Published in final edited form as:

Nat Methods. 2009 October ; 6(10): 759. doi:10.1038/nmeth.1375.

On an Approach for Extensibly Profiling the Molecular States of Cellular Subpopulations

Lit-Hsin Loo, Hai-Jui Lin, Robert J. Steininger III, Yanqin Wang, Lani F. Wu, and Steven J. Altschuler*

Green Center for Systems Biology and Department of Pharmacology, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., ND 9.214, Dallas, Texas 75390, USA.

Abstract

Microscopy often reveals the existence of phenotypically distinct cellular subpopulations. However, further characterization of observed subpopulations can be limited by the number of biomolecular markers that can be simultaneously monitored. Here, we present a computational approach for extensibly profiling cellular subpopulations by freeing up one or more imaging channels to monitor additional probes. In our approach, we train classifiers to re-identify subpopulations accurately based on an enhanced collection of phenotypic features extracted from only a subset of the original markers. Subpopulation profiles were then constructed step-wise from replicate experiments, in which cells were labeled with different but overlapping marker sets. We applied our approach to characterize molecular differences among subpopulations, and functional groupings of markers within populations of differentiating mouse preadipocytes, polarizing human neutrophil-like cells, and dividing human cancer cells.

INTRODUCTION

Visual examinations of cellular populations using microscopes often reveal striking phenotypic heterogeneity, with subpopulations distinguishable by differences in expression or subcellular localization of biomolecules, or in cell morphology¹⁻³. An increasing body of evidence suggests that characterization of cellular subpopulations may provide important insights into biological processes such as disease progression, cellular differentiation, or response to external perturbations²⁻⁵. To better understand the molecular underpinnings of observed cellular heterogeneity, methods are required for probing and comparing large numbers of biomolecular readouts on cellular subpopulations.

Two major obstacles have limited our ability to investigate the molecular states of cellular subpopulations. First, it can be difficult to physically and reliably isolate subpopulations that are identified according to spatially-defined phenotypic patterns such as cellular morphology or biomolecular localization. Second, conventional filter-based fluorescence microscopy is currently limited in practice to four to five simultaneous markers. This limitation is an obstacle for identifying molecular differences among subpopulations; even simple transcriptional networks may consist of more than four regulators and tens of target biomolecules⁶. Regardless of new labeling technologies^{7, 8} and computational algorithms⁹ that increase the number of simultaneous fluorescence markers, a need will exist to monitor ever larger numbers of biomolecules.

*corresponding author = steven.altschuler@utsouthwestern.edu.

AUTHOR CONTRIBUTIONS L.H.L. designed the profiling methods and performed the analysis. L.H.L. and H.J.L. performed the 3T3-L1 experiments. R.J.S. performed the H460 experiments. Y.W. performed the HL-60 experiments. L.H.L., L.F.W., and S.J.A. contributed to the conception of the overall approach, statistical analysis of the methods and writing of the manuscript.

Here, we propose a computational method for building extensible phenotypic profiles of cellular subpopulations based on immunofluorescence microscopy. We started by assuming that subpopulations had been identified through specific features of an initial biomolecular marker set and that no free fluorescence channels were readily available for monitoring additional biomolecules. In order to monitor additional biomolecules in these subpopulations, our strategy was to: 1) select one or more of the initial biomolecular markers to be dropped, thereby freeing one or more of the fluorescence channels; 2) train a classifier to re-identify the subpopulations using the retained marker(s) by making better use of available information (e.g. bright field images or expanded “high-content” features¹⁰⁻¹²); and 3) analyze replicate experiments co-stained with the retained markers, and one or more new markers in place of the dropped markers. In this process, while no individual cell had been simultaneously labeled with all markers, phenotypic information on the subpopulation was combined *in silico* via a common set of retained markers. This approach, which we refer to as “Virtual Subpopulation Profiling,” provided an extensible strategy for obtaining deeper phenotypic and functional characterization of the identified subpopulations.

In the following, we first provide an overview of our method. Then, we demonstrate its application to studying three different biological processes: adipogenesis of 3T3-L1 mouse preadipocytes, polarization of HL-60 human neutrophil-like leukemia cells, and cell cycle of H460 human lung cancer cells. We selected these systems because phenotypic heterogeneity has been observed^{1, 2}, markers for many of the involved molecular components are readily available^{13, 14}, and dramatic changes in cellular morphology and/or localization of biomolecules occur, enabling separation of subpopulations *in silico*. For each model system, we identified cellular subpopulations with distinct phenotypes, trained classifiers to re-identify the subpopulations using a reduced set of markers, and constructed Virtual Phenotypic Profiles for these subpopulations based on seven to nine additional biomolecular markers. We experimentally validated that Virtual Phenotypic Profiles constructed from our approach closely resembled profiles obtained from actual co-staining.

RESULTS

Overview of the subpopulation profiling method

Our approach consisted of three steps (Fig. 1). In the first step, we identified a collection of phenotypically distinct subpopulations from microscopy images of cells stained with an initial marker set (Fig. 1a). Manual, computer-assisted, or fully automated^{5, 15} methods can be used to group cells based on similarities of specific phenotypic features. The choice of the initial feature set and thus the method for identifying subpopulations is study-dependent. Subsequent steps in our approach will provide an assessment of whether phenotypic distinctions observed among subpopulations in the initial feature set are present for features defined using other markers.

In the second step, we trained a classifier to distinguish the identified subpopulations using only a “reference” feature set. A natural choice of reference feature set would be the initial feature set. However, one or more of the initial markers will be dropped to make “room” for new markers (Fig. 1b). Ideally, the marker(s) that provides the least discriminative feature would be dropped. We reasoned that the inclusion of additional phenotypic readouts from the retained markers and/or from bright-field images could make up part of the information lost from the dropped markers (Fig. 1b). For this purpose, we chose to first make use of a large set of general phenotypic features, and then automatically select features that provided the highest classification accuracy.

In the third step, we performed replicate experiments in which each set of cells was co-stained with a different new marker together with the retained markers (Fig. 1c). We then used the

trained classifier to match cells to the corresponding subpopulations based on the reference features. For each subpopulation, we extracted and combined phenotypes from all new markers to form a Virtual Phenotypic Profile. Our approach provided the ability to summarize, for each subpopulation, the simultaneous phenotypes of larger numbers of markers than could be accommodated with current technology.

Building Virtual Phenotypic Profiles for adipogenesis

We first applied our method to monitor differentiating 3T3-L1 preadipocytes. Previously, we had used a Gaussian mixture model (GMM) to approximate the heterogeneity observed among differentiating 3T3-L1 cells², and identify four subpopulations based on the average cellular levels of lipid droplets (LD) or adiponectin (AdipoQ) (Fig. 2a,b² Supplementary Methods; see also Fig. 2c in Loo *et. al.*²).

To make room for new markers, we trained a classifier to recognize the identified subpopulations based on a reduced marker set. We used a conservative estimate of classification performance called minimum class accuracy (MCA), which reports the classification performance of the most poorly-classified subpopulations (Supplementary Methods). Classification performance depended primarily on four factors: the retained marker set, the classifier type, the classifier parameters, and the reference features. An exhaustive search for a globally optimum combination was computationally prohibitive. Instead, we used a greedy search approach that optimized one factor at a time (except for classifier parameters, which were optimized during each step of the greedy search).

First, in this data set, we found that dropping any of the initial markers significantly reduced subpopulation classification accuracy (Fig. 2c, and Supplementary Fig. 1a), indicating that the initial features derived from these markers were not redundant. Between the two initial markers, dropping the LD marker yielded a smaller decrease in classification accuracy (from 95.3% to 32.6% MCA). To compensate for the information loss, we dramatically extended our initial feature set to include 499 high-content features per cell, extracted from images of cells stained with AdipoQ and with DNA markers, and from bright-field images (Fig. 2d; Supplementary Methods). These general features were not biased towards any particular biological process, and formed the basis for our reference feature set.

Second, we searched for a classifier type that would provide high accuracy re-identification of the original subpopulations. We selected 10 standard classifier types and crudely considered all 15 possible combinations of features categories (intensity, morphology, texture, and moment; feature sets were refined further in the next step). We found that most top-performing classifiers had similar relative performances across the 15 tested feature combinations (Supplementary Fig. 1b). We selected the classifier type with the highest overall MCA, namely support vector machines with radial-basis-function kernel (SVM-RBF)¹⁶.

Third, given our choice of classifier type, we searched for a subset of features within the reference feature set that would give improved classification performance. We compared several feature-selection methods (Supplementary Fig. 2a,b). The best performance was given by the sequential floating forward search (SFFS)¹⁷. This approach identified a subset of 7 high-content features from AdipoQ, DNA, and bright field images (Supplementary Data) that increased subpopulation classification accuracy to 74.2% MCA (Fig. 2d, and Supplementary Fig. 2a). This result indicated that, in some cases, the inclusion of bright-field and DNA features can increase the power of subpopulation discrimination. Thus, for subsequent identification of the original four subpopulations, we trained a final SVM-RBF classifier using these 7 informative features.

Finally, we combined phenotypic measurements from replicate experiments to create Virtual Phenotypic Profiles of subpopulations. Cells were stained with the three initial markers (DNA, LD, and AdipoQ) together with one of the five new markers selected to monitor the adipogenesis and lipolysis processes (Fig. 3a). We chose to extract features measuring the nuclear or cellular levels for these markers, based on their known subcellular-localization patterns. The Virtual Phenotypic Profiles were constructed using the median values of these features for each subpopulation.

Validation of Virtual Phenotypic Profiling

We next compared the constructed Virtual Phenotypic Profiles both to “ground truth” (defined below) and to population-averaged profiles. The first comparison allowed us to measure noise introduced by our trained classifier. The second comparison allowed us to test whether novel information was contained in the subpopulation-level profiles.

To assess deviation from ground truth, we computed phenotypic profiles using three different definitions of cellular subpopulations: the original Gaussian-mixture-model based on all initial features (INT/ALL; the “ground truth”); a trained classifier based on the initial adiponectin feature alone (INT/AdipoQ; used to assess performance before the addition of new features); or a trained classifier based on the high-content adiponectin feature set (HC/SFFS; the Virtual Phenotypic Profile) (Fig. 3a). Deviation from ground truth was measured as $|10 \log_{10}(\text{Virtual Profile}/\text{ground truth})|$, traditionally assigned in units of decibels, or dBs. As expected, the phenotypic profiles for the INT/AdipoQ subpopulations deviated from ground truth (Fig. 3a); the maximum noise level of the median marker levels was 8.8dB (Fig. 3b). In contrast, the HC/SFFS subpopulations had similar single-cell statistics to the ground truth (Fig. 3a); the maximum noise level was only 1.7dB (Fig. 3b). Overall, HC/SFFS subpopulations had either significantly lower or insignificantly different noise levels than the corresponding INT/AdipoQ subpopulations for all markers ($P < 0.05$; Fig. 3c). These results showed that the high-content features from AdipoQ could compensate for the dropped LD features.

To measure the deviation of Virtual Phenotypic Profiles from population averages, we repeatedly partitioned cells into four “random” subpopulations using the same subpopulation percentages as the original S1 to S4 (Fig. 2b). As expected, profiles based on the new markers for these random subpopulations were distributed around the population-averaged profile (Fig. 3d). In contrast, Virtual Phenotypic Profiles for the original subpopulations remained distinct from each other, and profiles for S1, S2 and S3 differed significantly from the population-averaged profile ($P < 0.05$, Fig. 3d; see Supplementary Method for P -value estimation). In general, Virtual Phenotypic Profiles should only be expected to diversify and offer higher-resolution information than a population-averaged profile to the extent that the original subpopulations diversify with respect to the initial markers, and that the new markers measure molecular differences among these original subpopulations.

Functional Grouping of Biomolecules

We next used Virtual Phenotypic Profiles to compare biomolecules based on their expression patterns across the subpopulations (Fig. 4). Hierarchical clustering based on the original subpopulations found two main clusters that were consistent with current understanding of these pathways^{13, 18} (Fig. 4a; see also Fig. 4a in Loo *et al.*²), but that were missed by the population-averaged profile (Supplementary Fig. 3). After dropping the LD marker, the heatmap diverged from the original (Fig. 4b). However, as measured by the lower noise level (Fig. 3b), the heatmap obtained by using the additional high-content features was similar to the original heatmap (Fig. 4c). These results indicated that the high-content features gave similar functional grouping of biomolecules as the initial features.

Virtual Phenotypic Profiles for polarization

We next applied our method to differentiated HL-60 neutrophil-like leukemia cells^{14, 19} that were stimulated for 2 minutes with formyl-met-leu-phe (fMLP). Three initial markers were used: F-actin to detect the protruding front of the cells, phosphorylated myosin light chain 2 (pMLC2) to detect the retracting rear, and DNA to detect nuclear regions (Supplementary Fig. 4a). GMM clustering identified three subpopulations corresponding to different stages or states of polarization (Fig. 5a, and Supplementary Fig. 4b).

To free an imaging channel, we dropped the pMLC2 marker, which caused the classification accuracy to decrease from 94.3% to 48.9% MCA (Supplementary Fig. 1c, and 5a). We then identified 8 high-content features from the F-actin and DNA staining to increase the classification accuracy back to 72.9% MCA (Supplementary Fig. 5b). We also added five new markers germane to the polarization process, and computed Virtual Phenotypic Profiles based on two polarization-specific features: the “polarization index,” which captured the distance between a marker and cell centroids; and “F-actin co-localization,” which captured the spatial correlation between a marker and F-actin (Fig. 5b). The maximum noise level was 0.3dB on all subpopulations, and the profiles for S1 and S2 significantly diverged from the population-averages (Supplementary Fig. 6a-c).

Hierarchical clustering of the Virtual Phenotypic Profiles for the polarized HL-60 cells grouped the markers by known functional associations^{14, 20} (Fig. 5b). Interestingly, similar functional grouping could also be obtained from the Virtual Phenotypic Profile of subpopulation S3 alone (Fig. 5c). We obtained similar results by using the actual subpopulations identified by all the initial features (Supplementary Fig. 7a,b).

Virtual Phenotypic Profiles for cell cycle

Finally, we applied our method to H460 lung cancer cells. Here, we used prior knowledge to separate cells into G1, S, G2 and M phases in cell cycle (Fig. 5d, and Supplementary Fig. 8). We co-stained cells with 5-bromo-2-deoxyuridine (BrdU) to detect S-phase cells²¹, phosphorylated histone H3 (pH3) to detect M-phase cells²², and DAPI (for DNA) to separate G1- from G2-phase cells (Supplementary Methods). These three initial markers are commonly used as a “gold standard” for cell-cycle detection²³.

Here, instead of systematically identifying and dropping the least informative marker(s), we chose to drop a marker based on practical considerations. In this case, BrdU labeling requires a pulse-and-chase experiment; eliminating this marker reduced the complexity and cost of the experiment. Although total DNA level alone has been used for separating G1, S, and G2/M cells²³, it is well-known that S and non-S-phase cells can have similar total DNA levels²⁴ (Supplementary Fig. 8c). Indeed, after dropping the BrdU marker, classification accuracy dropped from 95.4% to 52.1% MCA (Supplementary Fig. 5c). As with the two previous systems, we found that SVM-RBF classifier and SFFS feature selection provided the best overall classification performances. However, in this case, the addition of 10 high-content features increased the classification accuracy only back to 60.3% MCA (though the average class accuracy was 74.9%) (Supplementary Fig. 5d). Most of the misclassified cells were in the S and G2 phases.

To profile the subpopulations, we monitored six additional markers with differential expressions during cell cycle (Fig. 5e). The maximum noise levels of the Virtual Phenotypic Profiles was 1.4dB, and, as expected, profiles (especially for S1, S3, and S4) significantly diverged from population-averages (Supplementary Fig. 6d-f). Here, hierarchical clustering of the Virtual Phenotypic Profiles separated the added markers into those that peaked during M^{22, 25, 26}, G2^{27, 28}, and S²¹ phases (Fig. 5e). We also computed an additional feature, “DNA

co-localization,” which captured the spatial correlation between a marker and DNA (Supplementary Methods). Despite some misclassification of S/G2 cells, we found that our profiles (Fig. 5f) conformed to the known subcellular translocation of these markers during cell-cycle progression^{26, 27}, and looked similar to the profiles of the actual subpopulations (Supplementary Fig. 7c,d).

DISCUSSION

We demonstrated a method for constructing extensible phenotypic profiles of cellular subpopulations. The approach is based on the ability to re-identify subpopulations in replicate experiments using a reduced marker set but increased number of phenotypic readouts. The approach of substituting features for markers is most effective when information from the dropped marker(s) can be extracted from properties of cellular morphology or from the remaining markers, such as when the original markers represent readouts of a common biological pathway.

The degree of heterogeneity present within a population affects the usefulness of our approach. In principle, if there were only one actual subpopulation, Virtual Phenotypic Profiling would provide no additional benefit because multiple biomolecules could be assayed independently on different replicates of cells (first column, Fig. 6). At the other extreme, if cell-to-cell variability was so high that every cell defined its own subpopulation, Virtual Phenotypic Profiling would be impossible because the similarity of a given phenotype between two cells would bear no information about the similarity of any other phenotype between the same two cells (third column, Fig. 6). In practice, between these two extremes, cellular populations may be well modeled by classifying them into limited, but non-trivial, numbers of subpopulations with high classification accuracies, low phenotypic-profile noise levels, and/or significant divergence from population averages. Our approach can be used for profiling the phenotypes of such subpopulations (second column, Fig. 6).

The performance of our profiling method additionally depends on three key choices: the assignments of cells into subpopulations; the biomolecular markers available for building the reference feature set; and the reference features used to recognize subpopulations. Newer microscopy or labeling techniques⁷⁻⁹ that allow a larger number of markers may improve the performance of our methods, and enable a more complex definition of subpopulations. Other more complex phenotypic features, such as Gabor and multi-wavelet features, may also help to increase the performance of our method.

Our approach offers an extensible method for profiling cellular subpopulations, even as new markers and better methods for resolving fluorescence channels are developed. Our method is general, and may be applied to profile heterogeneity observed in a broad range of biological systems. Our method will enable the identification of molecular and phenotypic differences among subpopulations, and provide comprehensive data needed for modeling the distribution of biological states contained within heterogeneous cellular populations.

METHODS

Cell culture and differentiation

For 3T3-L1 preadipocytes, our differentiation protocol was based on previously described protocols^{29, 30}. In brief, we obtained 3T3-L1 preadipocytes from ATCC (CL-173, lot number 4715281), and propagated the cells in DMEM with 4g/l glucose (Hyclone) and 10% bovine calf serum supplement (BCS, Hyclone) in a 37 °C / 5% CO₂ incubator. We regularly sub-cultured cells at 70% confluence and only used cells with less than six passages for differentiation. To initiate differentiation, we plated preadipocytes to a 100 mm culture dish at

high density (3-4 millions in 20 ml medium) and incubated the cells overnight. On the next day (Day 0), we gently replaced the medium with stage I differentiation medium of DMEM: 4 g/l glucose supplemented with 10% fetal bovine serum (FBS, Gemini Bio-Products), 160 nM insulin (Sigma-Aldrich), 250 nM dexamethasone (Sigma-Aldrich), and 0.5 mM 3-isobutyl-1-methylxanthine (Sigma-Aldrich). We renewed the medium once on Day 2. On Day 3, we switched the medium to stage II differentiation medium of DMEM with 10% FBS supplement and 160 nM of insulin. Starting on Day 5, we changed the medium to adipocyte maintenance medium of DMEM and 10% FBS supplement. We renewed the maintenance medium every other day until the cells were transferred into imaging plates.

For HL-60 neutrophil-like cells, we used previously described protocols to culture and differentiate the cells¹⁹. In brief, cells were cultured in RPMI 1640 plus L-glutamine and 25 mM HEPES (Fisher Scientific) supplemented with antibiotic/antimycotic (Invitrogen) and 10% FBS (HyClone) in a 37 °C / 5% CO₂ incubator. To induce differentiation, 1.3% of Dimethyl sulphoxide (DMSO) (Sigma-Aldrich) were added to cell culture at a density of 0.2 million cells/ml and maintained for 7 days.

For H460 lung cancer cells, we maintained the cells in RPMI 1640 medium supplemented with 10% FBS, 2 mM L-glutamine and penicillin-streptomycin in a 37 °C / 5% CO₂ incubator.

Cell plating

For 3T3-L1 preadipocytes, cells were harvested from culture dishes and transferred to 384-well imaging plates (Nalgene Nunc) pretreated with 0.01% sterile poly-L-Lysine solution (molecular weight 70,000-150,000, Sigma-Aldrich) 40 hr before fixation. After gentle aspiration of adipocyte maintenance medium, phosphate buffer saline (PBS) was applied to the dishes to rinse off residual FBS, which inhibited the trypsinization process. After trypsinization, we gently spun down the detached cells with centrifugation and resuspended them in low glucose adipocyte maintenance medium (DMEM with 1g/l glucose and 10% FBS supplement). The cell suspensions were well-mixed by gentle pipetting to reduce aggregations, and then immediately transferred to imaging plate using a multichannel pipette. The optimal cell density was ~7000 cells at 50 µl of medium per well. We centrifuged the plates at 200 rpm for 2 min to bring cells down to the bottom of the glass slides.

For HL-60 neutrophil-like cells, we coated 96-well imaging plates (Nalgene Nunc) with 30 µl of 100 µg/ml fibronectin (BD Bioscience) diluted in distilled water for 1 hr at RT. Around 6000 differentiated HL-60 cells were plated into each well of the coated plate. After 20 min incubation at 37 °C, cell media was removed from each well, followed by addition of 30 µl chemoattractant solution, which consisted of 10 nM formyl met-leu-phe (fMLP, Sigma-Aldrich) diluted into Hanks' balanced salt solution (HBSS, Invitrogen) with 1.2% of low endotoxin Bovine serum albumin (BSA, Sigma-Aldrich). Cells were then incubated at 37 °C for 2 min.

For H460 lung cancer cells, cells were plated at a density of 30,000 cells per well on 96-well (uncoated) imaging plates (Nalgene Nunc), and incubated at 37 °C overnight.

Cell fixation and immunofluorescence staining

For 3T3-L1 preadipocytes, to stain intracellular lipid droplets, we replaced the maintenance medium with 4,4-difluoro-1,3,5,7,8-pentamethyl-4-bora-3a,4a-diaza-s-indacene (BODIPY® 493/503, Invitrogen) solution in DMEM at the concentration of 1 µg/l. After incubation for 30 min, we rinsed the BODIPY dye off with one wash of DMEM, and immediately fixed the cells with 4% para-formaldehyde (PFA, Electron Microscopy Sciences) in PBS for 15 min at RT. We kept the PFA solution at 37 °C and added it to the plate using an automatic microplate

dispenser (Matrix WellMate, Thermo Scientific). At the end of incubation, the fixative was flicked out quickly, and quenching buffer of 50 mM Ammonium Chloride was immediately added to each well to stop the reaction of PFA. After 10 min, we gently rinsed the cell plate 3 times with Tris buffered Saline (TBS) using a plate washer (ELX405, BioTek).

We permeabilized fixed 3T3L1 cells with 0.2% Triton-X100 in TBS for 5 min and washed twice with TBS on ELX405. We added blocking solution of 5% BSA in TBST to each well. After 1 hr incubation, blocking solution was completely flicked out and replaced with primary antibody mixtures (one from mouse with one from rabbit). We used rabbit anti-HSL, anti-phospho-HSL (Ser565), anti-PPAR γ , anti-C/EBP α (all from Cell Signaling Technology), anti-perilipin (Abcam), and mouse anti-adiponectin (gift from Dr. Philipp E. Scherer, UT Southwestern Medical Center). The plate was tightly sealed with Para film and incubated at 4 °C. After overnight incubation, we thoroughly rinsed off the primary antibodies with 3 washes of TBS and one wash of blocking buffer. Each wash had 10 min incubation time. The fixed cells were further incubated with AlexaFluor 647 conjugated anti-rabbit and AlexaFluor 546 conjugated anti-mouse antibodies (Invitrogen) for 1 hr, and washed 3 times with TBST. Lastly, we introduced 2 μ g/ml Hoechst to the plate. After two wash of TBS, we preserved the plate in 0.1% freshly prepared Sodium Azide at 4 °C.

For HL-60 neutrophil-like cells, cells were fixed with 30 μ l of 2x intracellular buffer (1.4 M KCl, 10 mM MgCl₂, 20 mM EGTA, 200 mM Hepes pH 7.5; diluted from 10x with water), 640 mM sucrose and 8% PFA (Sigma) at RT for 15 min on each well. After a quick wash with 50 μ l of TBS, we permeabilized cells with 50 μ l of 0.2% Triton-X100 in TBS for 15 min at RT, and blocked the permeabilized cells with 50 μ l of 3% BSA in TBST for 1 hr at RT. Then, cells were stained with primary antibody mixtures (one from mouse with one from rabbit). We used rabbit anti-Hem1 (gift from Dr. Orion Weiner, UCSF), anti-pPTEN (Bioscience), anti-alpha tubulin, anti-Rac1/2/3, anti-pAkt (Thr308), or mouse anti-pMLC2 (all from Cell Signaling Technology). We diluted the primary antibodies in blocking buffer, and incubated cells with 30 μ l of primary antibodies at 4 °C. After overnight incubation, we rinsed off the primary antibodies with 3 washes of 50 μ l of 0.3% Triton-X100 in PBS for 5 min each. The fixed cells were further incubated with AlexaFluor-488 conjugated anti-mouse (Invitrogen), and AlexaFluor-546 conjugated anti-rabbit (Invitrogen) antibodies for 2 hr at RT, and washed 3 times with 50 μ l of 0.3% Triton-X100 in PBS. Lastly, we incubated cells with 30 μ l of AlexaFluor-647 phalloidin (1:40 dilution from stock, Invitrogen) plus Hoechst (Invitrogen) for 30 min at RT. After three washes, cells were preserved in 50 μ l of PBS with 0.1% Sodium Azide.

For H460 lung cancer cells, cells were analyzed for cell cycle phase using a bioimaging-certified cell-cycle kit (BD Bioscience) according to the manufacturer's protocol. Briefly, adherent cells were BrdU (Cell-cycle kit, BD Bioscience) loaded at a concentration of 104 μ M for 1 hr at 37 °C. At RT, cells were fixed with PFA for 15 min, permeabilized with methanol solution for 10 min, and blocked for 30 min with FBS. Cells were stored overnight in PBS at 4 °C. The next day, cells were DNase-treated (Cell-cycle kit, BD Bioscience) for 1 hr at 37 °C, and stained with either rabbit anti-Cyclin A2 (Invitrogen), anti-Cyclin B1 (Santa Cruz), anti-p21-Waf1/Cip1, or anti-pRb (S608) (all from Cell Signaling Technology). These primary antibodies were diluted in FBS and incubated with cells for 2 hr at RT. Next, AlexaFluor-488 conjugated anti-BrdU, AlexaFluor-647 conjugated anti-pH3 (S28), Hoechst (all from Cell-cycle kit, BD Bioscience), and either AlexaFluor-546 conjugated anti-rabbit (Invitrogen), AlexaFluor-555 phalloidin (Invitrogen), or AlexaFluor-555 β -tubulin (BD Bioscience), were diluted in FBS and incubated at RT for 2 hr with the cells. Cells were washed with PBS between steps.

Image acquisition and preprocessing

We acquired images using a 20x objectives on an inverted fluorescence microscope (TE-2000, Nikon), equipped with a 12-bit CCD camera (CoolSNAP HQ, Photometrics) and controlled by the Metamorph software (v7.1, Universal Imaging). Sixteen images were acquired for each imaging well, and saved as 1392×1040 16-bit TIFF files. Then, we subtracted background intensities from the images using the rolling ball algorithm in ImageJ software (v1.38l, NIH), and stitched the 16 images together using the TurboReg plugin for ImageJ (version Feb 14, 2007; <http://bigwww.epfl.ch/thevenaz/turboreg>).

Data analysis software

All data analysis was performed using custom software written in Matlab v2007a (Mathworks), unless otherwise indicated. The software we used for automatic subpopulation identification, feature selection, and Virtual Molecular Profiles construction can be downloaded from <http://www4.utsouthwestern.edu/altschulerwulab>.

Supplementary Methods

Information on cell segmentation and categorization, phenotypic feature measurement and preprocessing, unsupervised subpopulation identification, supervised subpopulation classification, classification performance estimation, feature subset selection, Virtual Phenotypic Profiling clustering, noise-level measurement, population-averaged divergence measurement, and statistical analysis is available in Supplementary Methods.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank all members of the Altschuler and Wu lab at the University of Texas Southwestern Medical Center for critical discussion and performing manual cell categorization; Drs. Philipp Scherer and Orion Weiner for the gifts of the adiponectin and Hem1 antibodies (respectively); Jurg Rhorer at BD Biosciences for the gift of the cell cycle kit; and Drs. Steven Kliewer, David Mangelsdorf, Joyce Repa, and Philipp Scherer for stimulating conversations. This work was funded by the National Institutes of Health (R01 GM081549 to L.F.W. and R01 GM085442 to S.J.A.), the Welch Foundation (I-1619 and I-1644 to L.F.W. and S.J.A.), and the University of Texas Southwestern Endowment for Scholars in Biomedical Research (to L.F.W. and to S.J.A.). S.J.A. is a Rita Allen Scholar and is supported in part by a grant from the Rita Allen Foundation.

REFERENCES

1. Gallin JI. Human neutrophil heterogeneity exists, but is it meaningful? *Blood* 1984;63:977–983. [PubMed: 6324930]
2. Loo LH, et al. Heterogeneity in the physiological states and pharmacological responses of differentiating 3T3-L1 preadipocytes. *J. Cell Biol.* (in press).
3. Rubin H. The significance of biological heterogeneity. *Cancer Metastasis Rev* 1990;9:1–20. [PubMed: 2208565]
4. Chang HH, Hemberg M, Barahona M, Ingber DE, Huang S. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature* 2008;453:544–547. [PubMed: 18497826]
5. Slack MD, Martinez ED, Wu LF, Altschuler SJ. Characterizing heterogeneous cellular responses to perturbations. *Proc. Natl. Acad. Sci. USA* 2008;105:19306–19311. [PubMed: 19052231]
6. Lee TI, et al. Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* 2002;298:799–804. [PubMed: 12399584]
7. Malik Z, Dishi M, Garini Y. Fourier transform multipixel spectroscopy and spectral imaging of protoporphyrin in single melanoma cells. *Photochem. Photobiol* 1996;63:608–614. [PubMed: 8628752]

8. Resch-Genger U, Grabolle M, Cavaliere-Jaricot S, Nitschke R, Nann T. Quantum dots versus organic dyes as fluorescent labels. *Nat. Methods* 2008;5:763–775. [PubMed: 18756197]
9. Tsurui H, et al. Seven-color fluorescence imaging of tissue samples based on Fourier spectroscopy and singular value decomposition. *J. Histochem. Cytochem* 2000;48:653–662. [PubMed: 10769049]
10. Boland MV, Murphy RF. After sequencing: quantitative analysis of protein localization. *IEEE Eng. Med. Biol. Mag* 1999;18:115–119. [PubMed: 10497746]
11. Loo LH, Wu LF, Altschuler SJ. Image-based multivariate profiling of drug responses from single cells. *Nat. Methods* 2007;4:445–453. [PubMed: 17401369]
12. Perlman ZE, et al. Multidimensional drug profiling by automated microscopy. *Science* 2004;306:1194–1198. [PubMed: 15539606]
13. Rosen ED, MacDougald OA. Adipocyte differentiation from the inside out. *Nat. Rev. Mol. Cell Biol* 2006;7:885–896. [PubMed: 17139329]
14. Weiner OD. Regulation of cell polarity during eukaryotic chemotaxis: the chemotactic compass. *Curr. Opin. Cell Biol* 2002;14:196–202. [PubMed: 11891119]
15. Yin Z, et al. Using iterative cluster merging with improved gap statistics to perform online phenotype discovery in the context of high-throughput RNAi screens. *BMC Bioinformatics* 2008;9:264. [PubMed: 18534020]
16. Cristianini N, Shawe-Taylor J. An introduction to support Vector Machines: and other kernel-based learning methods. Cambridge University Press; New York, USA: 2000.
17. Pudil P, Novovicová J, Kittler J. Floating search methods in feature selection. *Pattern Recognition Letters* 1994;15:1119–1125.
18. Brasaemle DL. Thematic review series: Adipocyte Biology. The perilipin family of structural lipid droplet proteins: stabilization of lipid droplets and control of lipolysis. *J. Lipid Res* 2007;48:2547–2559. [PubMed: 17878492]
19. Weiner OD, Marganski WA, Wu LF, Altschuler SJ, Kirschner MW. An actin-based wave generator organizes cell motility. *PLoS Biol* 2007;5:e221. [PubMed: 17696648]
20. Eden S, Rohatgi R, Podtelejnikov AV, Mann M, Kirschner MW. Mechanism of regulation of WAVE1-induced actin nucleation by Rac1 and Nck. *Nature* 2002;418:790–793. [PubMed: 12181570]
21. Gratzner HG. Monoclonal antibody to 5-bromo- and 5-iododeoxyuridine: A new reagent for detection of DNA replication. *Science* 1982;218:474–475. [PubMed: 7123245]
22. Goto H, et al. Identification of a novel phosphorylation site on histone H3 coupled with mitotic chromosome condensation. *J. Biol. Chem* 1999;274:25543–25549. [PubMed: 10464286]
23. Darzynkiewicz Z, Bedner E, Smolewski P. Flow cytometry in analysis of cell cycle and apoptosis. *Semin. Hematol* 2001;38:179–193. [PubMed: 11309699]
24. Dolbeare F, Gratzner H, Pallavicini MG, Gray JW. Flow cytometric measurement of total DNA content and incorporated bromodeoxyuridine. *Proc. Natl. Acad. Sci. USA* 1983;80:5573–5577. [PubMed: 6577444]
25. Mittnacht S. Control of pRB phosphorylation. *Curr. Opin. Genet. Dev* 1998;8:21–27. [PubMed: 9529601]
26. Pines J, Hunter T. Human cyclins A and B1 are differentially located in the cell and undergo cell cycle-dependent nuclear transport. *J. Cell Biol* 1991;115:1–17. [PubMed: 1717476]
27. Dulic V, Stein GH, Far DF, Reed SI. Nuclear accumulation of p21Cip1 at the onset of mitosis: a role at the G2/M-phase transition. *Mol. Cell Biol* 1998;18:546–557. [PubMed: 9418901]
28. Whitfield ML, et al. Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol. Biol. Cell* 2002;13:1977–2000. [PubMed: 12058064]
29. Engelman JA, et al. Constitutively active mitogen-activated protein kinase kinase 6 (MKK6) or salicylate induces spontaneous 3T3-L1 adipogenesis. *J Biol Chem* 1999;274:35630–35638. [PubMed: 10585441]
30. Sadowski HB, Wheeler TT, Young DA. Gene expression during 3T3-L1 adipocyte differentiation. Characterization of initial responses to the inducing agents and changes during commitment to differentiation. *J Biol Chem* 1992;267:4722–4731. [PubMed: 1371508]

Figure 1. Building extensible Virtual Phenotypic Profiles of cellular subpopulations

Schematic showing the three major steps: **(a)** Step 1, stain and image cells with an initial biomolecular marker set (M_A , M_B , and M_C), and identify subpopulations based on an initial phenotypic feature set (F_1 , F_2 , and F_3). **(b)** Step 2, free up an imaging channel by dropping a marker (M_C), and adding more high-content features (F_4 and F_5) to partially compensate for the information lost. Train a classifier based on the new “reference” feature set to separate subpopulations. **(c)** Step 3, co-stain cells with the reduced marker set (M_A and M_B) and new markers (M_D , M_E , ..., M_Z), and use the trained classifier to separate cells into subpopulations. For each subpopulation, measure the averaged values of a specific chosen feature (F_s) for every marker, and combine the values into a common Virtual Phenotypic Profile. ($F_i(M_j) =$

phenotypic feature i based on marker j ; $S1, S2$ = subpopulations 1 and 2; pink and yellow squares = cells from $S1$ and $S2$; ellipsoids = boundaries of subpopulations; dashed lines or hyperplanes = trained classifiers)

Figure 2. High-content features partially compensated the decrease in classification performance due to dropping a marker

(a) Immunofluorescence images of 3T3-L1 cells near the centroids of the four identified 3T3-L1 subpopulations (S1 to S4) during adipogenesis. White curves = automated segmentation boundaries; scale bar = 20 μ m. (b) 2D scatter plot showing the Day 9 distribution of the subpopulations identified based on LD and AdipoQ levels. The percentages of cells are shown in parentheses. Data shown was 25% sub-sampled for visualization (triangles = subpopulation centroids). (c) Average minimum class accuracy (MCA) of subpopulation classification after systematic dropping of each initial marker. The accuracies for the best classifier type for each feature set (Supplementary Fig. 1a) are shown. (d) Average MCA of subpopulation classification after dropping the LD marker, and adding high-content features from the AdipoQ marker. Shown are the accuracies for the best overall classifier type, SVM-RBF classifier (Supplementary Fig. 1b). All features implicitly contain information from the DNA marker, which was used for segmentation. (INT/ALL = all initial features; INT/AdipoQ = initial adiponectin feature, INT/LD = initial LD feature; HC/ALL = all additional high-content features; HC/SFFS = high-content feature subset selected by sequential floating forward search; error bars = standard errors of cross-validation; ** = $P < 0.010$; *** = $P < 0.001$; one-sided paired t -test; $n = 3$)

Figure 3. Virtual Phenotypic Profiles had low noise levels and were significantly different from population averages

(a) Distributions of the nuclear or cellular levels of five new markers in subpopulations identified based on all initial features (INT/ALL), initial adiponectin feature alone (INT/AdipoQ), or high-content adiponectin features (HC/SFFS). (PPAR γ = peroxisome proliferator-activated receptor gamma; C/EBP α = CCAAT/enhancer binding protein alpha; PLINA = perilipin A; HSL = hormone sensitive lipase; pHSL = HSL phosphorylated at S565; error bars = medians $\pm$ upper and lower quartiles). (b) Absolute noise levels of the median cellular levels for the dropped marker (LD, left subpanel) and a new marker (PLINA, right subpanel) on subpopulations identified from either INT/AdipoQ or HC/SFFS (error bars = means $\pm$ standard errors; ** = $P < 0.01$; *** = $P < 0.001$; one-sided paired t-test; $n = 3$). (c) Number of initial markers (left subpanel) or new markers (right subpanel) with HC/SFFS noise level significantly lower than INT/AdipoQ noise level (one-sided paired t-test; $n = 3$). (d) Multi-dimensional scaling plot showing the relative distances of the Virtual Phenotypic Profiles of the four identified subpopulations identified from HC/SFFS to the profiles of the whole cellular population, or randomly-selected subpopulations. P-value shown was the probability that a profile may be obtained from randomly selected subpopulations (Supplementary Methods).

Figure 4. High-content features from AdipoQ gave similar clustering and heatmap of Virtual Phenotypic Profiles as the initial features

Dendrograms and heatmaps showing the hierarchical clustering of the normalized Virtual Phenotypic Profiles for the subpopulations identified based on either (a) all the initial features (INT/ALL), (b) initial adiponectin feature alone (INT/AdipoQ), or (c) high-content adiponectin features (HC/SFFS). The profile of each marker was normalized by dividing the profile with its maximum value. The known functions of the tested biomolecules are indicated. The distance metric used was Euclidean distance, and the clustering algorithm used was average linkage.

Figure 5. Virtual Phenotypic Profiling of polarizing and of dividing cells

(a) Immunofluorescence images showing polarizing HL-60 cells near the centroids of the three identified subpopulations, S1, S2, and S3 (white curves = automated segmentation boundaries; scale bar = 20 μ m). (b) Dendrogram and heatmap of normalized Virtual Phenotypic Profiles showing the hierarchical clustering of the biomolecules based on polarization indices and F-actin co-localization (Supplementary Methods). The known functions of the biomolecules are indicated (pPTEN = phosphorylated phosphatase and tensin homolog; pAKT = phosphorylated AKT; Hem1 = hematopoietic protein 1). (c) 2D scatter plot of the original (un-normalized) Virtual Phenotypic Profile for S3 cells showed that molecules from related pathways had similar values of polarization index and F-actin co-localization. (d) Immunofluorescence images showing dividing H460 cells near the centroids of the four identified subpopulations, S1, S2, S3, and S4 (corresponding to G1, S, G2 and M-phase cells respectively; white curves = automated segmentation boundaries; scale bar = 20 μ m). (e) Dendrogram and heatmap of normalized Virtual Phenotypic Profiles showing the hierarchical clustering of biomolecules based on their total cellular levels in different subpopulations (pRb = phosphorylated retinoblastoma protein). (f) 2D scatter plots of the Virtual Phenotypic Profiles showing subcellular localization relative to DNA and expression levels of the tested biomolecules in different phases of the cell cycle.

Figure 6. Effectiveness of subpopulation profiling depends on the degree of cell-to-cell variability (a) Simulated co-staining levels of a set of biomolecular markers on individual cells at different degrees of cell-to-cell variability (as measured by the average cell-to-cell dissimilarity). F_{ref} , a reference feature used to divide the cells into two subpopulations. F_s , a specific feature of marker A or B. (dashed line = decision boundary of the classifier used.) (b) Medians and interquartile ranges of the single-cell distributions of F_s in each of the two subpopulations. These features formed the basis for the Virtual Phenotypic Profiles (error bars = interquartile ranges.) (c) The effectiveness of subpopulation profiling evaluated using three criteria.