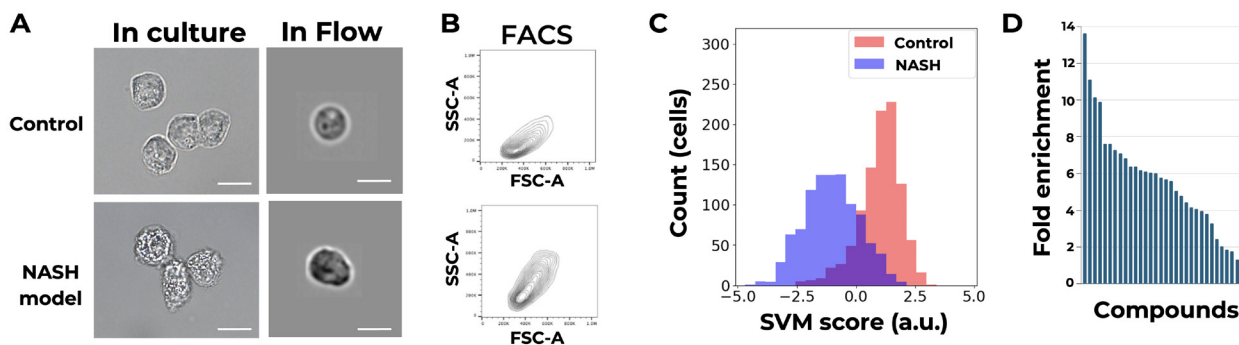


Figure 3. High content pooled phenotypic screening using VisionSort. (A) Bright field images of NASH induced and uninduced control HepG2 cells in culture and analyzed by conventional imaging flow cytometry. Scale bars are 10 μ m. (B) Conventional fluorescence-activated cell sorting (FACS) scatter plots of forward scatter (FSC) and side scatter (SSC). (C) Performance results for the ML-derived NASH phenotype classifier on VisionSort. The ML-based classification performance is displayed as histograms of SVM scores. Red and blue colors indicate ground truth labels. The classifier showed excellent performance with an AUC of 0.93 (D) Hit selection. Individual compounds used in the screen plotted against the value of fold enrichment from barcode read counts before and after sorting (Score > 1).

By screening with VisionSort, several unique hits were identified. Baccalin is reported to inhibit NASH in vitro and in vivo models, and inhibit lipid peroxidation through lipoxygenase inhibition^{8,9}. Vanillylacetone, is reported to have an ameliorative effect on a rat nonalcoholic fatty liver model¹⁰, suggesting that the label-free screening may detect compounds that broadly inhibit the pathology of NASH.

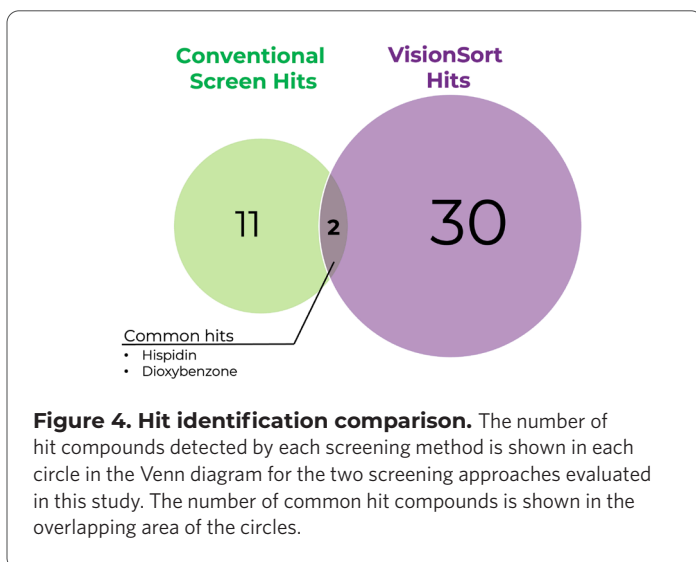
CONCLUSION

Here we present a novel high throughput compound screening method and evaluated the approach using a complex morphological phenotype. Compared to the results of using conventional fluorescence-based high content screening, we found differences in potential candidate compounds that inhibit the NASH phenotype. While each approach identified several hits, the number of hits detected by both approaches was minimal. Phenotypic screening using cell morphology as a readout has the potential to identify new compounds and novel pathways that could not be detected using conventional methods. Thus, the approach described here can be applied to screen for a wider range of complex phenotypes, giving researchers a new powerful tool for drug discovery.

MATERIALS AND METHODS

In vitro NASH induction from HepG2 cells

For induction preparation, HepG2 cells (American Type Culture Collection) were dissociated with 0.05 x Trypsin-EDTA solution, and then 3.5×10^6 cells were seeded in each well of 96-well plate with D-MEM High Glucose (Wako) and 10% FBS (Hyclone), 1% Penicillin-Streptomycin Solution (Wako). NASH induction was performed 16 hours after seeding. Cells were cultured for 48 hours in D-MEM Low Glucose (Wako) and 10% FBS, 1% Penicillin-Streptomycin Solution and supplemented with 65 μ M Oleic acid (Sigma), 45 μ M Palmitic acid (Sigma), 100 nM Insulin (Wako), 50 ng/ml TNF- α (Wako), 25 ng/mL IL-1 β (Wako) and 8 ng/mL TGF- β (Wako).



Compounds

960 compounds tested in both the BODIPY fluorescence intensity assay and label-free cell morphological assay were selected based on their biological activity reported in the public database (e.g. ChEMBL Database).

Sample preparation for fluorescence high content screening

Cells were fixed in 384-multiwell plates with 4% (w/v) paraformaldehyde in PBS for 15 minutes. Cells were stained with BODIPYTM 493/503 (Thermo Fisher) and Hoechst 33342 (Dojindo) for 30 minutes at room temperature.

Data acquisition and analysis for fluorescence high content screening

Fluorescence images were taken using IN Cell Analyzer 6000 (Molecular Devices) and analyzed using IN Cell developer Toolbox 1.9.2 software. The intracellular lipid load was calculated as BODIPYTM 493/503 lipid dye area and intensity. "BODIPY/Nuc evaluation" parameter was calculated as ratio of BODIPY staining area and the number of nuclei, which was quantified using Hoechst 33342. The inhibitory activity was evaluated by comparing the "BODIPY/Nuc evaluation" of vehicle-treated wells and compound-treated wells.

DNA Barcode tagging for label-free high content screening

DNA barcode tagging protocol was modified from a previous study³. For the 96-well plate experiment, cells were washed with D-PBS (Wako) supplemented with 0.5 mM EDTA solution (Nacalai) twice. Cells were treated with 90 μ l of 200 nM Cho-Anchor (5'-5'ChoI TEG - GTAACGATCCAGCTGTCACTTGAATTCTCGGGTGCCAAGG -3') dissolved in 0.05 w/v% Trypsin-EDTA solution (Wako) and incubated at 37°C for 5 min. Then, 10 μ l of 2 μ M Cho-co-Anchor (5'-AGTGACAGCTGGATCGTTAC- 3'ChoI TEG -3') dissolved in 0.05 w/v% trypsin-EDTA solution was added and incubated at 37°C for 5 min. In addition, 100 μ M of DNA barcode was added and incubated at 37°C for 5 min. See more information TechNote.

Classification and sorting for label-free high content screening

For the train sample, HepG2 control and NASH induced cells were individually prepared as previously described. NASH model cells were washed with D-PBS and treated with 0.05 w/v% Trypsin-EDTA solution for 5 min at 37°C to detach them from the culture dish. As a ground truth label, only control cells were stained with CellTracker Green CMFDA Dye (Thermo Fisher). For the classification of HepG2 control and NASH induced cells, an equal concentration of each cell type was mixed, and this mixture was passed through the Ghost Cytometry technology. FSC/SSC scatter plot gating was employed to

eliminate doublets and debris from the training dataset. 2,000 cells from each cell were randomly selected for training, with an additional 1,000 cells for testing. Using fsGMI and bsGMI, a trained model was created and an AUC score of 0.93 was obtained.

For screening, NASH induced HepG2 cells treated with each compound (final conc. 10 μ M) in well were washed twice with PBS and treated with 0.05 w/v% Trypsin-EDTA solution for 5 min at 37°C to detach them from the culture plates. Cells in each well were mixed and fixed with 4% formaldehyde in PBS for 20 min at room temperature. For sorting the NASH-induced, compound-treated, and then pooled cells, 6,265,695 cells were processed by the sorter, and approximately 402,022 cells predicted as positive were sorted within 30 mins using a classifier trained and implemented on a Field Programmable Gate Array (FPGA). To enhance purity, a threshold of SVM scores >0.5 was set for cell sorting. In cases where the percentage of target cells in the pre-sorting sample was 6.4%, the predicted purity was 89.5%, and the predicted recovery was 74.3% in the sorted sample. These experiments were performed as two independent replicates.

Quantification of DNA barcode enrichment for label-free high content screening

For amplicon sequencing for DNA barcodes, genomic DNA from fixed cells was extracted using the QIAamp DNA FFPE Tissue Kit (QIAGEN) or the hotshot method¹¹. For each sample, target regions were amplified using the extracted genomic DNA as the PCR template with the corresponding first HTS primer pair (see more information in the associated Tech Note). The PCR was conducted following the previously established protocol¹². The PCR product was purified using the FastGene Gel/PCR Extraction Kit (Nippon Genetics) and re-amplified with custom Illumina index primers. Each indexed library was electrophoresed in a 2% agarose gel, and the expected band was purified using the FastGene Gel/PCR Extraction Kit (Nippon Genetics). The sequencing libraries were quantified using qPCR with the KAPA Library Quantification Kit Illumina (KAPA Biosystems) for multiplexing. The multiplexed libraries were quantified using the same qPCR protocol and sequenced with 20% PhiX control using Illumina Miseq (Miseq v3 kit). For quantitative analysis of DNA barcodes, as modified from previous study¹³.

REFERENCES

- Ota, S., Horisaki, R., Kawamura, Y., Ugawa, M., Sato, I., Hashimoto, K., Kamesawa, R., Setoyama, K., Yamaguchi, S., Fujiu, K. and Waki, K., (2018). Ghost cytometry. *Science*, 360(6394), 1246-1251.
- Friedman, S.L., Neuschwander-Tetri, B.A., Rinella, M., Sanyal, A.J., (2018). Mechanisms of NAFLD development and therapeutic strategies. *Nature Medicine*, 24(7), 908-922.
- McGinnis, C.S., Patterson, D.M, Winkler, J., Conrad, D.N., Hein, M.Y., Srivastava, V., Hu, J.L., Murrow, L.M., Weissman, J.S., Werb, Z, Chow, E.D., and Gartner, Z.J. (2019) MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nature Methods*, 16(7), 619-626.
- Palkina, K.A., Ipatova, D.A., Shakhova, E.S., Balakireva, A.V., and Markina, N.M. (2012). Therapeutic potential of Hispidin - Fungal and plant polyketide. *Journal of Fungi*, 7(5), 323.
- Zhan, T., Zhang, L., Cui, S., Lui, W., Zhou, R., and Zhuang, S. (2021). Dioxibenzene triggers enhanced estrogenic effect via metabolic activation: in silico, in vitro and in vivo investigation. *Environmental pollution*, 268(Pt B), 115766.
- Clausen, D.M., Guo J., Parise, R.A., Beumer, J.H., Egorin, M.J., Lazo, J.S., Prochownik, E.V., and Eiseman, J.L. (2010). In vitro cytotoxicity and in vivo efficacy, pharmacokinetics, and metabolism of 10074-G5, a novel small-molecule inhibitor of c-Myc/Max dimerization. *The Journal of pharmacology and experimental therapeutics*. 335(3), 715-727.
- Ma, P.C., Jagadeeswaran, R., Jagadeesh, S., Tretiakova, M.S., Nallasura, V., Fox, E.A., Hansen, M., Schaefer, E., Naoki, K., Lader, A., Richards, W., Sugarbaker, D., Husain, A.N., Christensen, J.G., and Salgia, R. (2005). Functional expression and mutations of c-Met and its therapeutic inhibition with SU11274 and small interfering RNA in non-small cell lung cancer. *Cancer Research*, 65(4), 1479-88.
- Shi, H., Qiao, F., Lu, W., Huang, K., Wen, Y., Ye, L., and Chen, Y. Baicalin improved hepatic injury of NASH by regulating NRF2/HO-1/NRLP3 pathway. (2022). *European Journal of Pharmacology*, 934, 175270.
- Tsurusaki, S., Tsuchiya, Y., Koumura, T., Nakasone, M., Sakamoto, T., Matsuoka, M., Imai, H., Kok, C.Y., Okochi, H., Nakano, H., Miyajima, A., and Tanaka, M. (2019) Hepatic ferroptosis plays an important role as the trigger for initiating inflammation in nonalcoholic steatohepatitis. *Cell Death and Disease*, 10(6), 449.
- Shahid, R., lahtisham-UI-Haq, Mahnoor, Awan, K.A., Iqbal, M.J., Munir, H., and Saeed, I. Diet and lifestyle modifications for effective management of polycystic ovarian syndrome (PCOS). (2022). *Journal of Food Biochemistry*, 46(7), e14117.
- Jaitin, D.A., Weiner, A., Yofe, I., Lara-Astiaso, D., Keren-Shaul, H., David, E., Salame, T.M., Tanay, A., van Oudenaarden, A., and Amit, I. (2016). Dissecting Immune Circuits by Linking CRISPR-Pooled Screens with Single-Cell RNA-Seq. *Cell*, 167(7), 1883-1896. e15.
- Sakata, R.C., Ishiguro, S., Mori, H., Tanaka, M., Tatsuno, K., Ueda, H., Yamamoto, S., Seki, M., Masuyama, N., Nishida, K., Nishimasu, H., Arakawa, K., Kondo, A., Nureki, O., Tomita, M., Aburatani, H., and Yachie, N. (2020). Base editors for simultaneous introduction of C-to-T and A-to-G mutations. *Nature Biotechnology*, 38(7), 865-869.
- Tsubouchi, A., An, Y., Kawamura, Y., Yanagihashi, Y., Nakayama, H., Murata, Y., Teranishi, K., Ishiguro, S., Abratani, H., Yachie, N. and Ota, S. (2024). Pooled CRISPR screening of high-content cellular phenotypes using ghost cytometry. *Cell Reports Methods*, 4(3), 100737.



THINKCYTE INC.

UNITED STATES
1100 Island Drive, STE 203
Redwood City, CA 94065

JAPAN
7-3-1 Hongo,
Bunkyo, Tokyo

CONTACT@THINKCYTE.COM WWW.THINKCYTE.COM

FOR RESEARCH USE ONLY. © Copyright 2024. All rights reserved. ThinkCyte, VisionSort and Ghost Cytometry are trademarks of ThinkCyte K.K. All other trademarks are the property of their respective owners. AN-017.24.1-US
Unauthorized posting of this material in the public domain is prohibited.