Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Applications in image-based profiling of perturbations

Juan C Caicedo[1,2], Shantanu Singh[1] and Anne E Carpenter[1]

A dramatic shift has occurred in how biologists use microscopy images. Whether experiments are small-scale or high-throughput, automatically quantifying biological properties in images is now widespread. We see yet another revolution under way: a transition towards using automated image analysis to not only identify phenotypes a biologist specifically seeks to measure ('screening') but also as an unbiased and sensitive tool to capture a wide variety of subtle features of cell (or organism) state ('profiling'). Mapping similarities among samples using image-based (morphological) profiling has tremendous potential to transform drug discovery, functional genomics, and basic biological research. Applications include target identification, lead hopping, library enrichment, functionally annotating genes/alleles, and identifying small molecule modulators of gene activity and disease-specific phenotypes.

Addresses
[1] Imaging Platform of the Broad Institute of Harvard and Massachusetts Institute of Technology, 415 Main Street, Cambridge, MA, USA
[2] Fundación Universitaria Konrad Lorenz, Bogotá, Colombia

Corresponding author: Carpenter, Anne E (anne@broadinstitute.org)

Current Opinion in Biotechnology 2016, **39**:134–142

This review comes from a themed issue on **Systems biology**

Edited by **Mark P Styczynski** and **Fabian J Theis**

For a complete overview see the Issue and the Editorial

Available online 17th April 2016

http://dx.doi.org/10.1016/j.copbio.2016.04.003

0958-1669/© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Through shifts in both technology and culture, biology is increasingly a quantitative science. Experimental methods that capture the activity or state of multiple distinct biological processes ('multiplexed' assays) are thus increasingly valued. The quantitative increase in the number of independent measures that can be collected in a single assay has brought with it a qualitative change in experimental strategies. In fact, '*profiling*' technologies enable measuring hundreds to thousands of distinct properties from biological samples, an approach quite distinct from '*screening*', which refers to traditional, targeted experiments that seek to quantify a single process or cell function. In this paper, we draw a distinction between these two experimental designs: profiling vs. screening. Profiling aims to capture and encode as many properties of a sample as possible, while screening focuses only on capturing known properties of interest, usually just a few (see Box 1).

Whereas classical biological assays might measure a particular feature of a biological sample in response to perturbation (e.g., ATP consumption, cell size, or phosphorylation state of a single protein), profiling experiments capture a wide range of readouts and use techniques from machine learning and data mining to identify similarities and differences among the measured patterns (sample properties). Thus, typically, the identity of each measured feature is not of particular importance (as in screening experiments), but instead the discovered difference itself is the crucial readout. The particular measured features themselves become relevant only when informative similarities/differences in patterns have been identified. Profiling is a powerful approach enabling high-throughput experimentation and multiplexed readouts to generate massive amounts of mineable data for use in systems biology and drug discovery.

Microscopy, followed by image processing, is one of the few profiling methodologies suited to relatively inexpensive, large-scale experiments involving hundreds of thousands of tested samples. It is compatible with many scales of biological samples: cells, tissues, or organisms (for simplicity in this review we refer to the most common case: cells). In image-based profiling (also known as morphological profiling or cytological profiling [1]), large amounts of quantitative morphological data are extracted from microscopy images of cells to generate a profile comprised of various measures of the shape and size of various cellular compartments and the intensity, texture, and colocalization of various markers (Figure 1). The goal is to identify biologically relevant similarities and differences among samples based on these profiles using appropriate computational models (see Box 2). Profiles of biological populations can be compared to predict previously unrecognized cell states induced by different experimental perturbations of interest.

Alternate highly multiplexed assays for biological systems include the measurement of gene expression, protein levels, and metabolites [2,3]. While powerful, they tend to be low-throughput to medium-throughput (hundreds to thousands of samples per experiment) [4] and characterize the average response of a population of cells (with important exceptions: high-throughput techniques for gene-expression are emerging [5] and RNA-seq can measure mRNA at single-cell resolution albeit currently for only a few samples per experiment). Measuring the response of an arrayed panel of cell lines, for example, the NCI-60 panel, or a panel of RNAi-perturbed lines, to

Box 1 High-throughput image-based screening vs. profiling.

Screening is a distinct strategy from profiling. Although both involve large-scale (high-throughput) imaging experiments, the goals differ: in *screening*, the researcher aims to measure one or more phenotypes that are visually discernible, and choose a subset of hits for further investigation [29]. In *profiling*, a broad spectrum of measurements is captured from each sample (unguided by prior knowledge) in order to reveal important differences and similarities with other samples. Screening depends on a biologist's expertise to interrogate a particular phenomenon whereas profiling takes an unbiased approach to grouping samples, with a higher potential to capture unknown mechanisms.

Image-based *profiling* experiments remain relatively rare [30]. By far, the most common application of high-throughput imaging is *screening* large collections of small molecules in order to identify research probes and therapeutic leads with useful biological properties (often called high-content screening, HCS). High-content screening is becoming more widespread in recent years, in part due to the realization that screens based on cellular phenotypes are on average more fruitful than higher-throughput but less physiological screens on isolated protein targets [31,32]. High-throughput image-based screens involve the development of assays that measure particular morphological properties of single cells. This requires flexible software tools for extracting measurements from images and robust computational models for subsequent data analytics [33], whether a single morphological feature is the basis of the screen, or whether machine learning is used to combine multiple morphological features in order to 'score' the relevant phenotype based on expert input from biologists [34,35]. Given that image-based compound screening is now relatively routine, we refer the reader to prior comprehensive reviews [29,32,36–40].

Although somewhat less common, genetic perturbations are screened in a similar manner as small molecules. The major limiting factor is the initial construction and validation of libraries of perturbation reagents; for completed screens to date, RNA interference is used most often but, depending on the organism, alternatives include direct genetic manipulation (e.g., yeast deletion strains) and overexpression libraries. For loss-of-function screens, CRISPR-Cas9 and related technologies are an exciting prospect; relative to RNAi these are currently thought to have lesser off-target effects, thus improving the reliability of results [41–43].

each perturbation is another form of profiling [6–9] but requires a separate well for *each* measurement in the profile and is thus not generally practical for experiments with thousands of perturbations.

In contrast, high-content imaging techniques can measure hundreds of biologically meaningful features with single-cell resolution in a single assay well, and can be scaled to high-throughput assays with relative ease (Figure 2). There is therefore significant interest in devising appropriate computational techniques specifically for image-derived profiles, which come with technical challenges (Box 2). There is also great potential for combining profiles from multiple methodologies (e.g., imaging + gene expression) in the same experiment to capture a broader range of cell activities.

In this review, we aim to introduce an array of applications that can be achieved using image-based profiling, the collective potential impact of which is immense. Studies in this field are shifting from proof-of-principle to biological discovery; their collective breadth spans research in drug discovery and functional genomics. Microscopy is thus moving from a qualitative assessment tool to a powerful high-capacity quantitative modality.

We focus here on applications that involve systematically profiling large numbers of perturbations interrogated by microscopy imaging; outside this scope are other important applications such as high-throughput image-based screening (Box 1), pathology applications involving human tissue samples [10], studies of population heterogeneity [11–14], engineering extracellular microenvironments [15–17], location proteomics [18–22], and expression and architecture mapping [23–28].

Drug discovery

Identifying mechanisms of action, targets, and toxicity for small molecules

Small molecule perturbations can produce morphological changes detectable by microscopy, and these changes can reveal similarities among compounds in terms of their phenotypic impact in a cellular context. Many studies have demonstrated that morphological profiles can correctly predict the mechanism of action (plus toxicity in some cases) for blinded compounds, by grouping each unknown compound with already-annotated compounds, based on their phenotypic similarity [1,44,49,55–61]; several have made novel predictions [62–65,66••,67,68]. This builds on a foundation of earlier work that identified targets based on visual similarities, for example, the identification of the mitotic kinesin Eg5 as the target of the small molecule monastrol based on a distinctive monopolar spindle phenotype [69] and the phenotypic matching of gene-compound pairs related to cytokinesis using parallel RNA interference (RNAi) and small molecule screens [70] or suppressor/enhancer screens for an RNAi-sensitized phenotype [71]. These studies often focused on oncology/cell cycle, which is not surprising given their dramatic visual phenotypes. A more recent study on hundreds of compounds and several isogenic cell lines revealed novel gene–drug interactions, which were also mapped using image-based phenotypes [72]. As well, methods for identifying individual reporter cell lines that are most useful for grouping compounds with similar mechanisms of action have also been developed, using phenotypic image-based profiles [73]. Grouping compounds by their phenotypic effects is not only feasible for static images of cells but also for videos of complex behaviour in whole organisms; the locomotor response of zebrafish correctly predicted many small molecules' mechanism of action, some previously poorly characterized [74].

Although some studies use the term 'screening' when describing the measurement of phenotypic properties of cells, they may be referring to 'profiling' (e.g. [68,73]).

Download English Version:

https://daneshyari.com/en/article/6487503

Download Persian Version:

https://daneshyari.com/article/6487503

Daneshyari.com