Loechel, F., Bjorn, S., Linde, V., Prastegaard, M., and Pagliaro, L. (2006). High content translocation assays for pathway profiling. *Methods Mol. Biol.* **356,** 401–414.

Mittal, V. (2004). Improving the efficiency of RNA interference in mammals. *Nature Rev. Genet.* **5,** 355–365.

Robinson, M. K., and Cobb, M. H. (1997). Mitogen-activated protein kinase pathways. *Curr. Opin. Cell. Biol.* **9,** 180–186.

Terry, R., Cheung, Y. F., Praestegaard, M., Baillie, G. S., Huston, E., Gall, I., Adams, D. R., and Houslay, M. D. (2003). Occupancy of the catalytic site of the PDE4A4 cyclic AMP phosphodiesterase by rolipram triggers the dynamic redistribution of this specific isoform in living cells through a cyclic AMP independent process. *Cell. Signal.* **15,** 955–971.

van Roessel, P., and Brand, A. H. (2002). Imaging into the future: Visualizing gene expression and protein interactions with fluorescent proteins. *Nature Cell Biol.* **4,** E15–E20.

Vassilev, L. T., Vu, B. T., Graves, B., Carvajal, D., Podlaski, F., Filipovic, Z., Kong, N., Kammlott, U., Lukacs, C., Klein, C., Fotouhi, N., and Liu, E. A. (2004). *In vivo* activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **603,** 844–848.

Wang, Y. L. (1989). Fluorescent analog cytochemistry: Tracing functional protein components in living cells. *Methods Cell Biol.* **29,** 1–12.

Wilson, C. (1995). ''The Invisible World: Early Modern Philosophy and the Invention of the Microscope.'' Princeton Univ. Press, Princeton, NJ.

Zhang, J. H., Chung, T. D., and Oldenburg, K. R. (1999). A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screen.* **4,** 67–73.

# [28] High-Content Screening of Functional Genomic Libraries

*By* DANIEL R. RINES, BUU TU, LOREN MIRAGLIA, GENEVIEVE L. WELCH, JIA ZHANG, MITCHELL V. HULL, ANTHONY P. ORTH, and SUMIT K. CHANDA

## Abstract

Recent advances in functional genomics have enabled genome-wide genetic studies in mammalian cells. These include the establishment of high-throughput transfection and viral propagation methodologies, the production of large-scale cDNA and siRNA libraries, and the development of sensitive assay detection processes and instrumentation. The latter has been significantly facilitated by the implementation of automated microscopy and quantitative image analysis, collectively referred to as high-content screening (HCS), toward cell-based functional genomics application. This technology can be applied to whole genome analysis of discrete molecular and phenotypic events at the level of individual cells and promises to

METHODS IN ENZYMOLOGY, VOL. 414 0076-6879/06 $35.00

DOI: 10.1016/S0076-6879(06)14028-8

significantly expand the scope of functional genomic analyses in mammalian cells. This chapter provides a comprehensive guide for curating and preparing function genomics libraries and performing HCS at the level of the genome.

## Introduction

A recent census of the human genome sequence estimated that our cells contain between 20,000 and 25,000 protein-encoding genes (Pennisi, 2005). If this tally is accurate, then the genetic complexity of *Homo sapiens* lies somewhere in between that of a roundworm (*Caenorhabditis elegans* ~18,000) and a flowering plant (*Arabidopis thalisala* ~25,000). If we consider ourselves at the pinnacle of the evolutionary process, then it stands to reason that gene function, not gene number, at least partially dictates biological complexity. However, our understanding of genome function remains sparse. The last five decades of molecular biology research have led to the functional characterization of less than half of the genes in our cells (Su *et al*., 2004). A major challenge for the biological community in the coming years will be the complete functional annotation of the human genome.

A number of technologies have emerged that facilitate the systematic assessment of various aspects of gene function at the level of the genome. These include expression (microarray) profiling and immunoprecipitation mass spectrometry (IPMS) (Chanda and Caldwell, 2003). Microarray studies can detect spatial or temporal changes in gene regulation on a global scale, whereas IPMS can be utilized to understand protein–protein associations at the level of the proteome. Until recently, genome-scale genetic experimentation was limited to studies in tractable model organisms, such as yeast, Drosophila, or *C. elegans*. However, a number of advances in the manipulation of mammalian cells have made genome-wide gain-of-function and loss-of-function analyses possible. The development of high-throughput transfection technologies allowed for the introduction of large-scale arrayed libraries of nucleic acids into mammalian cells in a rapid and economical fashion (Chanda *et al*., 2003; Ziauddin and Sabatini, 2001). Thus, one could monitor the effects of gene dosage on particular cellular phenotypes for hundreds to thousands of genes in a single experiment. Furthermore, discovery of the RNA interference (RNAi) pathway, which allows systematic suppression of gene expression in mammalian cells, has greatly facilitated loss of function studies (Elbashir *et al*., 2001). Large collections of cDNAs, transcribed siRNAs, and chemically synthesized siRNAs are currently being compiled and distributed. These arrayed libraries can be used to interrogate various cellular processes, including differentiation, apoptosis, oncogenesis, and inflammation. The repertoire

of these assays is only limited by the ability to detect a particular molecular or cellular phenotype.

High-throughput microscopy (HCS) is a developing technology that enables the rapid capture and analysis of cellular images and is a promising tool that can be used to significantly expand the scope of high-throughput cellular genetic analysis. It offers many of the capabilities and advantages of fluorescent-activated cell sorting (FACS) analysis with a few important exceptions. First, adherent cells do not require from their growth surface. This reduces intricate sample preparation procedures, which may be rate limiting in large-scale assays, and also enables kinetic reads. Furthermore, HCS allows for the analysis of subcellular events such as cytoplasmic to nuclear translocation, which is currently untenable through FACS measurements. Thus, the application of high-throughput microscopy toward large-scale cellular genetic experimentation will facilitate the interrogation of a large number of biological processes that have previously not been addressable in mammalian systems. This chapter addresses considerations and provides methodologies found to be critical in this area of functional genomics investigation.

## Development of Large-Scale Genomic Libraries

A number of academic and commercial institutions are currently building and distributing mammalian genome-scale libraries, or sublibraries, that target families of proteins (i.e., kinases) (Table I). Until more enabling technologies are established, the availability of these reagents largely obviates the need for the construction of these libraries for screening. Libraries can be categorized by the delivery methodologies utilized to introduce them to mammalian cells. These libraries can be transfected utilizing various high-throughput and/or automated processes or, alternatively, cells that are recalcitrant to the transfection process may be infected with various viruses harboring short-hairpin RNAs or cDNAs. These libraries have distinct advantages and limitations, which are discussed next.

### *cDNA Libraries*

Performing HCS gain-of-function screens using cDNAs enables the identification of proteins, which, when expressed ectopically, overcomes a rate-limiting step in the biological process being investigated (Chanda *et al.*, 2003; Harada *et al.*, 2005; Ziauddin and Sabatini, 2001). While multiple libraries of full-length cDNAs are available commercially, only those cloned into vectors bearing a mammalian expression promoter such

TABLE I
COMMERCIALLY AVAILABLE LIBRARIES FOR GENOME-SCALE FUNCTIONAL GENOMIC STUDIES

| Vendor | Web site | Collection | Species | Reagent Type | Number of Genes Targeted | Coverage | Comments |
|---|---|---|---|---|---|---|---|
| Ambion | http://www.ambion.com/ | Druggable genome | Human/mouse | siRNA | 5495 | 3 siRNAs/gene | Genome wide Libraries In process |
| Sigma | http://www.sigmaaldrich.com/ | MISSION TRC shRNA | Human/mouse | Lentiviral shRNA | 9651 | 5 shRNAs/gene | 5x Coverage of Human & Mouse Genome in process. Also available from OpenBiosystems |
| Open Biosystems | http://www.openbiosystems.com | shMir libraries | Human/mouse | Lentiviral and retroviral | ~35,000 | ~6 shRNAs/gene | |
| Qiagen | http://www.qiagen.com | Genome wide | Human/mouse | siRNA | 28,500 | 4 siRNAs/gene | |
| Invitrogen | http://www.invitrogen.com | Protein families | Human/mouse | siRNA | 49–636 | 3 siRNAs/gene | |
| Dharmacon | http://www.dharmacon.com | Genome wide | Human/mouse/rat | siRNA | ~22,000 | 4 siRNAs/gene | siRNAs in pooled format |
| IDT | http://www.idtdna.com | Protein families | Human/mouse | siRNA | TBD | TBD | 27-mer oligos, protein family libraries in process |
| Open Biosystems | http://www.openbiosystems.com | Assay-ready MGC | Human/mouse | cDNA | 16,000 | N/A | |

as the CMV promoter can be directly introduced by transfection into cells with the expectation that the encoded protein will be produced. Collections of CMV-driven, fully sequenced mammalian clones in arrayed-well format have been produced by the Mammalian Genome Consortium (Strausberg *et al*., 1999, 2002) and are available commercially from several vendors, including ATCC, Invitrogen, and Openbiosystems. Notably, ready-to-screen cDNA collections are available from the latter company.

From a technical standpoint, gain-of-function HCS assays based on forced expression of cDNAs have some advantages and disadvantages relative to loss-of-function assay conducted using RNAi. cDNAs are more difficult to produce in assay-ready format than synthetic siRNAs, as DNA requires preparation from living bacteria. Unlike siRNAs, however, cDNAs can be produced inexpensively once acquired in virtually limitless quantities, can be moved into viral vectors for delivery into difficult-to-transfect cell types, and suffer less experimental ambiguity associated with the persistent problem of off-target effects observed when using RNAi. However, results from cDNA screens may not necessarily be physiologically relevant, as these results are often prone to artifacts that result from simple overexpression of a protein.

### *Chemically Synthesized siRNAs*

Synthetic oligonucleotide libraries that target the entire human genome, or subsets thereof, are currently available from a number of vendors (Table I). These sets usually contain two to four siRNAs per gene. siRNAs are typically 21 to 27 bp in length, although it is currently unclear which length is most advantageous for screening. Once introduced within cells, double-stranded siRNAs are unwound by the activated RISC complex in an ATP-dependent process, and the antisense strand directs recognition of the target mRNAs sequences (Hannon and Rossi, 2004). The mRNA homologous to the introduced siRNA is then subsequently degraded, resulting in a loss of gene function (Hannon and Rossi, 2004).

Advantages of chemically synthesized siRNA libraries are that their synthesis is relatively standardized, which results in little variation in the purity of individual siRNAs in a collection. This, and the (comparatively) small size of these molecules, enables robust and consistent transfection throughout a large-scale screen. From a workflow standpoint, these libraries are much easier to deploy for screening than are either cDNAs or hairpin-encoding RNAi libraries, as they do not require significant investments in automation and technical expertise required for high throughput DNA preparation and normalization. With these collections, the reagents are not renewable. Thus, ordering additional RNAi for

validation and confirmation studies rapidly becomes an expensive prospect. Alternatively, siRNAs can be produced *in vitro* by a number of enzymatic methodologies (Betz, 2003; Kittler *et al.*, 2004; Luo *et al.*, 2004; Sen *et al.*, 2004), which can reduce the costs of large collections substantially.

The two major drawbacks of RNAi oligonucleotide libraries are the requirement that they be transfected, limiting their use to highly transfectable cell types, and the fact that gene suppression will occur transiently, limiting assay phenotypes to those that occur over the course of a few days. While long-term depletion of some genes may occur in some cell types with chemically synthesized siRNAs, robust cell division will gradually dilute siRNAs, and intracellular RNase activity will gradually contribute to siRNA metabolism as well.

### *Plasmid Short-Hairpin siRNAs*

It is also possible to clone DNA sequences that, when transcribed, will produce a single-stranded, hairpin-forming RNA giving rise after processing by the RNAi machinery to a functional siRNA. These short-hairpin RNAs (shRNAs) can be introduced into cells to mimic the effects of chemically synthesized siRNAs (Brummelkamp *et al.*, 2002). Briefly, the cloned double-stranded DNA oligonucleotide contains sense and antisense sequences targeting the gene of interest separated by a hairpin loop. Once inside the cell, a promoter—typically an RNA polymerase III promoter—drives the expression of an RNA, which forms a hairpin-like structure, which is modified by a protein called Dicer to produce nucleotide fragments analogous to chemically synthesized siRNAs (Hannon and Rossi, 2004). It has been shown that polymerase II promoters can also be utilized to drive the transcription of short hairpin siRNAs (Stegmeier *et al.*, 2005).

Due to the larger size of the vectors that harbor these short hairpins (typically >3 kb), they are more difficult to transfect than their chemically synthesized counterparts. However, this method of delivering RNAi constructs has several advantages over chemically synthesized siRNAs. First, by engineering these hairpins into viral vectors, such as lenti-, retro-, or adenoviral vectors, it is possible to introduce shRNAs into cell types refractive to transfection and also achieve long-term siRNA-mediated reduction in mRNA levels (Tiscornia *et al.*, 2003; Xia *et al.*, 2002). Second, because these vector-based libraries can be propagated in bacteria, they represent a renewable resource, which facilitates a reduction in long-term screening and validation costs. Short hairpin technology can additionally be utilized in *in vivo* model systems—either by direct viral injection or the creation of

transgenic organisms (Tiscornia *et al*., 2003). Although the methodology is not currently robust enough to replace conventional genetic approaches, it offers significant promise for characterization of organismal gene function in a rapid and robust manner.

### *Formats for RNAi Arrays*

The number of individual siRNAs to screen per well has been a matter of considerable debate. siRNAs may be arrayed such that there is one siRNA or multiple siRNAs targeting a single gene in a well (Aza-Blanc *et al*., 2003; Li *et al*., 2004). Alternatively, strategies employing multiple siRNAs targeting many different genes in a single well have also been used for successful loss-of-function screens (Berns *et al*., 2004; Paddison *et al*., 2004). The latter two approaches are variations of what is referred to as a pooling strategy. Overwhelming opinion currently mandates that, in order to be considered a validated activity, a phenotype must be observed using two nonoverlapping siRNAs targeting the same gene. This is largely due to the preponderance of nonspecific (or ''off-target'') RNAi effects in which short stretches of sequence homology can trigger the degradation or translational repression of an unintended message (Jackson and Linsley, 2004). Thus, to enable rapid discrimination between validated and potential off-target phenotype-to-target relationships it would be ideal to screen individual siRNAs in a well, with the library containing multiple siRNA per gene (called a ''high coverage'' library). However, some have hypothesized that pooling siRNAs against the same gene increases the efficiency of mRNA knockdown, arguing that since each RNAi is at approximately 25% of the concentration of a single siRNA transfection, there may be a corresponding reduction in potential off-target effects. If a pooled strategy is employed for screening, however, pools should subsequently be split and retested as individual siRNAs to identify those responsible for the activity seen in the original screen. If only a single active siRNA is obtained, it would be necessary to confirm the observed phenotype with a second, independent (non-overlapping) siRNA. Ultimately, decisions on assay format are dependent on both the goals of the project and the budget and infrastructure available for screening.

## Maintaining Large Arrayed-Well Plasmid cDNA or shRNA Clone Libraries

Clone library maintenance is extremely important and often involves a substantial investment of time (Fig. 1). Arrayed-well libraries, those with one clone species per well, are frequently costly to generate or obtain commercially. Therefore it is imperative to ensure that collections are free of phage, contaminant bacterial species, or fungi. Additionally, it is

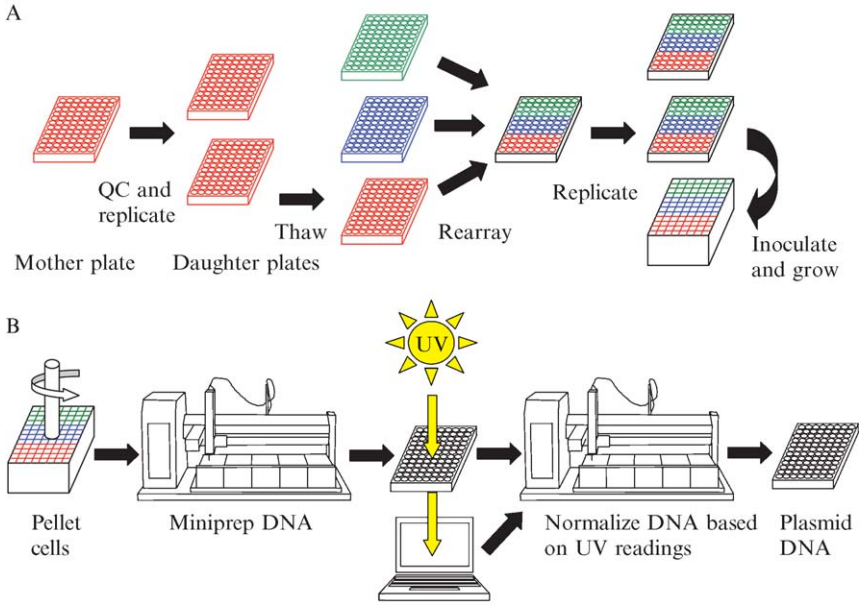


FIG. 1. Replication and preparation of clone collections used in functional genomic assays.

necessary to replicate archival copies both for general use and to provide a contaminant-free source should contamination with such organisms arise. Phage contamination, typically with T4 or T7 phage, can be detected by plaque assay (Sambrook, 1989) and, if detected early, can be contained by discarding affected plates and thorough bleaching (10%) of laboratory areas and instruments. Collections generated internally should be propagated in phage-resistant *Escherichia coli* strains such as DH10B-T1 (Invitrogen). Additionally, when propagating plasmid collections known to recombine, such as those containing inverted repeat sequences (i.e., viral LTRs), it is strongly recommended to use a recombination-deficient bacterial strains, such as STBL3 (Invitrogen) or Sure (Promega).

Contamination of libraries by bacterial species other than *E. coli* can be more difficult to detect and contain than contamination by phage. Contaminant bacteria such as *Pseudomonas aeruginosa* and *Bacillus cereus* may possess characteristics of *E. coli*, are frequently capable of outcompeting *E. coli* due to inherited, rather than episomal, antibiotic resistance, and may sporulate and contaminate neighboring cultures via aerosolization. Unusual colony morphology, color, smell, unusually large bacterial pellets, or failure to produce plasmid DNA may all indicate the presence of contaminant bacterial species. However, nonturbid cultures failing to yield plasmid DNA

are typically more indicative of phage. Methods for detecting contaminant bacteria include the spotting of inoculate using disposable pin replicators from plates onto solid, antibiotic-containing media, streaking out of sentinel wells on solid media, isolation and gram staining, and a number of other, more rigorous tests aimed at characterizing the contaminants (Krieg, 1984). Plates suspected of bacterial contamination should be discarded immediately and the laboratory area and instruments cleaned with a 10% bleach solution. Laboratory benches and equipment should be wiped and the wipes dotted onto antibiotic-containing solid media to test for the continued presence of live bacteria or bacterial spores.

### Collection Replication

Once satisfied that a collection is safe to use, at least two glycerol stock replicate (daughter) copies should be made as soon as practical (Fig. 1A). Generating ready-to-assay cDNA/shRNA and sublibraries for use in assay reconfirmation requires multiple freeze–thaw cycles tolerated well by some *E. coli* strains but less well by others. As a general rule, library copies should be tested for freeze/thaw tolerance and daughter copies discarded as soon as some wells begin to grow poorly. At any given time, one or more unthawed copies of a library should be reserved as a master copy. This copy should not be thawed until the next time the library needs to be replicated, and only a single copy should serve as a working copy. It is helpful to record which wells are empty in the original library plates in order to determine when a particular copy begins to become unviable. Additionally, it may be necessary to select (or ''hitpick'') specific clones from the original collections (i.e., full-length cDNAs, protein families, active genes, etc.). This rearray step enables the construction of customized collections (Fig. 1A).

Bacterial glycerol stocks in a 96- or 384-well format can be replicated using a variety of devices, all of which transfer liquid inoculate from one plate to another. The most robust and also most laborious method employs liquid inoculation using a multichannel pipettor fitted with disposable filter tips. Automated pipetting instruments can also be used if fitted for use of such tips. Fixed probe devices should be not be employed to pipette solutions containing bacteria, as probes can be corroded by frequent bleaching and heavy cultures may contain sticky bacterial colloids resistant to sterilization, thus facilitating the spread of contaminating organisms through the library if present anywhere in the clonal population. Fixed-pin inoculation employing either 96- or 384-pin disposable plastic pin replicators (Genetix), 96- or 384-pin reusable metallic pin replicators, or automated devices fitted with pin heads such as the QBot (Genetix) represents the least expensive and labor-intensive method for glycerol stock replication. Furthermore, unlike

fixed probe pipettors, fixed-pin devices are relatively easy to sterilize between inoculate steps, and most pin-tool heads, whether manual or automated, can be autoclaved periodically. The following protocol details how to make glycerol stock copies of 96- or 384-well bacterial libraries.

1. Before replicating, prepare sterile LB broth and, when cooled, supplement with appropriate antibiotic.
2. Add media to prelabeled, flat-bottomed 96- or 384-well plates fitted with lids to no more than half the total well volume.
3. Mother plates, if frozen, should be removed from 80° and unfoiled immediately (but relidded) to prevent liquid on the foil seal from thawing and dripping into adjacent wells during unfoiling, thereby contaminating adjacent wells. Once thawed, culture media from mother plates should be carefully transferred to daughter plates containing fresh media either by dipping pins from a pin replicator from one to the other or by aspirating and dispensing 5 to 10 $\mu$l using a pipetting device fitted with disposable filter tips.
4. Inoculated daughter plates should be lidded but left unfoiled, stacked two plates high, wrapped loosely with plastic film to prevent desiccation, and incubated overnight at 37° or until turbid without shaking or agitation.
5. Mother plates should be refoiled and replaced at −80°.
6. Once turbid, daughter plates should be cataloged for empty wells, supplemented with 50% the original media volume of a 1:1 glycerol/sterile water solution (16.7% final glycerol concentration), foiled, shaken slightly to mix, placed at −80°, and retained there until thawed for subsequent use.

If a collection proves particularly hardy, it is also possible to inoculate into daughter plates filled with media already containing glycerol, thereby avoiding the additional pipetting step. A test inoculation into glycerol-containing media will usually indicate whether this is possible.

### Growth of Bacterial Cultures in High-Throughput Format for DNA Preparation

Plasmid DNA is most commonly prepared in a 96-well plate format. Most assay applications require submicrogram quantities of DNA and thus 384-well miniprep formats may require unrealistically frequent preparation of DNA. The following protocol, illustrated schematically in Fig. 1, can be followed to prepare 96-well bacterial cultures for purposes of preparing transfection-grade plasmid DNA:

1. Fill 96-well deep-well blocks with 1.5 ml of Terrific broth (Sambrook, 1989) containing antibiotic. If automation is to be employed in a miniprep DNA preparation, ensure that the chosen deep-well blocks are compatible with the automated preparation instrument.

2. Source glycerol stocks are removed from $-80^{\circ}$ and unfoiled immediately but relidded thereafter to prevent frozen liquid adhering to the foil from thawing and sliding into neighboring wells during unfoiling. Once thawed, inoculate the deep-well blocks containing TB media using either a disposable pin replicator or by transferring liquid using a multichannel pipettor or liquid handler fitted with disposable, filtered tips.

3. Cover the block with a film that is permeable to air but not liquid or bacteria and place in a shaking incubator fastened tightly into an angle tube rack holder. Poor bacterial growth may indicate insufficient oxygen, in which case forced oxygen may be required.

4. Incubate with shaking at 270 rpm at the appropriate temperature until the culture reaches an $OD_{600}$ reading that is within the recommendations of the miniprep kit that will be used for plasmid isolation.

While some libraries require only an overnight growth period, it is not uncommon for libraries containing low-copy expression vectors or those propagated in slow-growing strains of bacteria to take 32 to 48 h to reach proper turbidity. To minimize the number of low-yielding or dead wells, blocks should be cultivated at least 24 h to allow slowest-growing cultures to reach the plateau growth phase. This is particularly important for cDNA libraries, as larger inserts may grow slower than smaller ones and leaky expression from eukaryotic promoters may cause bacterial toxicity.

### Preparing DNA from 96-Well Deep-Well Block Cultures

In a 96-well format, plasmid DNA to be employed in transfection-based assays or used for producing virus is most commonly separated from lysed bacteria using silica-based 96-well filterware or, less commonly, magnetic beads. Silica-based technologies are currently favored due to their versatility—the same filterware can usually be used manually or adapted to automated preparation instruments—and to the high yields of protein-, RNA-, lipid-, and genomic DNA-free plasmid DNA produced. Magnetic bead-based DNA purification, although usually less expensive on a per-well basis, frequently suffers from poorer yields or higher levels of endotoxin and is not discussed in detail here. High-throughput, filtration-based plasmid DNA miniprep kits are available from a number of commercial vendors, as are automated instruments to run them. The following protocol, illustrated schematically in Fig. 1B, has been adapted from Macherey-Nagel 11/2003/Rev. 02, and applies specifically to the Macherey-Nagel Nucleospin Robot 96 plasmid kit run either using a vacuum manifold or an MWG 2500 RoboPrep instrument. However, it could apply with minor modification to similar products from Qiagen, Eppendorf, Fisher,

and other major reagent providers deployed on such instruments as the Qiagen-BioRobot9600, the Beckman-Coulter-Biomek2000, the Beckman-Counter FX, the Tecan Freedom EVO, and others.

1. Centrifuge the deep-well blocks containing the bacterial cultures at 1000*g* for 10 min. After centrifugation, remove the membrane and immediately decant culture supernatant to a container for proper disposal. Invert and gently tap the blocks on several layers of paper towels to remove residual culture supernatant.

2. If performing minipreps manually using a vacuum manifold, we recommend performing liquid-handling steps using a high-quality, large-volume, eight-channel digital electronic pipettor such as the Thermo Electron Corporation Finnpipette Novus. While nonmotorized multichannel pipettors can be employed, for large collections their use may increase the risk to laboratory staff of developing repetitive motion injuries. If performing minipreps using automated instruments, program the instruments to perform the following steps, with the exception of the resuspension step, which is typically performed before blocks are subjected to automated handling.

3. Resuspend bacterial pellets by adding the recommended volume of resuspension buffer containing RNase (buffer A1) to each well, usually 250 $\mu$l, and then mixing vigorously using a plate shaker until the bacterial cells are totally resuspended. Clumped bacteria do not lyse well and incomplete lysis can degrade DNA yields; therefore light vortexing can be employed to resuspend clumps. Vigorous vortexing is to be avoided as it will shear the bacterial genomic DNA, causing it to remain in solution with the plasmid DNA instead of being filtered out with the bacterial cell wall components to which it would otherwise adhere.

4. Lyse the resuspended bacteria by adding the appropriate volume of alkaline lysis buffer (buffer A2) to each well, usually 250 $\mu$l, and incubating 2 min at room temperature.

5. Neutralize the lysis solution by adding the appropriate volume of neutralization buffer (buffer A3), usually 350 $\mu$l.

6. Transfer the neutralized crude lysates onto the lysate filter module. This module, which stacks above the filter plate but rests on the rubber gasket of the manifold, permitting a seal to form when vacuum is applied. It is designed to filter bacterial cell-wall components, including the attached bacterial chromosomal DNA, permitting the plasmid DNA and soluble bacterial components to flow through to the binding plate below. Incubate for 1 min and then apply vacuum at 750 mbar for 2.5 min to draw the cleared crude lysate into the binding plate. If using a vacuum manifold attached to a house vacuum, take care to predetermine the amount of vacuum to apply, as the application of an excessively strong vacuum can rupture the lysis filter

plate or permit lysate to be aspirated through an intact filter, either of which will contaminate the cleared lysate with undesired bacterial components.

7. Disassemble the manifold, discard the filter plate, and place the plasmid DNA-binding module containing the cleared lysate atop the manifold, taking great care to avoid spilling the cleared crude lysate or sloshing it into neighboring wells.

8. Apply vacuum at 750 mbar for 2 min to aspirate the cleared lysate through the DNA-binding silica resin.

9. Wash the plasmid DNA adhering to the silica membrane by adding and pulling through first 600 $\mu$l of AW wash buffer, then 900 $\mu$l of A4 wash buffer, and finally another 900 $\mu$l of A4 wash buffer by applying vacuum at 750 mbar for 45 s to allow buffer to pass the columns.

10. Dry the plamid-binding module by applying 10 to 15 min of maximum vacuum. Even minor amounts of residual ethanol can have a negative impact on some downstream applications, including transfection into certain cell types, and in these cases it is wise to centrifuge the binding modules placed atop blotting paper at 1000*g* for 2 min.

11. Elute plasmid DNA into an ultraviolet (UV)-transparent 96-well plate with 100 to 200 $\mu$l Tris buffer or TE buffer. Apply vacuum at 500 mbar for 5 min.

### Normalization of Plasmid DNA

DNA yields from 1.5-ml TB cultures grown to saturation and prepared using 96-well miniprep kits can exceed 15 $\mu$g/well, but more typically only reach 7 to 10 $\mu$g/well or, for 100 $\mu$l of eluate, approximately 100 ng/$\mu$l. However, these concentrations frequently vary considerably from well to well. Assay concentrations for transfected plasmid DNA range between 20 and 60 ng/well in a 384-well format depending on application but generally should be fixed for all wells for a given assay to ensure consistent transfection efficiencies. Superior assay results require that DNA be normalized to a target DNA concentration, but unlike the preceding steps, which can be performed manually, the need to dilute each well differently to obtain a single concentration makes manual normalization virtually impossible to perform at any scale exceeding several 96-well plates. Normalization therefore requires at a minimum a 96-well spectrophotometer and a liquid handler capable of handling microliter levels of liquid (Fig. 1B). The following protocol illustrates a general method for normalizing plasmid DNA.

1. Blank the 96-well spectrophotometer using a dummy UV-transparent elution plate filled with the chosen elution buffer at levels representative of the contents of the actual elution plates. This represents the background UV correction figure.

2. For each elution plate, measure the OD 280 nm and 260 nm twice, saving all values. If the spectrophotometer has an automated path length correction feature relying on readings in the 900-nm range to normalize for slight variations in well volume, it should be enabled, although minor variation (~10%) around 100 $\mu$l will not seriously affect quantitation.

3. Take mean values for OD 260 and 280 and calculate based on the OD 260 the amount of DNA per well, noting that for a 1-cm path length, 50 $\mu$g of double-stranded DNA has a OD 260 nm of 1. Many 96-well spectrophotometers can be precalibrated to a dummy elution plate, providing one or more log files for each elution plate containing DNA concentrations.

4. Calculate the volume of eluate and 5 m*M* Tris-HCl (no EDTA if proceeding to sequencing) required to achieve for each well a final concentration at the assay level, nominally 40 ng/$\mu$l.

5. Format a file for import to the liquid handler capable of interpreting these volumes and import to the liquid handler. For transfection-based assays, fixed probe devices incorporating water tip flushing can be employed readily without concern regarding well-to-well carryover of DNA due to the low relative lability of double-stranded DNA in a solution of Tris-HCl. Downstream applications involving the more sensitive polymerase chain reaction or other amplification strategies should always avoid the use of fixed probes, however.

6. To achieve a final concentration per well of 40 ng/$\mu$l, program the liquid handler to transfer DNA eluate containing 3.2 $\mu$g of DNA per well to a U- or V-bottom 96-well plate, adding the amount of 5 m*M* Tris-HCl required to bring the total final volume to 80 $\mu$l. U- or V-bottom destination plates should always be employed, as this permits most of the DNA to be removed during the creation of assay-ready 384-well plates. Flat-bottom plates will form a meniscus such that the last 20 $\mu$l of DNA eluation cannot be removed reliably.

7. Foil the normalized destination plates and store indefinitely at $-20$ or $-80^{\circ}$. Alternatively, samples may be stored in sealed plasmid plates for several weeks in the refrigerator.

8. The precision and accuracy of the entire normalization process can be assessed using a variety of intercalating DNA dyes, such as Quant-iT, according to the manufacturer's specifications (see following section).

### Arraying Collections into High-Throughput Assay Plates

Large-scale genetic screens in mammalian cells are usually conducted in 384-well assay plates, as this format provides a balance between the economies of high-throughput assay formats and the ability to interrogate a

sufficient number of cellular events to extract biologically meaningful data. Depending on assays and cell types intended to be utilized, between 15 to 60 ng of DNA and 5 to 20 ng of siRNA are arrayed in each well. However, virtually all siRNA libraries and prepared cDNA libraries exist in the 96-well format, necessitating reformatting from a 96- to a 384-well format for those HCS assays that can be executed in a 384-well format (Fig. 2). These 96-well (source plates) to 384-well plate (destination plate) transfers are accomplished most efficiently using a reformatting device, such as the Perkin Elmer EP3 or Beckman FX. These instruments can be used to accurately and consistently deliver between 500 and 3000 nl of liquid to a dry 384-well plate (''dry touch off''). Before embarking on this step, it is essential that the reformatting instrument be calibrated properly. This can be accomplished using limiting dilutions of fluorescein isothiocyanate dye, followed by measurement on a standard fluorescent plate reader. Once the coefficient of variance (CV) and standard deviations for each tip are determined to be in an acceptable range (10% or less CV), multiple screening sets may be arrayed in a single run and kept in appropriate storage conditions ($-80^{\circ}$) for future use.

For each spotting run, a final quality control step should be undertaken to ensure that the process yielded plated sets of acceptable screening quality. To accomplish this, it is necessary to use two to three of the plated sets to measure the concentration of siRNA or cDNA, utilizing reagents such as Ribogreen (Promega) or Quanit-it (Invitrogen), respectively, across the collection. An example of how to use the latter to assess DNA preparation is shown next.

*To Determine the DNA Concentration*

1. Spot the DNA plates to quality control/validate final spotted DNA amounts/well.
2. Set up standard curves in 384-well white-bottom plate as follows.

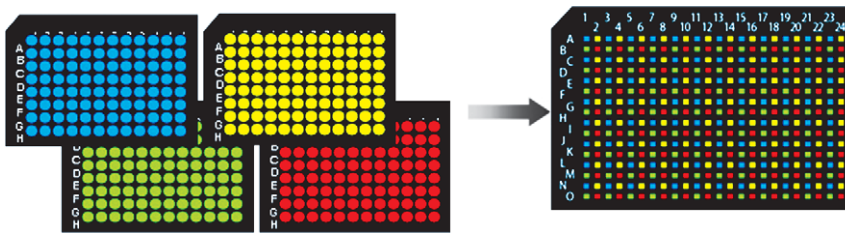

FIG. 2. Rearraying from 96-well to 384-well format.

a. Set up the standard curve using a known concentration DNA at 0, 5, 10, 20, 40, 60, 80, and 100 ng/well with a final volume of 20 $\mu$l/well in distilled water.
b. Set up the Quant-iT standard curve using the kit provided $\lambda$ DNA dilutions with the same setup as step 2a.

3. Remove the two spotted plates to be assayed from 20°/80° and equilibrate to room temperature.
4. Make a working solution by diluting Quant-iT DNA BR reagent 1:200 in Quant-iT DNA BR buffer.
5. Add 80 $\mu$l of diluted Quant-iT solution to each well and assay for fluorescence.

Generally, a screening set should have <20% CVs in concentration over 90% of the wells in a 384-well plate. If plated concentration varies significantly from these specifications, it is prudent to reprepare those plates.

## High-Throughput Transfections

For assay optimization, a cell line must be chosen that can be transfected at an adequate efficiency. This can be determined using a surrogate reporter such as green fluorescent protein (GFP) or a fluorescently labeled siRNA. In general, transfection efficiencies above 25% will yield acceptable results in HCS assays. Transfectability of cell lines varies, as do the efficiencies of different transfection reagents when combined with these cell lines. Performing transfections with a matrix of both different cell densities and transfection reagent/plasmid ratios to determine optimal conditions is recommended. Table II lists a number of reagents that may be initially tested for efficacy. Once a cell line/transfection reagent pair has been selected, it is necessary to determine if these conditions are compatible with high-throughput (or ''reverse'') transfection (Fig. 3).

### *Reagents*

FuGENE 6 transfection reagent or equivalent (Roche)
DMEM +10% fetal bovine serum (FBS) with antibiotics
DMEM
HEK 293T cells
384-well black clear-bottom plate (Greiner)
Sterile microfuge tubes
pCMV-GFP (or equivalent)

TABLE II
A SELECT LIST OF TRANSFECTION REAGENTS THAT CAN BE UTILIZED TO OPTIMIZE HIGH-THROUGHPUT TRANSFECTION CONDITIONS[a]

| Reagent | Vendor | Library |
|---|---|---|
| **Fugene** | **Roche** | **DNA/viral** |
| TransFectin | Bio-Rad | DNA/viral |
| CLONfectin | Clontech | DNA/viral |
| DreamFect | OZ Biosciences | DNA/viral |
| TransFast | Promega | DNA/viral |
| Escort | Sigma-Aldrich | DNA/viral |
| LipoGen | InvivoGen | DNA/viral |
| **Transit-Express** | **Mirus** | **DNA/viral** |
| GeneJuice | Novagen | DNA/viral |
| SuperFect | Qiagen | DNA/viral |
| GeneJammer | Strategene | DNA/viral |
| **Lipofectamine2000** | **Invitrogen** | **RNAi** |
| X-tremeGENE | Roche | RNAi |
| siIMPORTER | Upstate | RNAi |
| Block-it | Invitrogen | RNAi |
| RNAifect | Qiagen | RNAi |
| GeneEraser | Strategene | RNAi |
| RiboJuice | Novagen | RNAi |
| HiPerFect | Qiagen | RNAi |
| GeneSilencer | Genlantis | RNAi |
| siPORT | Ambion | RNAi |
| siLentFEC | Bio-Rad | RNAi |
| siFECTOR | B-Bridge | RNAi |
| TransIT-siQUEST | Mirus | RNAi |
| TransIT-TKO | Mirus | RNAi |
| jetSI | Polyplus | RNAi |
| Codebreaker | Promega | RNAi |

[a] Recommended starting reagents are in bold.

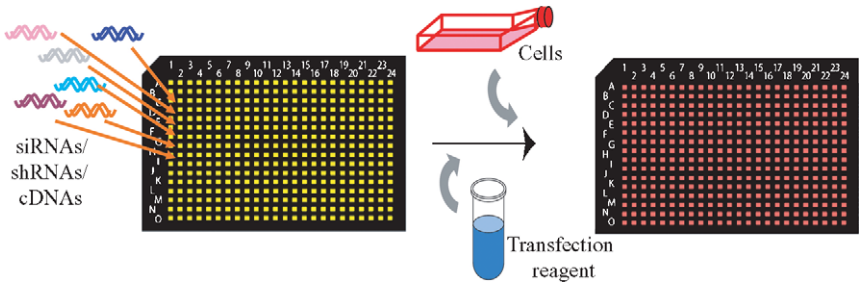


FIG. 3. Reverse transfection of siRNAs and cDNAs in multiwell plate format.

*Protocol*

1. Spot GFP plasmid at a desired concentration in a 384-well plate.
2. Prepare transfection cocktails in sterile microcentrifuge tubes by first adding FuGENE 6 to DMEM (20 $\mu$l/well) at a concentration of 3 $\mu$l per $\mu$g of spotted DNA and then mix gently by inverting the tubes, that is, 40 ng/well = 150 ng $\times$ 3 $\mu$l/$\mu$g DNA = 0.36 $\mu$l FuGENE 6.
3. Dispense 20 $\mu$l of reporter cocktail to each well.
4. Incubate at room temperature for 30 min.
5. While complexes are forming, prepare HEK 293T cells by harvesting and resuspending at a concentration of $5 \times 10^5$ cells/ml. Dispense 20 $\mu$l of cells/well.
6. Incubate for 24 to 48 h at 37°, 5% $CO_2$ depending on assay.
7. Analyze on a fluorescent plate reader or microscope (see p. 558).

*Note*: for a successful experiment, intrasample CVs should not exceed 20%.

## High-Throughput Retroviral/Lentivial Packaging

Large-scale production of lentiviruses or retroviruses harboring cDNAs or siRNAs can be accomplished using techniques based on the aforementioned high-throughput transfection protocol and transfection (assay)-grade DNA. Experimental methodologies for producing either lentivirus or retrovirus are effectively the same and only differ in the packaging constructs used (Fig. 4). Additionally, the choice of envelopes for viral pseudotyping is dependent on the types of assays that are intended to be run. However, we find that utilization of the vesicular stomatitis virus G protein (VSV-g) envelope enables the infection of a broad range of cell types and provides the most stability for viral manipulation and storage.

Finally, adequate safety precautions, which are, at minimum, in compliance with CDC recommended BSL-2 safety standards, should be in place before embarking on high-throughput production protocols. These viral particles are infectious to human cells, and the scale of the production runs presents additional risk factors not normally encountered in small-scale laboratory viral production experiments. Hazards can be reduced substantially through the use of automation, which limits contact with virus.

*Transfection Protocol (Lentiviral Packaging)*

1. Using a reformatter (i.e., Packard EP3), spot library viral DNA containing GFP into Greiner 96-well plates at a final concentration of 100 ng.

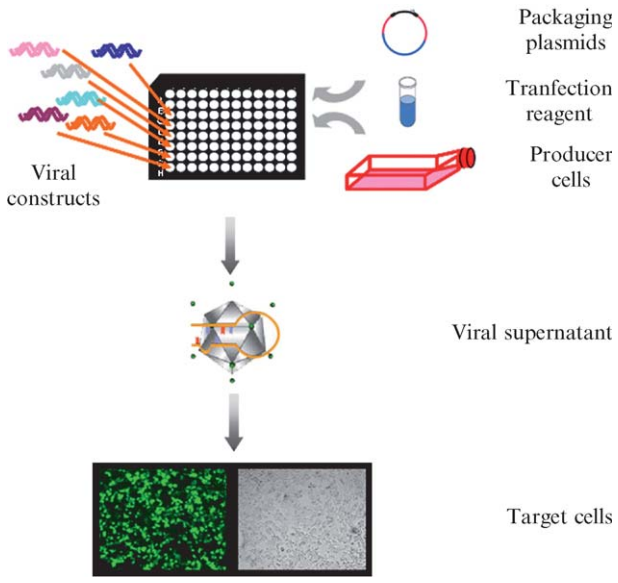


FIG. 4. A schematic for high throughput viral production.

2. Prepare the following master transfection mix based on per well:
   a. 50 $\mu$l of serum free-DMEM.
   b. 0.87 $\mu$l of FuGENE 6.
   c. 83.3 ng lentiviral packaging plasmid containing gag and pol.
   d. 32.1 ng lentiviral packaging plasmid containing rev.
   e. 44.9 ng envelope plasmid (i.e., VSV-g).
3. Using a Multidrop (or similar instrument), add 50 $\mu$l of the transfection mix.
4. Allow the plate to incubate at room temperature for at least 30 min.
5. Resuspend the 293T cells in 20% FBS DMEM, 2% p/s, 20 m*M* of nonessential amino acid, 20 m*M* HEPES at a concentration of $6.6 \times 10^5$ cells/ml.
6. After the transfection plate has incubated at room temperature for at least 30 min, dispense 50 $\mu$l of cells into the plate using the Multidrop or similar instrument.
7. Incubate the plate at 37° with 5% $CO_2$ for 24 h.
8. At 24 h posttransfection, replace media.
9. 96-well packaging plates should be spun for 10 min at high speed to ensure that non- or weakly-adherent 293T producer cells are pelleted (optional), after which the viral supernatant can be

transferred via an automated pipettor, such as a MiniTrack (Packard Instruments) fitted with filter tips into a 384-well plate.

*Infection Protocol for Titering Purposes*

1. Spot dilutions of virus (1 $\mu$L to 30 $\mu$L) from step 9 or step 10 into a 384-well plate.
2. Prepare HeLa cells at a concentration of $1.5 \times 10^6$ cells/ml.
3. Using the multiwell dispenser, dispense 10 $\mu$l of cells into the plate.
4. Using the multiwell dispenser, add 0.5 $\mu$l of 82.4 $\mu$g/ml of protamine sulfate in $H_2O$.
5. Incubate the plate at 37° with 5% $CO_2$ for 24 h.
6. After a 24-h incubation, add 80 $\mu$l of fresh media.
7. Incubate the plate at 37° with 5% $CO_2$ for another 24 h.
8. Analyze by FACS or HCS for marker expression. If the marker is a selectable marker, such as puromycin or neomycin, a selection step should be introduced after transduction of the HeLa cells.

To determine final titer, use the following formula:

$$([\%\text{GFP} + \text{cells}] * [\text{total cell number}])/(\text{volume of virus added [mL]}) = \text{infectious particles/ml}$$

## Instrumentation Required for Functional Genomics Screening

Automation required for functional genomic screen is largely based on the number of assays and depth of libraries intended to be screened. Small libraries (i.e., siRNAs against all kinases) can be interrogated with very limited investment in instrumentation. However, larger-scale efforts, including those undertaken by ''core'' facilities, require significantly greater investment in automation. Solutions range from integration of ''off-the-shelf'' instruments to fully customized robotic suites specifically designed for high-throughput functional genomics analysis and high-content screening (Table III). Advantages of partially or fully automated approaches include substantial increases in throughput, as well as data quality and reproducibility.

## Automated Microscopy

Morphological analysis of cell populations treated with cDNA or RNAi libraries requires automated microscopes, sophisticated image processing tools, and significant data storage. The second half of this chapter focuses on

TABLE III
A PARTIAL LIST OF INSTRUMENTS THAT CAN BE UTILIZED TO ENABLE HIGH-THROUGHPUT FUNCTION GENOMIC ANALYSES

| Procedure | Process | Instruments |
|---|---|---|
| DNA prep | Deep-well media dispense | μFill microplate dispenser (BioTek)<br>FlexDrop IV Exi (Perkin Elmer)<br>Multidrop DW (Titertek)<br>CyBi-Drop (Cybio) |
| | Minipreps | Biomek FX (Beckman Coulter)<br>MultiPROBE II (Perkin Elmer)<br>Freedom EVO (Tecan)<br>Biomek 2000 (Beckman Coulter)<br>BioRobot 9600 (Qiagen) |
| Reformatter | Rearray DNA from 96- to 384-well plates | Evolution EP3 (Perkin Elmer)<br>Bravo (Velocity11)<br>Janus (Perkin Elmer)<br>Beckman FX (Beckman Coulter) |
| Transfection | Microplate dispenser | BottleValve (GNF)<br>μFill microplate dispenser (BioTek)<br>NanoQuot microplate dispenser (BioTek)<br>BioRAPTR FRD (Aurora Discovery) |
| | Media changes | BottleValve (GNF)<br>ELx405 microplate washer (BioTek)<br>AquaMax DW4 (Molecular Devices) |
| Viral preps | Reformat virus from 96 to 384 | Same as DNA reformatter |
| | Seal plates for storage | PlateLoc (Velocity11)<br>ALPS 300 (Abgene)<br>Stacking heat sealer (Tecan) |
| Viral production and siRNA screening | Automated suite for high-throughput viral production, siRNA/cDNA transfection, cell fixation, and high-content screening | ALVS robotic suite (GNFSystems) |

the capabilities of these tools, provides protocols for sample preparation, and supplies image acquisition techniques to leverage library screening. Depending on the scientific application, HCS can be done with either live or fixed cell samples. Some microscope systems can also be equipped with environmental control chambers that regulate temperature and atmospheric $CO_2$ concentrations. Here we focus on the preparation and imaging of fixed time-point samples, but the techniques presented are equally applicable to live cell studies. In addition, all the commercial HCS systems now include some form of robust software that quantitatively analyze

image properties based on image segmentation or intelligent detection methodologies. When combined, these tools create a powerful bioinformatics platform where information can be extracted in a highly automated fashion.

## High-Throughput HCS Equipment

Unlike standard research microscopes, high-content or automated screening microscopes include specialized hardware and software for dealing with multiple samples arrayed into distinct regions of either a microtiter plate or a standard microscope slide. A broad range of imaging systems have been either modified or designed specifically to work with these arrayed sample formats (Table IV). At a minimum, an automated microscope must have the ability to automatically position each and every sample in a microtiter plate above the objective lens, contain a robust autofocusing system, and have the capacity to store image data without manual intervention. Many of the commercial systems available today also offer features such as automated filter switching, sophisticated image analysis tools, simultaneous multichannel image acquisition, and robotic plate loaders, to name a few. The selection of additional features must usually be balanced between the functional requirements of a system and the funds available for the acquisition system. The complete list of features and associated capabilities is beyond the scope of this chapter. Instead, we focus briefly on those features that are most advantageous for working with genomic libraries. In addition, we also briefly discuss additional resources typically required to successfully complete functional genomic screens such as clear-bottom microtiter plates and large-scale data storage systems.

Screening microscopes are subject to fundamental limitations in optical resolution, acquisition speed, and image contrast, all of which are of practical interest in HCS applications and must be balanced to achieve the highest quality outcome. For instance, short exposures increase the sampling rate but decrease the signal-to-noise ratio (SNR). Long exposures improve SNR but increase the overall time required to complete the screen. Compared to small molecule libraries that usually contain one million or more compounds, genomic libraries are relatively small and typically contain from 5 to 250,000 elements. Thus, speed is not usually the most important factor when evaluating a HCS system. Instead, having the ability to collect more fluorescent channels and to use more colors in a single sample is often an overriding concern. Determining the transfection efficiency on a well-by-well basis in a genomic screen requires one fluorescent channel and effectively reduces the information capacity associated

TABLE IV
A PARTIAL LIST OF HIGH-CONTENT SCREENING MICROSCOPES

| Company | Name | Web site | Confocal |
|---|---|---|---|
| Evotec | Opera | http://www.evotec-technologies.com | Yes |
| Molecular Devices | ImageXpress Ultra | http://www.moleculardevices.com/ | Yes |
| Molecular Devices | ImageXpress MICRO TM | http://www.moleculardevices.com/ | No |
| Cellomics | ArrayScan VTI HCS Reader | http://www.cellomics.com | No |
| Cellomics | KineticScan HCS Reader | http://www.cellomics.com | No |
| Cellomics | cellWoRx | http://www.cellomics.com | No |
| Applied Precision | cellWoRx | http://www.api.com | No |
| BD BioScience | BD Pathway 415 | http://www.bdbiosciences.com | No |
| BD BioScience | BD Pathway 435 | http://www.bdbiosciences.com | Yes |
| Compucyte | iCyte | http://www.compucyte.com | No |
| GE Healthcare | IN Cell Analyzer 1000 | http://www.amershambiosciences.com | Yes, confocal Focusing |
| GE Healthcare | IN Cell Analyzer 3000 | http://www.amershambiosciences.com | Yes |
| MAIA Scientific | MIAS-2 | http://www.maia-scientific.com | No |

with a biological response. Choosing a microscope capable of acquiring four or more colors per experiment, therefore, is typically justified when conducting genomic image-based screens.

The ability to select among different objective lenses when conducting genome library screens is also an important feature of a HCS system. A benefit of being able to choose between multiple magnifications is that the number of cells per image can be moderated. Some primary cell types, such as neuronal or epithelial cells, can spread out over large areas; a trade off in resolution may be essential to acquire the entire cellular region. While there is a loss in magnification when selecting lower power objectives, analyzing more cells can yield more statistically significant screening results. Conversely, higher magnification objectives with larger numerical apertures (NA) are more capable of collecting more photons per exposure, thus providing a brighter image with higher contrast. This allows for shorter exposures, faster overall acquisition times, and higher resolving capacity. However, the freedom to choose between different objectives ultimately rests with a need to measure a particular cellular phenomenon. More often than not, a 20× objective is the standard choice and provides an excellent compromise among magnification, resolution, and information capacity for most screens. We have also used our 10× objective extensively for many assays that do not require higher resolving capacities to obtain more statistically significant results while using fewer image frames per well.

### *Clear-Bottom Microtiter Plates*

The introduction of 96-, 384-, and 1536-well microtiter plate formats into biological research creates an opportunity to examine many biological samples in parallel while using a limited amount of reagent. However, this format can be particularly challenging to use in an HCS environment, as the cell fixation and staining steps frequently incorporated into successful assays dictate midassay aspiration or dispensing steps that can be physically difficult to accomplish using standard bench-top equipment. In contrast to whole-well assays where aggregate activity is assayed, the distribution of cells within the well is more important in high-content screens, as it is often difficult to focus on those cells that are along the edge of the well. Thus, specialized equipment is often needed to dispense reagents into wells that without disrupting cells.

### *Automated Microtiter Plate Washers*

Automated liquid-handling robots can consistently dispense and aspirate both media and fluorescent reagents for the best HCS assay results.

Plate washers that dispense toward the side of the well (side shooters) are usually more desirable for HCS applications than those that dispense directly into the bottom of the well (straight shooters). The ability to adjust the dispensing pressure is another important consideration when deciding on a plate washer. The ability to run the equipment at low pressure is often necessary when working with loosely adherent cell cultures. Finally, an aspiration head is also a very useful feature. During incubation, wells along the edges of the plate will often experience some significant evaporation. Thus the volume of liquid across all the wells in the plate can change dramatically depending on the time of incubation. An aspiration head provides the ability to first level the volume of media across the entire plate prior to the addition of exogenous reagents, ensuring an even concentration for all wells in the plate. We have found that this step is especially important for obtaining better $z$ scores and ultimately more consistent screening results with fewer false negatives.

### *Data Storage*

Conducting HCS screens also involves trade-offs among the number of optical slices, fields per well, channels acquired per exposure, and the number of cells required to reasonably determine a significant effect on a population. In addition, collecting images for every single cell per well can be time-consuming and must be balanced with the data storage available. Image data are much more information rich and also require more overall storage capacity than typical HTS data collections.

The cost of disk storage necessary for screening large-scale genome-wide libraries can be expensive. Unlike reporter gene assays, which only require megabytes (MB) to gigabytes (GB) of disk storage, HCS screens typically require GB to terabytes (TB) of storage space. These noteworthy data requirements also necessitate high-speed disk arrays and gigabit networking connections to accommodate the high rate of HCS data acquisition. Therefore, before venturing into a HCS screen, it is important to determine the data storage needed to facilitate image-based screening.

## Fluorescent Biomarkers for HCS Applications

For most HCS assays, the addition of an exogenous fluorescent organelle marker is often necessary prior to imaging. The most common reagent for HCS assays is a fluorescent DNA stain. Even cell lines that have been engineered with GFP-tagged proteins usually require DAPI, Hoechst33352, or other DNA stain so that all the cells in a plate can be identified. Some systems demand the addition of the DNA signal, as they

use this fluorescent channel for autofocusing. Moreover, unless the automated microscope contains an environmental chamber, it is also best to formaldehyde fix the cells, since unwanted biological effects may be introduced due to a change in temperature or reduced $CO_2$ concentrations.

The selection of fluorescent markers to determine organelle localization or to quantitate a cellular response is an important part of executing a successful high-content screen. Choice of markers is typically dictated by the assay readout and by the ability of the imaging equipment to discriminate between different wavelengths. Unlike standard research microscopes, automated screening microscopes tend to have a limited number of fluorescent settings. Microscopes that switch filters between each exposure tend to provide more flexibility when it comes to selecting fluorophores. However, some automated microscopes use a series of beam splitters (dichroics) in combination with multiple cameras to acquire different fluorescent channels simultaneously. While these systems can increase the rate of data acquisition rapidly, as there is no wait time for mechanical changes in filter wheel positions, they tend to limit the combination of different fluorescent reagents that can be used together.

When developing assays where multiple fluorescent markers will be employed in a single well, it is always best to select combinations where maximal emission spectra are as far apart as possible. It is also important to consider how broad excitation or emission spectra are for all the fluorescent elements used in a screen. Selecting two fluorophores that exhibit excessive overlap in wavelengths or wide emission spectra can seriously reduce the contrast of the resulting images. This loss in contrast may also extend the overall time needed to complete a screen if excessively long exposures are required to compensate for the high background signals. More importantly, overlapping signals between different channels can produce misleading results and could yield many false positives in a screen.

Biomarkers that exhibit a difference in their emission spectra based on molecular interactions can be extremely useful in high-content screening applications. Calcein AM is used commonly in HCS applications, as it is converted into a fluorescent conjugate via the activity of cellular esterases. This particular dye is permeable to cell membranes and is an excellent cytoplasmic marker. Once inside a cell, the molecule is cleaved by cellular esterases and results in a fluorescent derivative that is also nonmembrane permeable. This is ideal for live-cell applications, as there is little or no background signal from outside the cell, meaning that little or no washing is required after the dye is applied. Other fluorescent dyes depend on changes in pH for maximal fluorescence emission. These classes of dyes are quite useful for activity or viability assays. However, they may only work with

live cells and, in some cases, lose their ability to selectively highlight a cellular organelle if fixatives are added after treatment. One particularly helpful source of information concerning organelle identification, fluorescent spectra, and fixation compatibility can be found on the Web site of Invitrogen (Molecular Probes, http://probes.invitrogen.com). The advantages and disadvantages must be weighed when choosing among the various options. Careful selection of organelle markers can save time and money during the screening process, not to mention a great deal of frustration.

The use of antibodies for the demarcation of protein localization in cellular studies is an important part of cellular biology. However, in order to detect intracellular proteins, cells must first be treated with a fixative and their membranes permeabilized so that the antibody can gain direct access to these proteins. This process is usually extensive and requires multiple incubation and washing steps. With the advent of automated microtiter plate washers, the use of antibodies for protein localization studies becomes more tractable. Some washers can even be programmed to complete multiple rounds of dispensing and aspiration such that little or no manual intervention is required, especially when connected directly to plate stackers that load and unload the plates. However, the use of fixatives and the number of dispensing and aspirating steps demand that the cells are treated gently, even for tightly adherent cell lines. Thus it is best to use directly conjugated fluorescent antibodies if possible.

## Preparing Samples for Automated Microscopy

We have optimized a set of protocols for HCS studies using antibody-based protein localization in aldehyde-fixed cells. Cell-permeable reagents can also be used with these protocols by simply ignoring the detergent permeabilization steps. This set of protocols has been used extensively with cultures dispensed into 384-well plates (commonly used in the HTS cDNA and RNAi screens). However, given an optimized assay, these protocols could be adapted easily to other plate formats (e.g., 96- or 1536-well), with the only difference being that the amounts of reagents must be scaled up or down as necessary.

### *Selecting Appropriate Cell Lines for HCS Assays*

Pipetting fixation or fluorescent reagents directly onto the cultured cells can often disrupt their attachment to the surface material depending on the cell line and pipetting or dispensing equipment being used. This treatment

can dislodge the cells and introduce large empty spaces that prevent consistent cell counts from well to well. Thus it is important to consider the cell line being used and to optimize the reagent addition step before proceeding with a full-scale library screen.

Strongly adherent cells usually work best for HCS assays. HeLa, U2OS, MCF7, and other strongly adherent cell lines typically provide the best opportunity to conduct a successful image-based screen (Harada *et al.*, 2005). These cells grow flat, do not generally pile up on one another, and adhere to coated tissue culture plates very well. Conversely, we have found that HEK-293 or suspension cells are less desirable for HCS assays, as any disruption to the culture media may rearrange the cell distribution. Since we do not always have the luxury to work with strongly adherent cell lines, it is possible to work with loosely attached cells, although this demands that reagent addition or aspiration be done at low pressure and flow conditions, therefore extending assay duration, which, in itself, can be an important consideration for either very large screens or those with brief assay windows. Using appropriate coating of plates (i.e., poly-D-lysine) and enough care, even suspension cells can be used successfully with HCS assays.

*Protocols for Applying Cell Fixatives*

1. Remove the 384-well plates containing the transfected/infected cells from the $CO_2$ incubator after the appropriate incubation period.
2. Aspirate all but 20 $\mu$l of media from the wells using an automated plate washer, taking care not to disturb the cells, which should have adhered or settled to the bottom of the plate during incubation. (This step helps create equal volumes in each well and reduces edge effects often created from evaporation.)
3. Make an 8% (v/v) stock fixation solution of paraformaldehyde (PFA) suspended in phosphate-buffered saline (PBS).
4. Dispense 20 $\mu$l the stock PFA solution into each well using an automated liquid dispenser or plate washer. (The final concentration will be 4% paraformaldehyde in the cell culture media.)
5. Incubate the plate at room temperature for approximately 10 min.
6. Aspirate 30 $\mu$l of the PFA/cell culture media using an automated plate washer. (*Do not* turn the plate upside down and shake out media, as this will often dislodge many cells.)
7. Add 40 to 60 $\mu$l of PBS to each well using an automated liquid dispenser and incubate for approximately 5 min.
8. Repeat steps 6 and 7 two more times to wash out any remaining PFA.

*Protocols for Antibody Staining*

1. Add 40 to 60 $\mu$l of PBS supplemented with 0.5% Triton-X100 to each well.
2. Incubate the plate at room temperature for approximately 10 min.
3. Aspirate most of the buffer without disturbing the cells.
4. Add 40 to 60 $\mu$l of PBS supplemented with 0.1% Triton-X100 (called PBST) and 1% BSA to each well.
5. Incubate the plate at room temperature for at least 30 min.
6. Aspirate off all but 20 $\mu$l of the buffer and add the 1° antibody diluted a small volume. We typically add 10 $\mu$l of antibody solution at 3× concentration to each well. (If a 1:1500 final concentration is desired, add 10 $\mu$l of a 1:1500 stock.) The antibodies are always resuspended in PBST + 1% BSA and 0.5% sodium azide ($NaN_3$).
7. Incubate the plate for at least 60 min at room temperature or overnight at 4°.
8. Aspirate 30 $\mu$l of the buffer using an automated plate washer. (*Do not* turn the plate upside down and shake out media, as this will often dislodge many cells.)
9. Add 40 to 60 $\mu$l of PBS to each well using an automated liquid dispenser and incubate for approximately 5 min.
10. After the incubation period, repeat steps 8 and 9 three times using PBST with 10-min incubation periods between each wash.
11. After the incubation period is complete, add Hoechst 3342 (nuclear dye) to the existing buffer volume at 1:10,000 final and incubate for an additional 5 min.
12. Again, repeat steps 8 and 9 three times using PBST with 10-min incubation periods between each wash.
13. After the last wash, leave about 30 to 40 $\mu$l in the wells and seal the plate using an adhesive aluminum seal.
14. Remove any dust or large particles from the bottom of the plate before starting the imaging process.

## Determining Transfection or Transduction Efficiency

Unlike chemical screening studies, functional genomic applications do not necessarily affect all the cells in a well equally. Depending on the cell type, some cells are more capable of internalizing exogenous oligonucleotide-based reagents or of accepting viral particles. This limitation can create a disparity in the number of cells responding to a particular treatment. When combined with fluorescent tags, however, functional genomic libraries can be analyzed more efficiently using a HCS approach as opposed

to other HTS assays. Individual cells that have been transfected efficiently with a particular cDNA or chemically synthesized siRNA can first be identified from other nontransfected cells. Plasmid-based transfections such as cDNAs or lentiviral delivery systems can usually include a fluorescent protein expression cassette such as GFP or, in the case of straight plasmid transfections, involve cotransfection with a marked plasmid. Alternatively, synthetic nucleotide treatments with double-stranded RNAs can also be measured by tagging the oligonucleotides with fluorescent molecules, such as an Alexa™ fluorescent dye (Fig. 5). This may be of questionable value for several reasons. First, it may be very inconvenient or impractical to conjugate each siRNA is an arrayed library—all current commercially available libraries are unconjugated. Additionally, the effect that different conjugates may have on individual siRNA potency has not been systematically tested. Co-expressed markers in viral shRNA constructs may be more useful in marking transduced cells, since as the independently expressed marker is not likely to impact shRNA potency and shRNA function can be measured over a longer time frame than can be done with siRNA. Fluorescently labeled siRNAs are best employed during the optimization phase of an assay or for sentinel wells in live assay plates to mark plates that may have suffered from poor handling during siRNA transfection. At any rate, transfected cells can then be identified positively by thresholding on only those cells that exhibit a certain fluorescent intensity before determining the differential response of a cell to a condition or stimulus. With our siRNA-based screens, we commonly add a fluorescently tagged negative control (ds)RNA to one well per plate for establishing a baseline transfection efficiency. In several of our plasmid based libraries, we use GFP expression to measure transfection efficiency (Fig. 5). Minimum transfection efficiencies need to be empirically determined and will vary among assays. Typically, we favor cell types that have median high-throughput transfection efficiencies of 25% or more, and have found it difficult to extract statistically significant data from wells where less than 10% of the cells (~50 events/well) have been co-transduced with an shRNA/cDNA and a marker.

## Quantitative Image Analysis

After establishing the assay conditions and the transfection efficiencies for a particular genomic screen, the next step is to image and analyze the biological effect on cell populations in a quantitative manner. Quantitative image analysis is typically approached in one of two different ways using either image segmentation or supervised learning algorithms. All the commercial HCS readers come equipped with some form of image segmentation software. More recently, supervised and pattern recognition techniques have

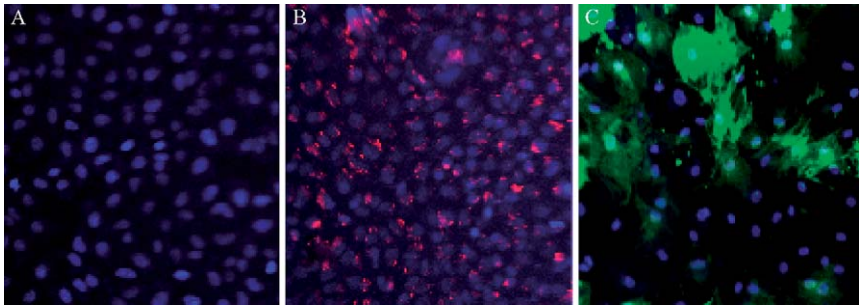


FIG. 5. Use of fluorescent transfection markers in functional genomic screens. (A) Negative control transfection well; DNA (blue, Hoechst33342). (B) siRNA tagged with Alexa-647 fluorescent molecule and transfected into HeLa cells; siRNA (red), DNA (blue). (C) Lentiviral infections with GFP expressing plasmid; GFP (green), DNA (blue).

been added to the HCS analysis packages. The latter are generally quite sophisticated and are beyond the scope of this chapter. Instead, we focus on how image segmentation is used to measure a cellular response. In particular, we present several examples that show how cell populations can be measured for cell cycle differences and cytoplasmic protein redistribution.

### *Image Segmentation*

The most common approach to HCS analysis is done by segmenting multicolor images first on one channel, typically the nuclei signal, and then measuring intensity properties in the other color channels to determine a cellular response. Once the multicolor image is acquired, quantitative image analysis is achieved using a class of image-segmentation techniques for morphological and cellular identification (Fig. 6). The original fluorescent nuclear channel (e.g., Hoechst 33342 signal) is converted into a binary image by thresholding the grayscale image; this is then used as a mask for discriminating between nuclear and background signals. Once binarization is complete, regions of interest (ROI) are established using contour-based detection to identify the perimeter of each nucleus. Each perimeter map is converted into a geometrical feature and is designated as a nuclear ROI. These regions are filtered based on size and fluorescent intensity to remove ROIs from nonnuclear signals sometimes created from particulate or fluorescent cellular debris. The average fluorescent intensity in the other fluorescent channels is then determined (e.g., Alexa 647 signal) and gating on only those nuclear ROIs that have a strong positive signal, called the

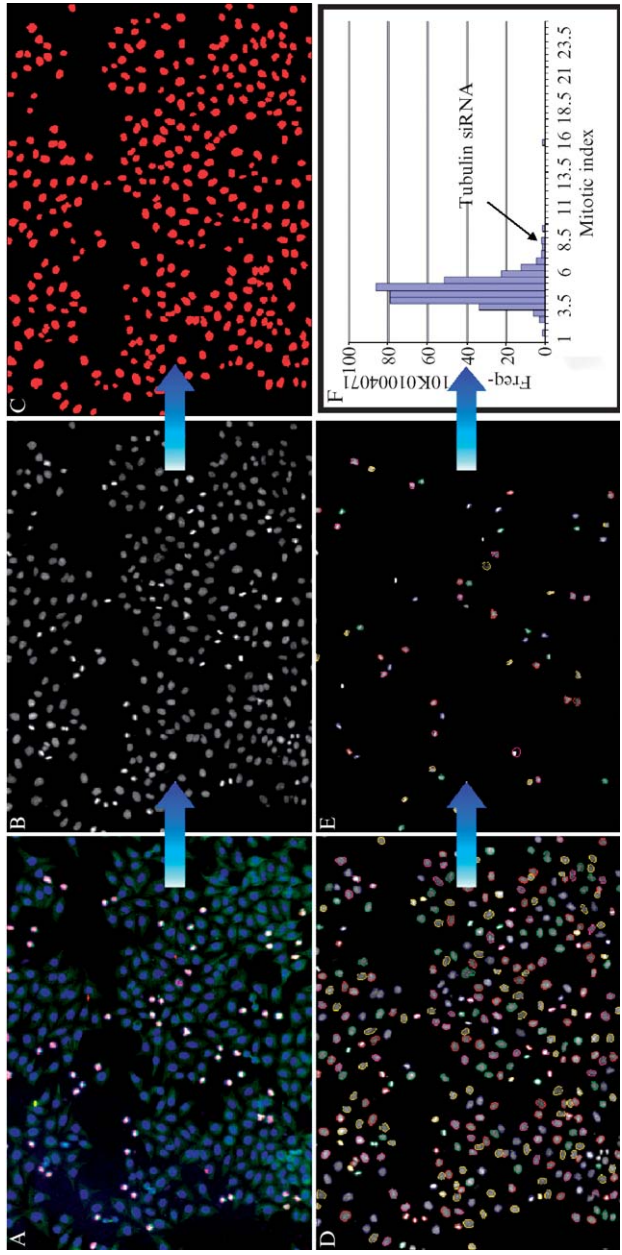


FIG. 6. Automated identification of mitotic cells.

response index. In addition to the response index, several other parameters are calculated for characteristics, including roundness of the nuclei, cellular proliferation, fluorescent DNA intensities, and the average size of the nuclear ROIs.

### *Protein Relocalization*

The ROIs can also be manipulated to measure the fluorescent signals in other portions of the cell. For example, if each of the nuclear ROI is dilated so that it includes a portion of the cytoplasmic region, a quantitative nuclear translocation assay becomes possible. If a protein, such as a transcription factor, is marked with a fluorescent molecule, the fluorescent intensity in the expanded ROI can be compared with the original nuclear ROI based on the fluorescent channel of the protein (Fig. 7A and B). This measurement can then serve as a cellular response to external stimulus that might normally drive the transcription factor from the cytoplasm into the nucleus. A genomic library can then be screened for inhibitors or activators of this translocation event (Cho *et al.,* 2006).

Even more complex pattern changes in protein localization can be determined using HCS assays. Upon activation, some cytoplasmic proteins, such as $\beta$2-arrestin, can be driven to relocalize into cellular vesicles (Fig. 7C–E). These changes in protein localization can be assayed readily by first determining a ROI for the entire cytoplasmic region. This is usually best done with the addition of a cytoplasmic dye that does not overlap with the nuclei or protein channels, such as calein AM. Once a cytoplasmic ROI is established, then it is possible to look for discrete areas of maximal intensity, also known as punctuate spots. These can be measured and counted per cell to determine a cellular response to the stimulus. Those cells with more or brighter spots would be considered to respond stronger.

The number of HCS assays possible is virtually limitless. Assuming that a fluorescent measurement can be determined and a change in a morphological parameter established, a change in the cells response to stimuli can be output. This creates an opportunity to expand biological discovery and will also require the development of future pattern identification methods as we probe for more complex biological phenomenon.

## Summary

This chapter focused on the preparation and application of genomic libraries with respect to high-content screening. These two technologies are particularly complementary and offer a powerful approach to conducting genomic analyses. The technology is robust and demonstrates the

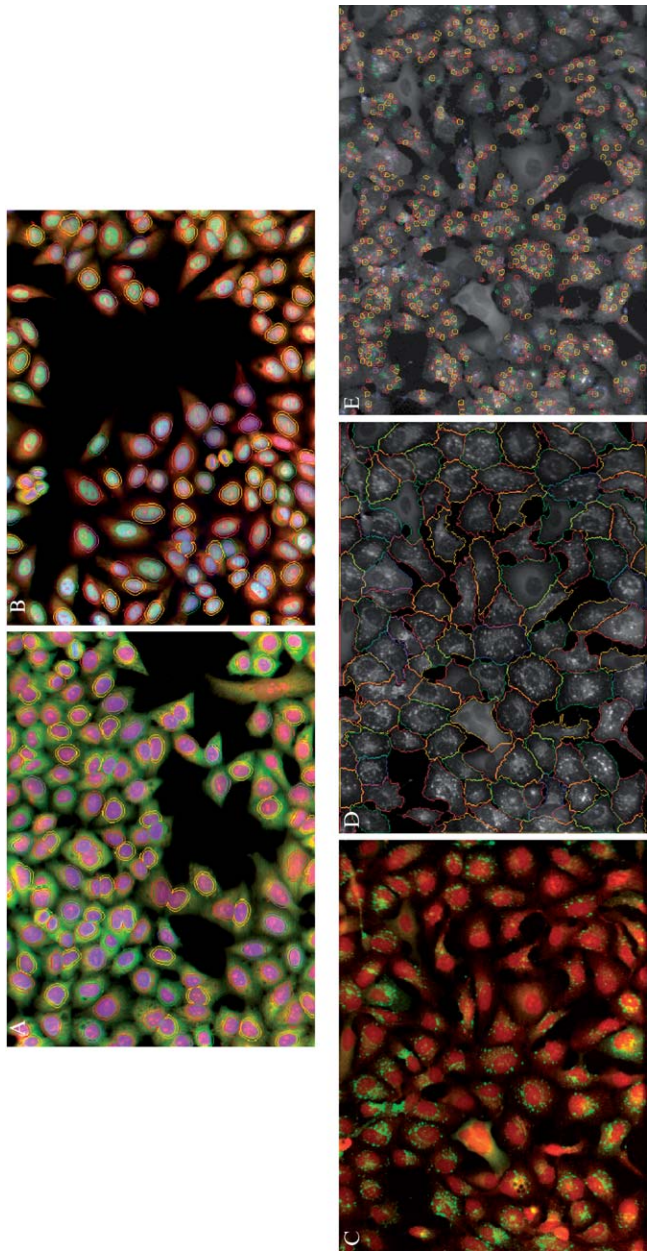


FIG. 7. High-content image analysis to determine protein redistribution.

appropriateness of image-based technology for enhancing analysis based on those cells in a population that have the highest degree of potential silencing or expression, depending on the screen. While each interrogated cellular phenotype will require additional modifications, these protocols and summaries provide a framework from which large-scale genomic screens using HCS can be achieved. Because each component of this approach involves technologies that are in their infancy, we anticipate continual advances that will transform this methodology for understanding gene function on a global scale.

## References

Aza-Blanc, P., Cooper, C. L., Wagner, K., Batalov, S., Deveraux, Q. L., and Cooke, M. P. (2003). Identification of modulators of TRAIL-induced apoptosis via RNAi-based phenotypic screening. *Mol. Cell* **12,** 627–637.

Berns, K., Hijmans, E. M., Mullenders, J., Brummelkamp, T. R., Velds, A., Heimerikx, M., Kerkhoven, R. M., Madiredjo, M., Nijkamp, W., Weigelt, B., Agami, R., Ge, W., Cavet, G., Linsley, P. S., Beijersbergen, R. L., and Bernards, R. (2004). A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* **428,** 431–437.

Betz, N. (2003). Produce functional siRNAs and hairpin siRNAs using the T7 RiboMAX express RNAi system. *Promega Notes* **85,** 15–18.

Brummelkamp, T. R., Bernards, R., and Agami, R. (2002). A system for stable expression of short interfering RNAs in mammalian cells. *Science* **296,** 550–553.

Chanda, S. K., and Caldwell, J. S. (2003). Fulfilling the promise: Drug discovery in the post-genomic era. *Drug Disc. Today* **8,** 168–174.

Chanda, S. K., White, S., Orth, A. P., Reisdorph, R., Miraglia, L., Thomas, R. S., DeJesus, P., Mason, D. E., Huang, Q., Vega, R., Yu, D. H., Nelson, C. G., Smith, B. M., Terry, R., Linford, A. S., Yu, Y., Chirn, G. W., Song, C., Labow, M. A., Cohen, D., King, F. J., Peters, E. C., Schultz, P. G., Vogt, P. K., Hogenesch, J. B., and Caldwell, J. S. (2003). Genome-scale functional profiling of the mammalian AP-1 signaling pathway. *Proc. Natl. Acad. Sci. USA* **100,** 12153–12158.

Cho, C. Y., Koo, S. H., Wang, Y., Callaway, S., Hedrick, S., Mak, P. A., Orth, A. P., Peters, E. C., Saez, E., Montminy, M., Schultz, P. G., and Chanda, S. K. (2006). Identification of the tyrosine phosphatase PTP-MEG2 as an antagnist of hepatic insulin signaling. *cell. Metab.* **3,** 367–378.

Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411,** 494–498.

Hannon, G. J., and Rossi, J. J. (2004). Unlocking the potential of the human genome with RNA interference. *Nature* **431,** 371–378.

Harada, J. N., Bower, K. E., Orth, A. P., Callaway, S., Nelson, C. G., Laris, C., Hogenesch, J. B., Vogt, P. K., and Chanda, S. K. (2005). Identification of novel mammalian growth regulatory factors by genome-scale quantitative image analysis. *Genome Res.* **15,** 1136–1144.

Jackson, A. L., and Linsley, P. S. (2004). Noise amidst the silence: Off-target effects of siRNAs? *Trends Genet.* **20,** 521–524.

Kittler, R., Putz, G., Pelletier, L., Poser, I., Heninger, A. K., Drechsel, D., Fischer, S., Konstantinova, I., Habermann, B., Grabner, H., Yaspo, M. L., Himmelbauer, H., Korn, B., Neugebauer, K., Pisabarro, M. T., and Buchholz, F. (2004). An endoribonuclease-prepared

siRNA screen in human cells identifies genes essential for cell division. *Nature* **432,** 1036–1040.

Krieg, N., and Holt, J. (1984). ''Bergey's Manual of Systematic Bacteriology.'' Williams & Wilkens, Baltimore, MD.

Li, T., Chang, C. Y., Jin, D. Y., Lin, P. J., Khvorova, A., and Stafford, D. W. (2004). Identification of the gene for vitamin K epoxide reductase. *Nature* **427,** 541–544.

Luo, B., Heard, A. D., and Lodish, H. F. (2004). Small interfering RNA production by enzymatic engineering of DNA (SPEED). *Proc. Natl. Acad. Sci. USA* **101,** 5494–5499.

Paddison, P. J., Silva, J. M., Conklin, D. S., Schlabach, M., Li, M., Aruleba, S., Balija, V., O'Shaughnessy, A., Gnoj, L., Scobie, K., Chang, K., Westbrook, T., Cleary, M., Sachidanandam, R., McCombie, W. R., Elledge, S. J., and Hannon, G. J. (2004). A resource for large-scale RNA-interference-based screens in mammals. *Nature* **428,** 427–431.

Pennisi, E. (2005). Why do humans have so few genes? *Science* **309,** 80.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). ''Molecular Cloning, a Laboratory Manual.'' Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sen, G., Wehrman, T. S., Myers, J. W., and Blau, H. M. (2004). Restriction enzyme-generated siRNA (REGS) vectors and libraries. *Nature Genet.* **36,** 183–189.

Stegmeier, F., Hu, G., Rickles, R. J., Hannon, G. J., and Elledge, S. J. (2005). A lentiviral microRNA-based system for single-copy polymerase II-regulated RNA interference in mammalian cells. *Proc. Natl. Acad. Sci. USA* **102,** 13212–13217.

Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G., Klausner, R. D., Collins, F. S., Wagner, L., Shenmen, C. M., Schuler, G. D., Altschul, S. F., Zeeberg, B., Buetow, K. H., Schaefer, C. F., Bhat, N. K., Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A. A., Rubin, G. M., Hong, L., Stapleton, M., Soares, M. B., Bonaldo, M. F., Casavant, T. L., Scheetz, T. E., Brownstein, M. J., Usdin, T. B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S. S., Loquellano, N. A., Peters, G. J., Abramson, R. D., Mullahy, S. J., Bosak, S. A., McEwan, P. J., McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S., Worley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulyk, S. W., Villalon, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A., Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Young, A. C., Shevchenko, Y., Bouffard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M., Butterfield, Y. S., Krzywinski, M. I., Skalska, U., Smailus, D. E., Schnerch, A., Schein, J. E., Jones, S. J., and Marra, M. A. Mammalian Gene Collection Program (2002). Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc. Natl. Acad. Sci. USA* **99,** 16899–16903.

Strausberg, R. L., Feingold, E. A., Klausner, R. D., and Collins, F. S. (1999). The mammalian gene collection. *Science* **286,** 455–457.

Su, A. I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K. A., Block, D., Zhang, J., Soden, R., Hayakawa, M., Kreiman, G., Cooke, M. P., Walker, J. R., and Hogenesch, J. B. (2004). A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl. Acad. Sci. USA* **101,** 6062–6067.

Tiscornia, G., Singer, O., Ikawa, M., and Verma, I. M. (2003). A general method for gene knockdown in mice by using lentiviral vectors expressing small interfering RNA. *Proc. Natl. Acad. Sci. USA* **100,** 1844–1848.

Xia, H., Mao, Q., Paulson, H. L., and Davidson, B. L. (2002). siRNA-mediated gene silencing *in vitro* and *in vivo*. *Nature Biotechnol.* **20,** 1006–1010.

Ziauddin, J., and Sabatini, D. M. (2001). Microarrays of cells expressing defined cDNAs. *Nature* **411,** 107–110.